

COMPARISON OF GENERATIVE ADVERSARIAL NETWORKS AND DIFFUSION MODELS FOR SYNTHETIC STEADY-STATE VISUAL EVOKED POTENTIAL SIGNAL GENERATION

Przemysław Wiszniewski, Marcin Kołodziej, Andrzej Majkowski

Warsaw University of Technology, Faculty of Electrical Engineering, ul. Koszykowa 75, 00-662 Warsaw, Poland
(✉ marcin.kolodziej@pw.edu.pl)

Abstract

Steady-state visual evoked potential (SSVEP) signals are widely used in brain–computer interface (BCI) applications. Artificial generation of SSVEP signals can reduce the time required for experimental data collection, thereby minimizing user fatigue and improving comfort during electroencephalographic recordings. Synthetic SSVEP data enable data augmentation, training and validation of BCI classification algorithms, and benchmarking of SSVEP-based BCI systems without extensive subject-specific recordings. However, realistic modelling and generation of SSVEP signals remain challenging due to their nonlinear and non-stationary characteristics and the presence of harmonic components. In this study, we present a comparative analysis of two approaches to synthetic SSVEP signal generation: generative adversarial networks and diffusion models. The performance of both methods is evaluated using qualitative and quantitative metrics to assess the fidelity of the generated signals in both the time and frequency domains. The evaluation framework includes the frequency-domain Euclidean distance, Wasserstein distance, signal-to-noise ratio (SNR), and principal component analysis (PCA). The results indicate that diffusion models produce synthetic SSVEP signals with higher fidelity in both the time and frequency domains than the evaluated GAN configuration. Overall, diffusion models achieved better agreement with real data according to the adopted metrics while preserving key signal characteristics.

Keywords: SSVEP, EEG, BCI, diffusion models, GANs, data augmentation.

1. Introduction

Steady State Visual Evoked Potentials (SSVEP) are widely used signals in *Brain–Computer Interfaces* (BCI) [1, 2] due to their high *information transfer rate* (ITR) and non-invasive nature as an electroencephalographic-based method [3]. Collecting *electroencephalographic* (EEG) signals is costly, time-consuming, and requires careful experimental preparation and monitoring [4]. Generating synthetic SSVEP data can potentially improve the performance and reliability of SSVEP-based BCIs. Augmenting training datasets with synthetic SSVEP signals can enhance the performance of SSVEP classifiers [5]. Synthetic signals can also reduce

the time-consuming calibration required for each subject and help minimize fatigue during experiments [6, 7]. Additionally, synthetic data can help mitigate privacy issues associated with sharing EEG recordings, facilitating broader data dissemination and collaboration [8, 9]. However, modelling SSVEP signals is challenging due to their nonlinear characteristics [10] and the presence of harmonic and sub-harmonic components in EEG signals caused by rhythmic stimulation [11, 12]. Moreover, SSVEP signals exhibit non-stationarity, with amplitude and energy fluctuating over time [13]. Another challenge is the reliable assessment of the quality of generated data and the development of robust validation methods [14].

From the perspective of metrology and measurement systems, accurate acquisition of EEG signals – particularly SSVEP – remains a significant challenge. EEG signals are extremely weak, typically in the microvolt range, and highly susceptible to various sources of interference, which complicates reliable measurement. Measurement uncertainty arises from factors such as variability in electrode placement, differences in scalp conductivity, environmental noise, and individual physiological variability. Furthermore, EEG acquisition requires specialized equipment and controlled experimental conditions. Collecting data from large groups of participants also involves substantial organizational and ethical requirements, including obtaining informed consent, securing approval from bioethics committees, and ensuring appropriate data protection. Consequently, large, high-quality, and standardized EEG datasets remain limited and are not widely available.

In this study, we investigate the generation of synthetic SSVEP signals using generative adversarial networks and diffusion models and provide a comparative evaluation of both approaches using time- and frequency-domain criteria.

2. Related articles

Generative Adversarial Networks (GANs) are a class of deep learning models [15] first introduced in 2014 [16]. In its basic form a GAN consists of two components: a generator and a discriminator, which compete in a minimax game [17, 18]. The generator maps noise, often sampled from normal distribution, to generate data close to real examples. Both components are trained simultaneously, with the generator improving its ability to mimic real data and the discriminator enhancing its ability to distinguish between real and generated data. To address these issues, several extensions have been proposed [19]. In *Wasserstein* GANs (WGANs) the objective function was modified by replacing the Jensen-Shannon divergence with the Wasserstein distance, further stabilizing training [20, 21]. GANs have been used to reconstruct high-resolution signals from low-resolution data acquired with low-cost equipment in [22]. This approach incorporates Temporal-Spatial-Frequency Loss, computing loss across time-series, common spatial patterns and power spectral density features. MIEEG-GAN is a model for generating artificial EEG signals for motor imagery tasks [23]. Deep convolutional GANs have also been adopted for EEG augmentation [24]. It uses various up- and down-sampling techniques and a multi-stage learning process with increasing data resolution. In [25] authors use GANs to augment EEG data and compare classification performance. The study found that GAN-generated data was most beneficial for smaller datasets. Another approach with EEGGAN-Net [26] combines *Conditional* GANs (CGANs) with the *Squeeze-and-Excitation* (SE) attention mechanism to enhance EEG signal classification, particularly for motor imager tasks in BCI Competition IV dataset.

Diffusion models (DM) were introduced in 2015 as generative models inspired by non-equilibrium thermodynamics [27]. The concept is to gradually add noise to the data during the forward process and learn the reverse process to reconstruct data from noise. *Diffusion Probabilistic Models* (DPMs), such as *Denoising Diffusion Probabilistic Models* (DDPMs) [28],

which incorporate deep networks to parameterize the inverse process, have achieved state-of-the-art performance in high-quality image synthesis [29,30]. While diffusion models excel at generating realistic data, they require significant computational resources and large datasets to train. *Denoising Diffusion Implicit Models* (DDIMs) are an efficient class of generative models that accelerate the sampling process of DDPMs while maintaining high image generation quality [31].

Diffusion models are increasingly being applied to time series data to forecast [32], refine and generate synthetic data [33]. Several diffusion models have been developed for time-series data. *TimeDiff* [34] is a non-autoregressive model for high-quality time series forecasting with new conditioning mechanisms. *TSDiff* [35] is an unconditional diffusion model with a self-guidance mechanism, useful for forecasting, refinement and synthetic data generation.

In EEG signal generation, diffusion models have been utilized for tasks such as noise reduction [36] and data augmentation [37]. DPMs have been successfully applied to generate high-quality EEG signals, including visually evoked potentials [38,39] and motor imagination signals from BCI competition datasets [40]. The researchers have used techniques such as implicit sampling and progressive distillation to speed up inference while using convolutional neural networks without up- and down-sampling layers to prevent frequency artifacts. The importance of appropriate metrics to assess the quality of generated EEG data has been highlighted due to the unique characteristics of EEG compared to image or audio data. Recent advances include brain-based representation learning [41], extending Motor Imagery data for classification tasks [42], and high-fidelity EEG generation using DDPMs with convolutional neural networks and the *Short-time Fourier Transform* (STFT) [43]. BCI modelling with stable diffusion models and generation of visual evoked potentials were also investigated.

For generated EEG signals evaluation, two main approaches are used: comparing generated and real data in both time and frequency domains [44], and comparing data projected into lower-dimensional spaces [45] using such techniques as the *Principal Component Analysis* (PCA) [46,47]. These domain-specific evaluation methods provide more meaningful insights into the quality of synthetic EEG signals compared to metrics designed for image-based tasks.

3. Aim of the Article

Artificially generated steady-state visual evoked potential signals can significantly reduce the time required for experimental data collection, thereby minimizing user fatigue and increasing comfort during electroencephalographic recording procedures. Furthermore, they serve as an effective tool for training data augmentation, supporting the development of classification algorithms in cases where the availability of real electroencephalographic samples is limited. The aim of this study is to assess the quality of synthetic electroencephalographic data generated using two advanced deep learning techniques: GANs and diffusion models. These two approaches represent fundamentally different paradigms for data generation. In diffusion models, the learning process is based on sequential denoising, that is, reconstructing the signal from successive levels of noise, which leads to a stable and well-defined approximation of underlying data distribution. In contrast, GAN-based generation relies on adversarial competition between two networks – a generator and a discriminator – which often results in an unstable training process. To the best of the authors' knowledge, no prior study has compared these two methods in the context of SSVEP signal generation, highlighting the novelty and originality of the present work.

The comparison includes both quantitative and qualitative aspects of the generated signals. The main contributions of this article are as follows:

- Comparative analysis of generative approaches: We conduct a detailed comparison of GANs and diffusion models for synthetic steady-state visual evoked potential signal generation.

- Multi-domain evaluation framework: We assess the quality of generated signals using a set of quantitative and qualitative metrics in both the time and frequency domains.
- Computational comparison: We report training and inference costs of both approaches to contextualize the quality–efficiency trade-off.

4. Materials

SSVEP responses are typically strongest at stimulation frequencies around 10 Hz. To investigate responses near this frequency range, we selected a dataset including stimuli at 5, 6, 7 and 8 Hz. This study utilized a previously published online dataset of steady-state visual evoked potentials [48]. The dataset included recordings from 5 participants aged 23, 25, 31, 42, and 46. The EEG data was recorded at a sampling rate of 256 Hz using a *g.USBAmp* 2.0 biosignal amplifier (*g.tec*) with 24-bit A/D converters. Visual stimulation was provided by a green LED (1 cm diameter) placed 1 m from the participants’ eyes. The LED was driven by a Siglent SDG1062X function generator with frequencies of 5, 6, 7, or 8 Hz. EEG signals were recorded from 16 electrodes placed according to the international 10–20 system.

The data were pre-processed using a Butterworth bandpass filter (0.1–100 Hz) to remove low-frequency drifts and a notch filter (48–51 Hz) to eliminate power line interference. We chose the Oz electrode for our study because it is widely recognized as giving strong SSVEP response [49, 50]. For each SSVEP stimulation frequency (5, 6, 7, and 8 Hz), the first 20 seconds of the 30-second recordings were used for training purposes. The remaining 10 seconds of each recording were reserved exclusively for testing. The EEG signals were segmented into 1-second epochs (256 samples) with an overlap of 245 samples. For a single subject, this resulted in 348 training examples and 165 test examples per class. In total, 1392 training examples and 660 test examples were obtained across all four classes. The use of overlapping windows increased temporal resolution, enabling a more fine-grained analysis of signal dynamics while reducing the impact of discontinuities that may occur at window boundaries. Prior to model training, all signals were normalized to the range $[-1, 1]$ to ensure consistency across different EEG recordings. This normalization procedure improved the models’ ability to accurately capture salient signal characteristics. Empirical observations showed that training both GANs and DMs on normalized data led to more stable training behaviour and, in many cases, faster convergence.

Models were trained separately for each subject using the full training set containing 348 examples per class. Subsequently, 1000 new synthetic samples per class were generated. The generated data were then adjusted to match the mean and variance of the original dataset for the corresponding subject, ensuring statistical consistency between synthetic and real samples. Test EEG data were not used at any stage in the training of either the GAN or DM generative models. This strict separation prevents data leakage and ensures that the evaluation of the generated signals faithfully reflects the true generalization capability of the models.

5. Methods

This study used both diffusion models and generative adversarial networks (GANs). These two methodologies are crucial in generative tasks such as generating images and time series data. The use of noise in both diffusion models and GANs is an integral part of them, although they use it differently. Below is a detailed description of the architectures and learning processes used for both networks in this study. The dataset used in this study, *Synthetic SSVEP Dataset Generated with GAN and Diffusion Models*, was created by the authors and is publicly available in the Zenodo repository (<https://doi.org/10.5281/zenodo.18874308>, accessed on 5 March 2026).

5.1. Diffusion models

Diffusion models are a type of machine learning framework that transforms data in a two-stage process. In the forward diffusion phase, structured data becomes increasingly noisy over time. The reverse diffusion phase then reconstructs the original data by systematically removing noise, based on learned probabilities. Visualization of the time steps for sinusoidal data highlighting the gradual degradation and subsequent recovery of the data structure is presented in Fig. 1. The first step in diffusion models is the diffusion process, which systematically transforms clean data into clean noise through a series of steps. The noise is introduced gradually perturbing the original data, which is modelled as a Markov chain. The data at each timestep x_t depends only on the previous step x_{t-1} .

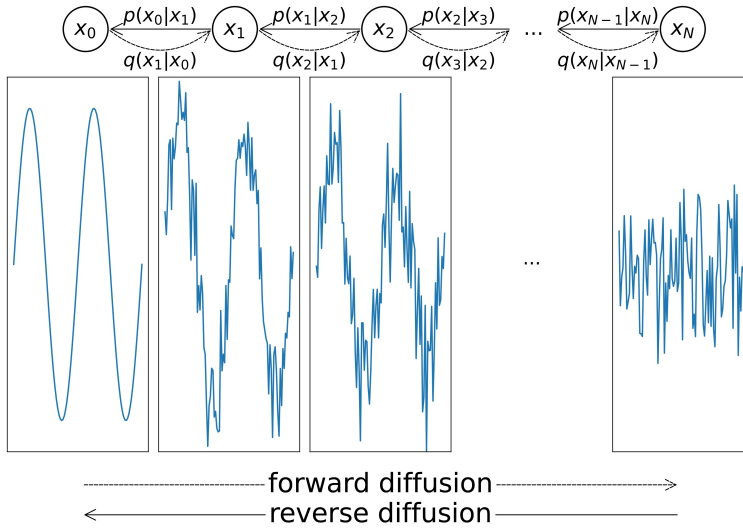


Fig. 1. Illustration of forward and reverse diffusion processes in a diffusion model.

The forward process is defined as follows:

$$q(x_t | x_{t-1}) = N\left(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I\right), \quad (1)$$

where: t – represents the timestep, β_t – noise variance schedule, $N(\mu, \Sigma)$ – Gaussian distribution with mean μ and covariance Σ .

The noise schedule determines how much noise is added at each time step. In this implementation there was a cosine schedule used to ensure smooth transitions between time steps.

Later, during the reverse process, the aim is to reconstruct the data by gradually eliminating noise in an iterative manner, effectively reversing the transformations applied during the forward diffusion phase in a learnable way. The reverse process learns to de-noise the data by predicting added noise or other parameterizations of the original data. The goal is to learn a transition function that reverses this process, moving from a Gaussian distribution back to the original data distribution. The inverse process is also modelled as a Markov chain:

$$p_\theta(x_{t-1} | x_t) = N\left(x_{t-1}; \mu_\theta(x_t, t), \sigma_t^2 I\right), \quad (2)$$

where: $\mu_\theta(x_t, t)$ – predicted mean, σ_t^2 – predicted or fixed variance, θ – are the model parameters learned during training.

As the backbone for the diffusion model, a modified U-Net network was used, using a hierarchical structure to effectively denoise the data over multiple time steps (Table 1). U-Net is an architecture widely used in diffusion models due to its symmetric encoder-decoder structure and the use of skip connections. These skip connections are crucial for preserving spatial information and enabling efficient back propagation, as they directly connect the relevant layers in the encoder and decoder paths [51]. The U-Net followed a classical encoder-decoder structure with skipped connections, adapted to temporal signals. It begins by embedding input sequences into higher-dimensional feature representations using convolutional layers and sinusoidal positional embeddings as proposed in [52], which inject temporal context into models for time conditioning. The encoder progressively down-samples the input data through multiple levels of resolution, each using residual blocks and linear attention mechanisms to effectively capture local and global dependencies. At the bottleneck, the network integrates multi-head attention to model complex temporal dependencies. The decoder then reconstructs the original sequence resolution by up-sampling and combining encoder features using skip connections, ending with a final convolution layer that maps features back to the original data dimensionality.

Table 1. Architecture components of the 1D U-Net model within the diffusion model.

Component	Modules	Details
Initial projection	Conv1d (kernel=7, padding=3)	256 samples $\rightarrow$ 64 features
Time Embedding	SinusoidalPosEmb, Linear, GELU	Scalar $\rightarrow$ 256-dim vector
Downsampling (Level 1)	2×ResnetBlock, Linear Attention, Downsample	Feature dim: 64
Downsampling (Level 2)	2×ResnetBlock, Linear Attention, Downsample	Feature dim: 128
Downsampling (Level 3)	2×ResnetBlock, Linear Attention, Downsample	Feature dim: 256
Downsampling (Level 4)	2×ResnetBlock, Linear Attention, Downsample	Feature dim: 512
Bottleneck	ResnetBlock, Multi-head Attention, ResnetBlock	Feature dim: 512
Upsampling (Level 4)	2×ResnetBlock, Linear Attention, Upsample	Feature dim: 512 $\rightarrow$ 256
Upsampling (Level 3)	2×ResnetBlock, Linear Attention, Upsample	Feature dim: 256 $\rightarrow$ 128
Upsampling (Level 2)	2×ResnetBlock, Linear Attention, Upsample	Feature dim: 128 $\rightarrow$ 64
Upsampling (Level 1)	2×ResnetBlock, Linear Attention, Conv1d	Feature dim: 64
Output Projection	ResnetBlock, Conv1d	64 features $\rightarrow$ 256 samples

In this work, we used a DDPM for one-dimensional data. The forward process involved gradually adding Gaussian noise to the input data over a fixed number of 50000 training steps. The training was performed using the Adam optimizer with a learning rate of $1e^{-4}$ and beta parameters set to (0.9, 0.99).

The loss function was based on mean squared error between the model’s prediction and target using v -parameterization:

$$\mathcal{L} = E_{t, x_t, \epsilon} \left[w(t) \cdot \|v - v_{\theta}(x_t, t)\|^2 \right], \quad (3)$$

where: $SNR = \frac{\bar{\alpha}_t}{1-\bar{\alpha}_t}$ - signal-to-noise ratio, $w(t) = \frac{SNR}{SNR+1}$ - time-dependent weighting function, with SNR being the signal-to-noise ratio, $v = \sqrt{\bar{\alpha}_t}\epsilon - \sqrt{1-\bar{\alpha}_t}x_0$, $\epsilon \sim N(0, I)$ is Gaussian noise, $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$ - cumulative product of alpha values up to the timestep t , $\alpha_t = 1 - \beta_t$ - alpha values derived from the noise schedule.

In our implementation, we used the v - parameterization objective [53], where the model predicts a parameterization that combines aspects of both noise prediction and data prediction.

5.2. Generative Adversarial Networks

The GAN framework consists of two competing neural networks – a generator and a discriminator – trained simultaneously in an adversarial learning process (Fig. 2). As part of this study, a systematic analysis and comparison of multiple GAN architectures were conducted to identify the model that produces the highest-quality synthetic data. Both classical approaches and advanced variants were considered, including Basic GAN (binary cross-entropy loss), LSGAN (mean squared error loss), WGAN with a Lipschitz constraint enforced via weight clipping, WGAN-GP with gradient penalty, conditional architectures (WGAN-Cond), progressive models (WGAN-Progressive), approaches incorporating frequency-domain losses (WGAN-FFT), as well as models specifically designed for temporal and neurophysiological data, such as TimeGAN, EEG-GAN, and MIEEG-GAN. For each architecture, extensive experiments were carried out involving the selection and evaluation of training hyperparameters, loss functions, regularization strategies, and training stability. The process of selecting the optimal model was time-consuming and required multiple iterations of training and validation to assess the quality of the generated signals. Nevertheless, this comprehensive evaluation enabled the identification of the architecture best suited to the characteristics of the analysed data. Among the evaluated GAN variants, Basic GAN provided the best overall performance, including the highest classification accuracy within the GAN group; therefore, it was selected as the representative GAN baseline for comparison with diffusion models.

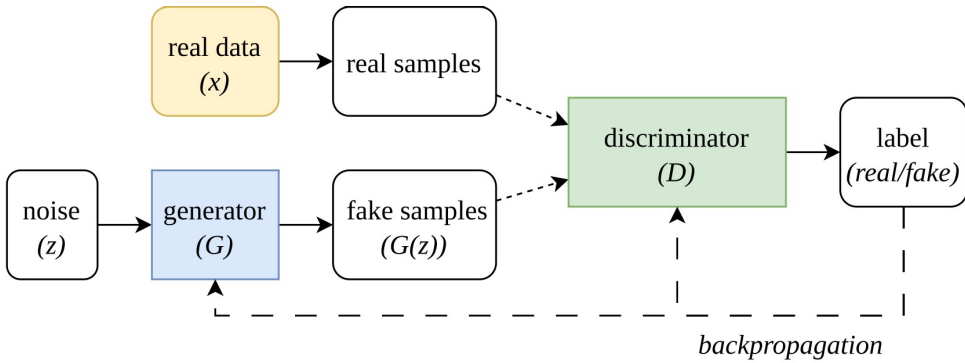


Fig. 2. Schematic representation of the Basic GAN architecture.

The generator network aims to synthesize artificial data samples that closely resemble actual data distribution, while the discriminator network acts as a binary classifier that distinguishes between real and fake samples. The training procedure involves a minimax game in which the discriminator aims to maximize the classification accuracy, while the generator attempts to minimize it. This adversarial relationship is formulated mathematically using the following loss functions:

$$L_{G_{\text{time}}} = -\frac{1}{n} \sum_{i=1}^n \log (D (G (z_i))), \quad (4)$$

$$L_{D_{\text{time}}} = -\frac{1}{n} \sum_{i=1}^n [\log (D (x_i)) + \log (1 - D (G (z_i)))] , \quad (5)$$

where: $D(\cdot)$ represents the discriminator, $G(\cdot)$ represents the generator, x_i – real data samples, z_i – random noise inputs, n - number of examples in the batch.

We have introduced additional spectral losses for the generator and discriminator to efficiently generate SSVEP signals that exhibit characteristic frequency patterns in response to repeated visual stimuli, calculated as follows:

$$L_{G_{\text{spectral}}} = L_{D_{\text{spectral}}} = \frac{1}{n} \sum_{i=1}^n |F(x_i) - F(G(z_i))|^2, \quad (6)$$

where $F(\cdot)$ represents the Fourier transform operation, capturing the characteristics in the frequency domain of both the real and generated signals.

To balance the characteristics in the time and frequency domains, we formulate the final loss functions as weighted combinations:

$$L_G = 0.1 \cdot L_{G_{\text{time}}} + 0.9 \cdot L_{G_{\text{spectral}}}, \quad (7)$$

$$L_D = 0.1 \cdot L_{D_{\text{time}}} + 0.9 \cdot L_{D_{\text{spectral}}}. \quad (8)$$

This weighted approach prioritizes spectral fidelity (90%) while maintaining temporal coherence (10%), which is consistent with the frequency-sensitive nature of SSVEP signals.

The GAN training process utilized a generator and discriminator network, with a latent dimension of 20 for input noise. The model was trained for 3000 epochs using a mini-batch size of 8 and a learning rate of 0.00005. The Adam optimizer was employed with a gradient decay factor of 0.7 and a squared gradient decay factor of 0.999. The neural network architecture used in this study was selected through an empirical optimization process, guided by performance metrics across multiple configurations.

The generator network used an architecture designed to transform random noise into realistic SSVEP patterns (Table 2). The network starts with a 20-dimensional noise input vector, processes it through a hidden, fully connected layer with 256 neurons and ReLU activation.

Table 2. Architecture of the generator used in the GAN model.

Layer	Type	Output Shape	Activation
Input	Feature Input	20	None
Hidden	Fully Connected	256	ReLU
Output	Fully Connected	256	Tanh

The discriminator consists of a deeper network structure to effectively distinguish authentic and synthetic SSVEP signals (Table 3). It processes 256-dimensional input signals through two hidden fully-connected layers with ReLU activations, gradually reducing the dimensionality to extract discriminative features.

Table 3. Architecture of the discriminator used in the GAN model.

Layer	Type	Output Shape	Activation
Input	Feature Input	256	None
Hidden 1	Fully Connected	256	ReLU
Hidden 2	Fully Connected	64	ReLU
Output	Fully Connected	1	Sigmoid

6. Results and discussion

In this study, we compared the performance of GANs and DMs in generating artificial EEG data. Both models were trained to generate 1000 examples per class.

6.1. Evaluation in the Time Domain

The data generated by both GANs and diffusion models were compared with the original data used for training. A visual inspection of the sample data for the 5 Hz stimuli illustrated in Fig. 3 shows the similarities between the generated and original data in the time domain.

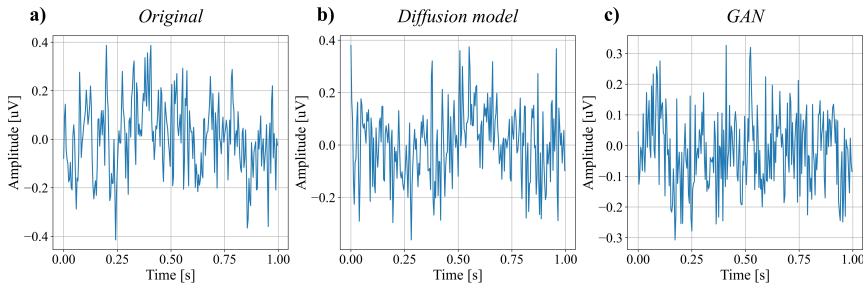


Fig. 3. Comparison of single examples of original SSVEP data (left), diffusion model-generated data (middle), and GAN-generated data (right) in the time domain. The plots show amplitude variations over a 1-second interval, illustrating the temporal characteristics of the signals for each dataset.

The distributions of amplitude values for the original and generated SSVEP data for four stimulation frequencies are compared in Fig. 4. The distribution analysis showed that the diffusion models generated data more closely matched the original distribution, while the GAN-generated data showed a smoother and shifted distribution compared to the original data, suggesting less precise replication of amplitude characteristics.

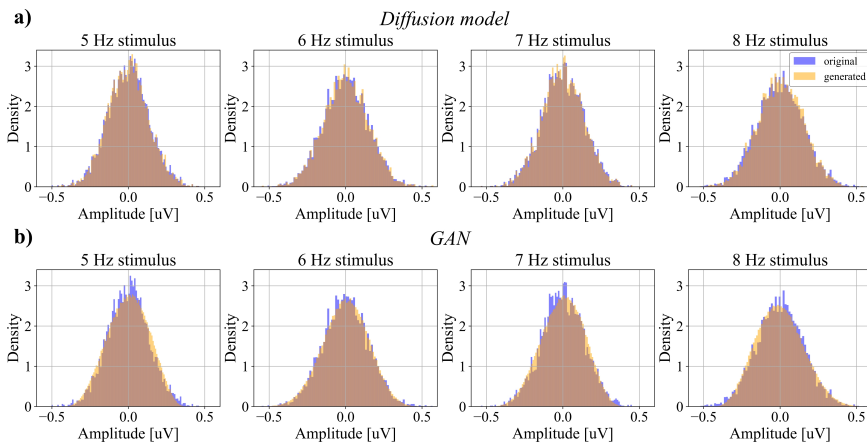


Fig. 4. Distribution of amplitudes for original (blue) and generated (orange) SSVEP data for the diffusion model (top row) and GAN (bottom row) across stimulus frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz). Each panel shows the density of amplitude values, highlighting the similarity in distribution between original and generated data for each model.

6.2. Evaluation in the Frequency Domain

A spectral analysis of the generated EEG waveforms was conducted. A comparison of the mean Fourier transform spectra of original SSVEP data and the data generated by the diffusion model and GAN across four stimulus frequencies is presented in Fig. 5. The diffusion model shows excellent performance in maintaining spectral fidelity at all frequencies of the stimulus. The GAN shows reduced accuracy in replicating the higher frequency components. This suggests that the diffusion model is more effective in capturing key features of SSVEP signals in the frequency domain.

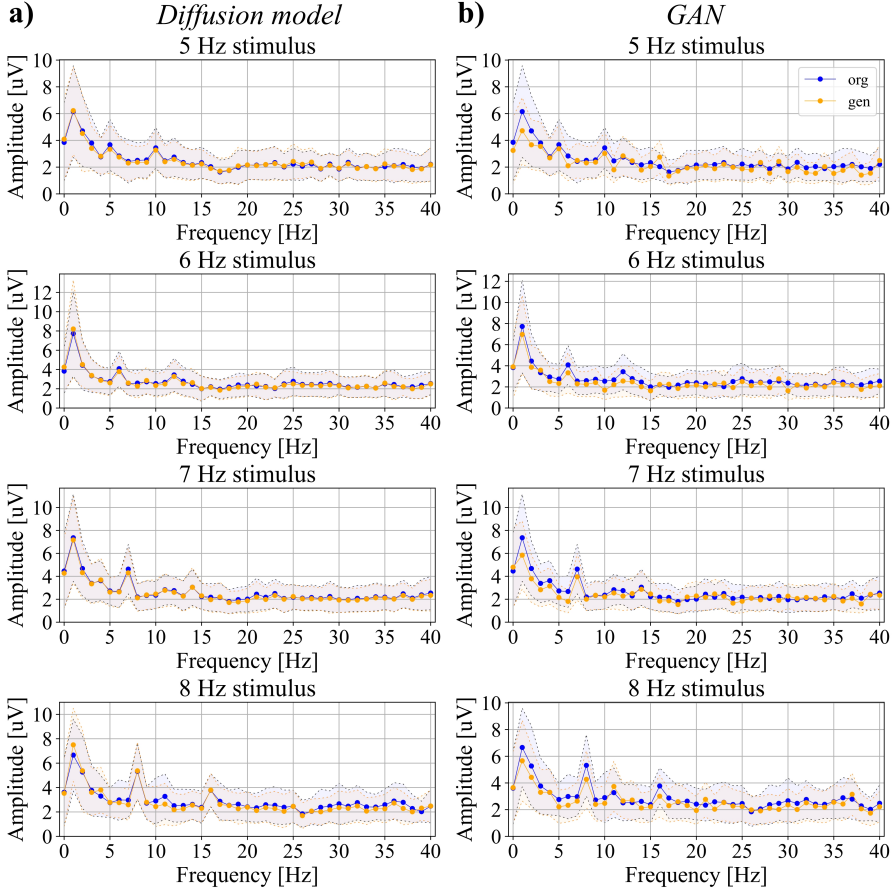


Fig. 5. Fourier transform spectra (mean and standard deviation) of original SSVEP data (train org) and generated data (gen) for the diffusion model (left panels) and GAN (right panels) across stimulus frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz). Each row corresponds to a specific stimulus frequency, showing amplitude variations in the frequency range of 0–40 Hz.

In the context of SSVEP signal analysis, the spectral components up to 40 Hz represent the most significant frequency band. It captures the fundamental frequencies and up to the 5th harmonic in our experimental paradigm. The spectra generated by the diffusion model are very close to the original spectra, with peaks at the corresponding stimulation frequency. The standard deviation bands for the diffusion model indicate a more consistent reproduction of the spectrum across samples. The spectra for the GAN-generated data also show peaks at the corresponding stimulation frequency but with less precise amplitude alignment compared to the diffusion model.

6.3. Evaluation Using Quantitative Similarity Metrics

To objectively evaluate the quality of the generated SSVEP data, we used three different metrics: Euclidean distance of the Fourier transform, Wasserstein distance, and SNR. These metrics provide complementary insights into the similarity of the generated and original data, taking into account different aspects of data fidelity.

Euclidean distance was chosen as the primary metric for quantitative comparison because of its well-established utility in signal processing applications. Given that steady-state visual evoked potentials are essentially frequency-domain phenomena, characterized by characteristic peaks at the stimulation frequency and corresponding harmonics, we implemented measurements of the Euclidean distance between representations of signals in the frequency domain. First, we applied the Fourier transform to convert the time-domain signal into its frequency-domain representation. Next, we computed the mean spectral magnitude within the 0–40 Hz range. Finally, we assessed the difference between the generated and reference signals by calculating the Euclidean distance between their average mean spectral magnitudes obtained in the previous step.

In SSVEP analysis, the Fourier transform is crucial for identifying the frequency components of the signal. By calculating the Euclidean distance between the Fourier transforms of the original and generated data, it is possible to assess how well the generated data reproduce the spectral features of the original data.

Table 4 shows a comparison of the Euclidean distances between the original SSVEP data and the data generated by two models, DM and GAN. Euclidean distances are listed for each of the five tested subjects at four different stimulus frequencies. The diffusion model shows a greater consistency and yields lower Euclidean distances (on average 1.4) across frequencies and subjects. For each subject, the Euclidean distances vary according to frequency, with GAN generally showing larger distances (on average 3.5) compared to the diffusion model.

Table 4. Comparison of Euclidean distances between original and generated SSVEP data for the DM and GAN across different stimulus frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz) for five subjects. The final column shows the overall mean Euclidean distance for each model.

	5Hz		6Hz		7Hz		8Hz		Mean	
	DM	GAN	DM	GAN	DM	GAN	DM	GAN	DM	GAN
Subject 1	1.13	3.59	1.39	3.51	1.09	3.35	1.59	4.28	1.30	3.68
Subject 2	1.42	2.44	1.38	1.82	1.60	2.36	1.32	2.78	1.43	2.35
Subject 3	1.37	4.65	1.45	5.06	2.34	6.76	1.53	6.14	1.67	5.65
Subject 4	1.36	3.22	1.39	3.58	1.38	3.40	2.07	2.89	1.55	3.27
Subject 5	0.88	2.80	1.12	2.61	0.88	2.74	1.91	2.79	1.20	2.74
Mean									1.43	3.54

To assess the similarity in terms of probability distributions, the Wasserstein metric was used. This measure quantifies the difference between the distributions of generated and original data distributions by calculating the cost of transforming one distribution into another. The Wasserstein distance provides a comprehensive picture how well the generative models capture the distribution of the underlying data. This is particularly useful for assessing the overall shape and spread of data distributions, offering insights into the fidelity of the generated data. Table 5 presents a comparison of Wasserstein distances between original SSVEP data and data generated by two models: the

diffusion model and the GAN. The Wasserstein distances are listed for each of five subjects across four different stimulus frequencies. The diffusion model generally shows consistent and low Wasserstein distances (on average 0.0057), indicating a good fit with the original data in terms of similarity of distribution. The GAN shows slightly higher Wasserstein distances (on average 0.0067), suggesting that it generates data distributions that are slightly less similar to the original data.

Table 5. Comparison of Wasserstein distances between original and generated SSVEP data for DM and GAN across different stimulus frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz) for five subjects. The final column shows the overall mean Wasserstein distance for each model. All values are multiplied by 1000 for clarity.

	5Hz		6Hz		7Hz		8Hz		Mean	
	DM	GAN	DM	GAN	DM	GAN	DM	GAN	DM	GAN
Subject 1	3.0	7.5	4.6	5.1	3.8	7.8	10.3	6.1	5.4	6.6
Subject 2	6.7	4.0	5.8	4.4	7.4	3.6	5.9	4.7	6.5	4.1
Subject 3	11.7	11.4	2.1	11.5	6.0	12.1	10.0	10.1	7.4	11.3
Subject 4	4.9	3.2	3.2	7.0	5.8	8.4	8.8	5.4	5.7	6.0
Subject 5	2.8	6.3	1.8	3.5	3.2	5.3	5.8	6.7	3.4	5.4
Mean									1.43	3.54

Particularly relevant for EEG data, the SNR measures the ratio of signal power to noise power within a specified frequency range, providing an indication of signal clarity. In SSVEP analysis, SNR is a critical metric for assessing signal strength and reliability. By evaluating the SNR of both original and generated data, it is possible to determine how effectively the synthetic signals reproduce relevant signal features while minimizing noise, which is essential for accurate classification and analysis in BCI applications.

In this article, SNR estimation was computed as follows:

1. Identifying the amplitude of the Fourier transform at the stimulation frequency.
2. Dividing this amplitude by the average of the amplitudes of two neighbouring frequencies

$$\text{SNR}(f) = \frac{A(f)}{\frac{1}{2} [A(f - \Delta f) + A(f + \Delta f)]}, \quad (9)$$

where f is the stimulation frequency, $A(f)$ is the amplitude of the Fourier transform at a given frequency, and Δf represents the spectral resolution (equal to 1 Hz in our case).

The SNR was computed separately for each stimulation frequency (5, 6, 7, and 8 Hz). For each of these frequencies, the noise was estimated locally using only two neighbouring frequencies, rather than across the entire spectrum.

Table 6 presents a comparison of the mean signal-to-noise ratio for the original SSVEP data, and the data generated by the DM and GAN. Additionally, Figure 6 shows box plots comparing the SNR values for five subjects at stimulation frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz). The mean SNR across subjects is 0.698 for the original data, 0.689 for the diffusion model, and 0.678 for GAN. This indicates that the signals generated by the DM and GAN exhibit a mean SNR very close to that of the real data. It should be noted, however, that in some cases there are slight discrepancies. For example, a clear overestimation is observed at 7 Hz for Subject 1 (1.220 vs. 0.957), whereas an underestimation occurs at 8 Hz for Subject 3 (0.458 vs. 0.620).

Table 6. Mean SNR values for original SSVEP data (real) and data generated by the DM and GAN across different stimulus frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz) for five subjects. The final column shows the overall mean SNR for each model.

	5Hz			6Hz			7Hz			8Hz			Mean		
	real	DM	GAN	real	DM	GAN	real	DM	GAN	real	DM	GAN	real	DM	GAN
Subject 1	0.536	0.560	0.542	0.713	0.587	0.580	0.957	1.013	1.220	0.758	0.738	0.745	0.741	0.724	0.772
Subject 2	0.649	0.628	0.652	0.532	0.577	0.538	0.702	0.654	0.719	0.786	0.736	0.691	0.667	0.648	0.650
Subject 3	0.504	0.512	0.484	0.700	0.715	0.717	0.514	0.521	0.420	0.620	0.625	0.458	0.584	0.593	0.520
Subject 4	0.656	0.638	0.611	0.705	0.686	0.659	0.590	0.548	0.739	0.893	0.834	0.752	0.711	0.676	0.690
Subject 5	0.590	0.584	0.558	0.739	0.736	0.701	0.915	0.859	0.880	0.903	1.032	0.889	0.787	0.803	0.757
Mean													0.698	0.689	0.678

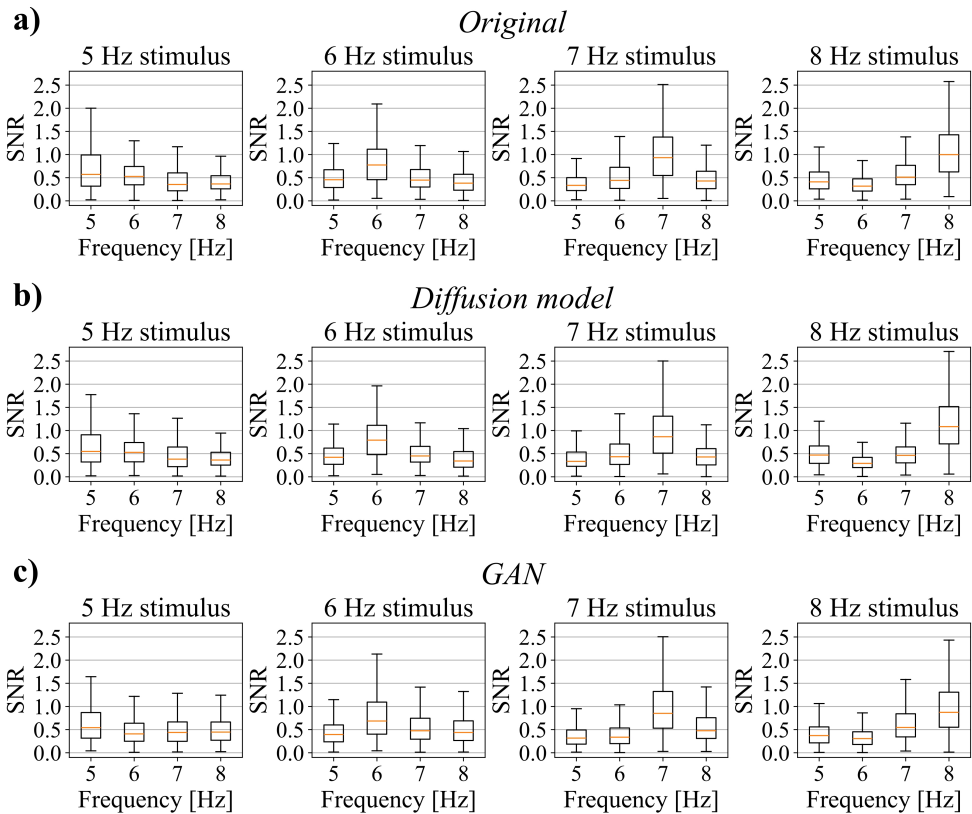


Fig. 6. Box plots of signal-to-noise-ratio (SNR) at stimulation frequencies (5 Hz, 6 Hz, 7 Hz and 8 Hz) for the original SSVEP data (top row), data generated by the diffusion model (middle row) and data generated by GAN (bottom row). Each column corresponds to a specific stimulus frequency (5 Hz, 6 Hz, 7 Hz and 8 Hz).

The highest SNR values are generally observed at 7 Hz and 8 Hz, depending on the subject, indicating that higher stimulation frequencies tend to yield stronger signal dominance over noise. The SNR values obtained from the diffusion model are consistently close to those of the original data across subjects and frequencies, with only minor deviations and a nearly identical overall mean. In contrast, the SNR values of the GAN-generated data exhibit greater variability relative to both the real and diffusion-generated signals, including both overestimations and underestimations at specific frequencies. Nevertheless, the differences in the overall mean SNR values remain small.

6.4. Evaluation Using Principal Component Analysis

PCA was employed as a quantitative and visual tool for assessing the quality of synthetic data with particular emphasis on the agreement between their distribution and that of the real data, as well as on the degree of preserved diversity. PCA enables the projection of high-dimensional data into a lower-dimensional space while retaining the maximum possible variance, thereby allowing the structures of real and generated data distributions to be compared within a common feature space. The PCA transformation was derived exclusively from real training data to ensure direct comparability and to avoid distortions arising from differences in feature representations.

The analysis was performed on spectral data obtained via the Fourier transform in the 0–40 Hz frequency range. By definition, the leading principal components captured the largest proportion of the dataset variance, representing the dominant spectral patterns of signals. The results of the PCA are shown in Fig. 7 as a two-dimensional projection (PC1 vs. PC2) of both real and synthetic data. Each point in the plot corresponds to a single sample projected onto the first two principal components while contour plots illustrate the density of the data distributions. Data generated by the diffusion model exhibited a high degree of agreement with the distribution of the real data while simultaneously preserving increased diversity. These samples occupied regions of feature space that partially extended beyond the original data yet were primarily located between the clusters of real samples, indicating the generation of realistic and diverse signals. In contrast, data generated by the GAN covered a substantially broader area in the PCA space, including points distant from the main distribution of real data. Despite this higher variability, the resulting distribution reflected the structure of the original data to a lesser extent. Consequently, the dispersion observed in the diffusion-model-generated data was assessed positively as an indication of enhanced diversity while maintaining distributional consistency, whereas the broader coverage of the feature space by the GAN was considered less favourable.

6.5. Evaluation Using Classification Performance

One of the key measures of signal quality in brain–computer interface systems is classification accuracy. High classifier performance indicates that the signals contain salient information necessary for correctly recognizing user intent. In this study, classification accuracy was used as an indirect measure of synthetic data quality, understood as the ability of the generated signals to preserve features that are essential for SSVEP signal recognition.

The aim of this analysis was not to evaluate the effectiveness of data augmentation or to optimize the classifier training process, but rather to perform a comparative assessment of the quality of signals generated by different generative models. In particular, the analysis examined the extent to which synthetic signals, generated independently for each subject, retain information that enables accurate classification of real EEG signals. To this end, classification accuracy was compared across three EEG datasets: real signals from the training set, synthetic signals generated using generative adversarial networks, and synthetic signals generated by diffusion

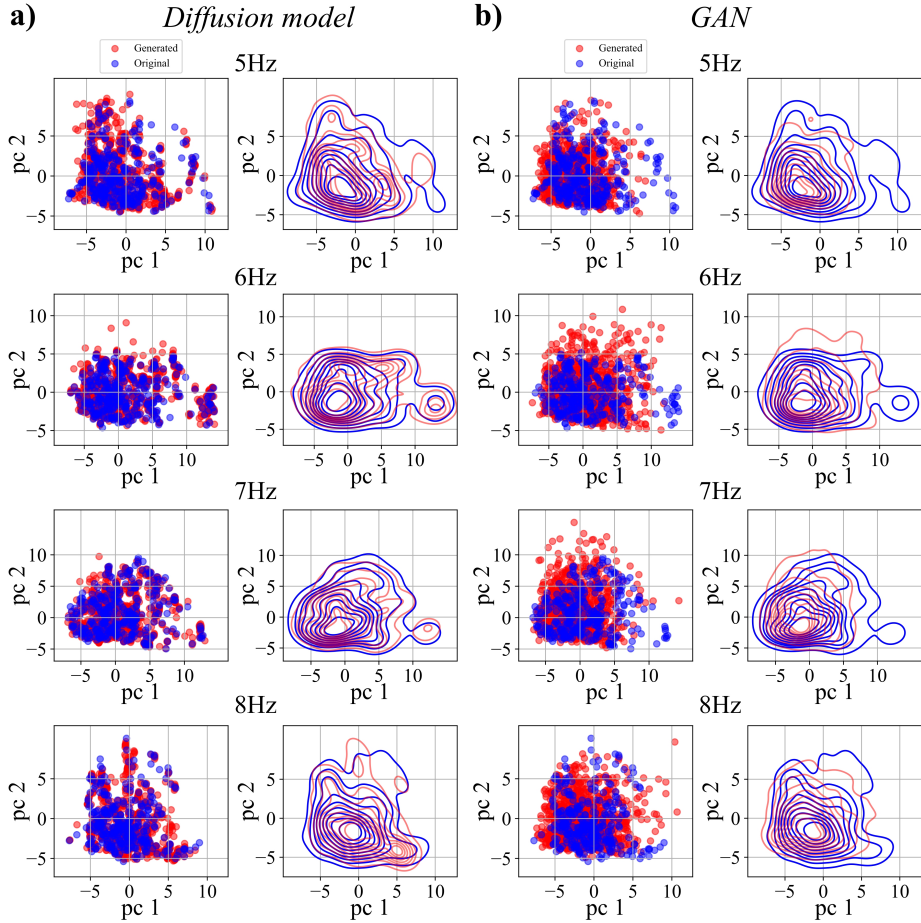


Fig. 7. PCA of the Fourier transform for 0-40Hz frequency range analysis. Scatter and contour plots compare the principal component distributions of original SSVEP data (blue) and generated data (red) for the diffusion model (left panels) and GAN (right panels) across stimulus frequencies (5 Hz, 6 Hz, 7 Hz, and 8 Hz). Each row corresponds to a specific stimulus frequency, while the columns show scatter plots (left) and kernel density estimate (KDE) plots (right).

models. Classifiers were trained separately on real training data, GAN-generated synthetic data, and DM-generated synthetic data. In all cases, testing was performed exclusively on real EEG test data which were not used at any stage in the training of either the generative models or the classifiers. This approach enables a direct assessment of whether synthetic data contain sufficient information to train a classifier that can subsequently generalize to real SSVEP signals. The analysis thus addresses the extent to which data generated by a given model can serve as a substitute for real data from the perspective of the classification task.

Classification was performed using a *support vector machine* (SVM) with a *radial basis function* (RBF) kernel applied to EEG data after the Fourier transformation. The entire frequency spectrum was treated as a feature vector, ensuring consistent comparison conditions across all data types. Based on the obtained results (Table 7), the highest mean classification accuracy was achieved using real signals, while signals generated by diffusion models attained a very similar level of performance.

In contrast, signals generated by GANs exhibited substantially lower classification accuracy. These findings indicate that diffusion models preserve the key spectral features of SSVEP signals required for accurate classification to a greater extent than GANs. It should be emphasized that the presented analysis is comparative in nature and does not constitute a comprehensive evaluation of data augmentation effectiveness. The investigation of other classifiers, alternative feature representations, and different proportions of real and synthetic data is left for future work. Nevertheless, the results suggest that, for the considered SSVEP signal generation task, diffusion models produce data whose quality is closer to that of real data than that achieved by conventional GAN-based approaches.

Table 7. Classification accuracy (in percentage points) for real and synthetic EEG signals generated using DM and GANs. Classification using an SVM with Fourier transformed data.

	Real	DM	GAN
Subject 1	69.4	67.3	41.4
Subject 2	66.2	65.6	35.8
Subject 3	27.7	28.8	33.6
Subject 4	55.3	50.9	40.5
Subject 5	57.4	53.8	42.9
Mean	55.2	53.3	38.8

6.6. Computational Cost Analysis

GAN training took approximately 3.5 hours per subject, while diffusion models training required about 12 hours. Diffusion models have a significantly larger number of parameters (14 853 505) compared to GANs (generator: 71 168, discriminator: 82 305). Both models were trained using an NVIDIA GeForce RTX 4090 GPU and a 12th generation *Intel*^(R) Core^(TM) i7-12700F CPU.

The inference time was 6.08 s for generating 1000 samples using the diffusion model and 0.157 s for the GAN model. This corresponds to per-sample generation times of 6.08 ms and 0.157 ms for the diffusion model and GAN, respectively. The differences in inference time arise directly from the nature of the generation process. In diffusion models, generating a single sample requires iterative denoising over multiple time steps, which substantially increases the computational cost. In contrast, GANs generate data in a single forward pass through the generator network, resulting in a very short inference time.

7. Conclusions

The study revealed a trade-off between data quality and computational requirements. Both diffusion models and GANs were able to successfully generate artificial EEG data to augment the original dataset. However, the EEG signals produced by diffusion models demonstrated superior fidelity and stability compared to those generated by GANs. Specifically, diffusion-based synthetic EEG data better preserved critical frequency characteristics and temporal patterns which are essential for accurate representation of brain activity.

Despite their quality advantages, diffusion models required significantly more computational resources and longer training times than GANs.

These findings suggest that while both methods are viable for EEG data generation, diffusion models offer higher-quality outputs at the expense of greater computational demands. Their stability and ability to capture fine-grained signal details make them especially suitable for generating complex EEG data.

Ultimately, the generation of synthetic SSVEP signals using diffusion models and GANs can reduce the need for extensive data collection, increase user comfort and enhance the performance of BCIs, all without altering the physical experimental setup.

Acknowledgements

Research was funded by Warsaw University of Technology within the Excellence Initiative: Research University (IDUB) program.

Data Availability Statement

The dataset used in this study, *Synthetic SSVEP Dataset Generated with GAN and Diffusion Models*, containing original data, training datasets, and data generated using GAN and diffusion models, was created by the authors and is publicly available in the Zenodo repository (<https://doi.org/10.5281/zenodo.18874308>, accessed on 5 March 2026).

References

- [1] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767–791. [https://doi.org/10.1016/s1388-2457\(02\)00057-3](https://doi.org/10.1016/s1388-2457(02)00057-3)
- [2] Kołodziej, M., Majkowski, A., Oskwarek, Ł., & Rak, R. J. (2016). Comparison of EEG Signal Preprocessing methods for SSVEP recognition. In *2016 39th International Conference on Telecommunications and Signal Processing (TSP)* (pp. 340–345). <https://doi.org/10.1109/tsp.2016.7760893>
- [3] Saha, S., Mamun, K. A., Ahmed, K., Mostafa, R., Naik, G. R., Darvishi, S., Khandoker, A. H., & Baumert, M. (2021). Progress in brain computer interface: challenges and opportunities. *Frontiers in Systems Neuroscience*, 15, 578875. <https://doi.org/10.3389/fnsys.2021.578875>
- [4] Boudewyn, M. A., Erickson, M. A., Winsler, K., Ragland, J. D., Yonelinas, A., Frank, M., Silverstein, S. M., Gold, J., MacDonald, A. W., Carter, C. S., Barch, D. M., & Luck, S. J. (2023). Managing EEG studies: How to prepare and what to do once data collection has begun. *Psychophysiology*, 60(11), e14365. <https://doi.org/10.1111/psyp.14365>
- [5] Aznan, N. K. N., Atapour-Abarghouei, A., Bonner, S., Connolly, J. D., Moubayed, N. A., & Breckon, T. P. (2019). Simulating brain signals: Creating synthetic EEG data via neural-based generative models for improved SSVEP classification. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8). <https://doi.org/10.1109/ijcnn.2019.8852227>
- [6] Huang, J., Wen, H., Yang, P., Yang, K., & Xiong, B. (2024). Shift aliasing signal data augmentation method for SSVEP-based BCIS. In *2024 30th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)* (pp. 1–6). <https://doi.org/10.1109/m2vip62491.2024.10746225>

- [7] Pan, Y., Li, N., Zhang, Y., Xu, P., & Yao, D. (2024). Short-length SSVEP data extension by a novel generative adversarial networks based framework. *Cognitive Neurodynamics*, 18(5), 2925–2945. <https://doi.org/10.1007/s11571-024-10134-9>
- [8] Tian, M., Chen, B., Guo, A., Jiang, S., & Zhang, A. R. (2024). Reliable generation of privacy-preserving synthetic electronic health record time series via diffusion models. *Journal of the American Medical Informatics Association*, 31(11), 2529–2539. <https://doi.org/10.1093/jamia/ocae229>
- [9] Debie, E., Moustafa, N., & Whitty, M. T. (2020). A Privacy-Preserving Generative Adversarial Network Method for Securing EEG Brain Signals. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8). <https://doi.org/10.1109/ijcnn48605.2020.9206683>
- [10] Labecki, M., Kus, R., Brzozowska, A., Stacewicz, T., Bhattacharya, B. S., & Suffczynski, P. (2016). Nonlinear Origin of SSVEP Spectra – A combined experimental and modeling study. *Frontiers in Computational Neuroscience*, 10, 129. <https://doi.org/10.3389/fncom.2016.00129>
- [11] Teng, F., Chen, Y., Choong, A. M., Gustafson, S., Reichley, C., Lawhead, P., & Waddell, D. (2011). Square or Sine: Finding a Waveform with High Success Rate of Eliciting SSVEP. *Computational Intelligence and Neuroscience*, 2011, 1–5. <https://doi.org/10.1155/2011/364385>
- [12] Chen, N., Qin, Y., Gao, H., Li, Y., & Sun, P. (2020). Study on induced factors and power variations of the steady-state visual evoked potential subharmonics. *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 10, 1256–1262. <https://doi.org/10.1109/itnec48623.2020.9085060>
- [13] Łabęcki, M., Nowicka, M. M., Wróbel, A., & Suffczynski, P. (2024). Frequency-dependent dynamics of steady-state visual evoked potentials under sustained flicker stimulation. *Scientific Reports*, 14(1), 9281. <https://doi.org/10.1038/s41598-024-59770-5>
- [14] Albahri, A. S., Al-Qaysi, Z. T., Alzubaidi, L., Alnoor, A., Albahri, O. S., Alamoodi, A. H., & Bakar, A. A. (2023). A Systematic Review of Using Deep Learning Technology in the Steady-State Visually Evoked Potential-Based Brain-Computer Interface Applications: Current Trends and Future Trust Methodology. *International Journal of Telemedicine and Applications*, 2023, 1–24. <https://doi.org/10.1155/2023/7741735>
- [15] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. In *MIT Press eBooks*. <https://dl.acm.org/citation.cfm?id=3086952>
- [16] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139–144. <https://doi.org/10.1145/3422622>
- [17] Wang, Y. (2020). A Mathematical Introduction to Generative Adversarial Nets (GAN). *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2009.00169>
- [18] Goodfellow, I. (2016). NIPS 2016 Tutorial: Generative Adversarial Networks. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1701.00160>
- [19] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1606.03498>
- [20] Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein GAN. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1701.07875>
- [21] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of Wasserstein GANs. *arXiv (Cornell University)*, 30, 5769–5779. <https://doi.org/10.48550/arxiv.1704.00028>

- [22] Luo, T., Fan, Y., Chen, L., Guo, G., & Zhou, C. (2020). EEG signal reconstruction using a generative adversarial network with Wasserstein distance and temporal-spatial-frequency loss. *Frontiers in Neuroinformatics*, 14, 15. <https://doi.org/10.3389/fninf.2020.00015>
- [23] Roy, S., Dora, S., McCreadie, K., & Prasad, G. (2020). MIEEG-GAN: Generating Artificial Motor Imagery Electroencephalography Signals. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8). <https://doi.org/10.1109/ijcnn48605.2020.9206942>
- [24] Fahimi, F., Dosen, S., Ang, K. K., Mrachacz-Kersting, N., & Guan, C. (2020). Generative adversarial networks-based data augmentation for brain–computer interface. *IEEE Transactions on Neural Networks and Learning Systems*, 32(9), 4039–4051. <https://doi.org/10.1109/tnnls.2020.3016666>
- [25] Williams, C. C., Weinhardt, D., Wierzberger, M., & Musslick, S. (2023). Augmenting EEG with Generative Adversarial Networks Enhances Brain Decoding Across Classifiers and Sample Sizes. In *Proceedings of the 45th Annual Conference of the Cognitive Science Society*, 45. <https://escholarship.org/uc/item/9gz8g908>
- [26] Song, J., Zhai, Q., Wang, C., & 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>
- [27] Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., & Ganguli, S. (2015). Deep Unsupervised Learning Using Nonequilibrium Thermodynamics. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1503.03585>
- [28] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*. (pp. 6840–6851). Curran Associates Inc. <https://dl.acm.org/doi/10.5555/3495724.3496298>
- [29] Hu, X., Jin, Y., Liang, J., Liu, J., Luo, R., Li, M., & Peng, T. (2023). Diffusion Model for Image Generation – A Survey. In *2023 2nd International Conference on Artificial Intelligence*. <https://doi.org/10.1109/aihcir61661.2023.00073>
- [30] Po, R., Yifan, W., Golyanik, V., Aberman, K., Barron, J. T., Bermano, A., Chan, E., Dekel, T., Holynski, A., Kanazawa, A., Liu, C., Liu, L., Mildenhall, B., Nießner, M., Ommer, B., Theobalt, C., Wonka, P., & Wetzstein, G. (2024). State of the art on diffusion models for visual computing. *Computer Graphics Forum*, 43(2). <https://doi.org/10.1111/cgf.15063>
- [31] Song, J., Meng, C., & Ermon, S. (2020). Denoising diffusion implicit models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2010.02502>
- [32] Rasul, K., Seward, C., Schuster, I., & Vollgraf, R. (2021). Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. *arXiv (Cornell University)*, 8857–8868. <https://doi.org/10.48550/arxiv.2101.12072>
- [33] Lin, L., Li, Z., Li, R., Li, X., & Gao, J. (2023). Diffusion models for time-series applications: a survey. *Frontiers of Information Technology & Electronic Engineering*, 25(1), 19–41. <https://doi.org/10.1631/fitee.2300310>
- [34] Shen, L., & Kwok, J. (2023). Non-autoregressive conditional diffusion models for time series prediction. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2306.05043>
- [35] Kollovich, M., Ansari, A. F., Bohlke-Schneider, M., Zschiegner, J., Wang, H., & Wang, Y. (2023). Predict, Refine, Synthesize: Self-Guiding Diffusion Models for Probabilistic Time Series Forecasting. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2307.11494>

- [36] An, Y., Lam, H. K., & Ling, S. H. (2022). Auto-denoising for EEG signals using generative adversarial network. *Sensors*, 22(5), 1750. <https://doi.org/10.3390/s22051750>
- [37] Habashi, A. G., Azab, M., Eldawlatly, S., & Aly, G. M. (2023). Generative adversarial networks in EEG analysis: an overview. *Journal of NeuroEngineering and Rehabilitation*, 20(1), 40. <https://doi.org/10.1186/s12984-023-01169-w>
- [38] Torma, S., & Szegletes, L. (2024). Generative modeling and augmentation of EEG signals using improved diffusion probabilistic models. *Journal of Neural Engineering*, 22(1), 016001. <https://doi.org/10.1088/1741-2552/ada0e4>
- [39] Torma, S., Tevesz, J., & Szegletes, L. (2023). Generating Visually Evoked Potentials Using a Diffusion Probabilistic Model. In *2023 14th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)* (pp. 61–66). <https://doi.org/10.1109/coginfocom59411.2023.10397513>
- [40] Blankertz, B., Müller, K. R., Curio, G., Vaughan, T. M., Schalk, G., Wolpaw, J. R., Schlögl, A., Neuper, C., Pfurtscheller, G., Hinterberger, T., Birbaumer, N., & Ramoser, H. (2004). The BCI Competition 2003: Progress and perspectives in detection and discrimination of EEG single trials. *IEEE Transactions on Biomedical Engineering*, 51(6), 1044–1051. <https://doi.org/10.1109/tbme.2004.826692>
- [41] Kim, S., Lee, S., Lee, Y., Lee, J., Park, J., & Lee, S. (2023). Brain-driven representation learning based on diffusion model. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2311.07925>
- [42] Soingern, N., Sinsamersuk, A., Chatnuntawech, I., & Silpasuwanchai, C. (2023). Data augmentation for EEG motor imagery classification using diffusion model. In *Communications in Computer and Information Science* (pp. 111–126). https://doi.org/10.1007/978-981-99-7969-1_9
- [43] Tosato, G., Dalbagno, C. M., & Fumagalli, F. (2023). EEG Synthetic data generation using probabilistic diffusion models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2303.06068>
- [44] Eilts, H., & Putze, F. (2022). Is that real? A multifaceted evaluation of the quality of simulated EEG signals for passive BCI. *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2639–2644. <https://doi.org/10.1109/smc53654.2022.9945093>
- [45] Li, X., Metsis, V., Wang, H., & Ngu, H. H. (2022). TTS-GAN: a Transformer-Based Time-Series Generative Adversarial Network. *Lecture Notes in Computer Science*, 133–143. https://doi.org/10.1007/978-3-031-09342-5_13
- [46] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441. <https://doi.org/10.1037/h0071325>
- [47] Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews. Computational Statistics*, 2(4), 433–459. <https://doi.org/10.1002/wics.101>
- [48] Marcin Kołodziej, Andrzej Majkowski (2022). SSVEP classification with a limited training dataset. IEEE Dataport. (accessed February 12, 2025). <https://iee-dataport.org/documents/ssvep-classification-limited-training-dataset>
- [49] Kołodziej, M., Majkowski, A., Oskwarek, Ł., & Rak, R. J. (2016b). Comparison of EEG signal preprocessing methods for SSVEP recognition. In *2016 39th International Conference on Telecommunications and Signal Processing (TSP)* (pp. 340–345). <https://doi.org/10.1109/tsp.2016.7760893>
- [50] Zheng, X., Xu, G., Han, C., Tian, P., Zhang, K., Liang, R., Jia, Y., Yan, W., Du, C., & Zhang, S. (2021). Enhancing performance of SSVEP-based visual acuity via spatial filtering. *Frontiers in Neuroscience*, 15, 716051. <https://doi.org/10.3389/fnins.2021.716051>

- [51] Prasad, V., Zhu-Tian, C., Vilanova, A., Pfister, H., Pezzotti, N., & Strobelt, H. (2023). Unraveling the temporal dynamics of the UNET in diffusion models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2312.14965>
- [52] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). Attention Is All You Need. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1706.03762>
- [53] Salimans, T., & Ho, J. (2022). Progressive distillation for fast sampling of diffusion models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2202.00512>



Przemysław Wiszniewski is a Ph.D. student of biomedical engineering at Warsaw University of Technology (WUT), affiliated with the Institute of Theory of Electrical Engineering, Measurement and Information Systems. He received his B.Sc. (2019) and M.Sc. (2020) in Electrical Engineering from WUT. His research focus lies at the intersection of artificial intelligence, digital signal processing, and brain-computer interfaces.



Marcin Kołodziej is an associate professor at the Institute of Theory of Electrical Engineering, Measurement and Information Systems, Faculty of Electrical Engineering, Warsaw University of Technology, Poland. His research interests include biomedical engineering, biomedical signal processing, brain-computer interfaces, machine learning, artificial intelligence, and expert systems. He has authored and co-authored numerous scientific publications in international peer-reviewed journals and has

contributed to research in biomedical engineering, biometrics, and intelligent data analysis.



Andrzej Majkowski received his M.Sc. degree in electrical engineering from Warsaw University of Technology in 1995, followed by a Ph.D. in 2000 and a postdoctoral degree (D.Sc.) in 2019 from the same institution. Since 1995, he has been with the Institute of Theory of Electrical Engineering, Measurement and Information Systems, Faculty of Electrical Engineering, where he currently holds the position of associate professor. His research interests include digital signal processing, multimedia

technologies, and bioengineering. He is a member of the IEEE Engineering in Medicine and Biology Society and has authored over 50 journal publications and numerous international conference papers.