A COMPARATIVE STUDY OF GENERATIVE ADVERSARIAL NETWORKS FOR IMAGE RECONSTRUCTION FROM BRAIN ACTIVITY

by

Jay Webb, 10704047

Thesis Proposal submitted to University of Plymouth
in partial fulfilment of the requirements for the degree of

MSc Electrical and Electronics Engineering

University of Plymouth
Faculty of Science & Engineering

September 2025

Table of Contents:

1 Introduction to the Research Area

The reconstruction of visual experiences from human brain activity stands as a compelling challenge at the intersection of neuroscience and BCI research (Shen et al., 2019; St-Yves & Naselaris, 2018). The ability to decode neural signals associated with vision, captured through neuroimaging techniques like fMRI and EEG, offers new insights into the neural basis of perception and enables the development of advanced BCI applications (Kumari et al., 2023; Shimizu & Srinivasan, 2022). Among the machine learning methodologies explored for this purpose, Generative Adversarial Networks (GANs) have emerged as particularly promising models, renowned for their ability to learn intricate data distributions and generate realistic synthetic data, including images (Goodfellow et al., 2014; Kumari et al., 2023).

Building upon the increasing interest in leveraging EEG's high temporal resolution and accessibility for neural decoding, GANs are being actively investigated for the reconstruction of visual stimuli from these brain signals (Khaleghi et al., 2022; Kumari et al., 2023; Shimizu & Srinivasan, 2022). The core idea involves conditioning a GAN's generator network on processed EEG data to synthesize images that approximate the visual input that elicited the observed neural activity (Kumari et al., 2023). The adversarial training paradigm, where a discriminator network learns to distinguish between generated and real images, compels the generator to produce increasingly accurate and plausible reconstructions (Khaleghi et al., 2022). While initial research has demonstrated the feasibility of GANs for reconstructing basic visual elements or salient aspects of perceived scenes from EEG (Khaleghi et al., 2022; Shimizu & Srinivasan, 2022), achieving high-fidelity reconstructions that capture the nuanced details of visual experience remains a significant hurdle. This necessitates a thorough investigation into the efficacy of diverse GAN architectures and processing techniques tailored for the unique characteristics of EEG data in the context of visual reconstruction.

This research endeavours to contribute to this area by conducting a comparative analysis of various GAN approaches to reconstruct seen images from EEG data, focusing on the extraction of relevant temporal and spatial features from pre-processed EEG signals to train and evaluate these models and ultimately identify potential benchmarks for future advancements in this field.

2 Aims and Objectives

2.1 Aims

To extract temporal and spatial features of pre-processed EEG data, using data recorded by the university as well as open-source data, to build, train and test various GAN architectures and evaluate their performance based on the quality of the reconstructed images compared to the original seen images.

2.2 Objectives

1. Analyse and evaluate existing literature relating to image reconstruction from brain activity using GAN networks.
2. Pre-process EEG data to select relevant frequencies and remove any noise in the signal.
3. Process the EEG data to obtain temporal and spatial features of the signal and format in latent structure.
4. Build, train and test various GAN networks for many epochs.
5. Evaluate the regenerated images against the original seen images using evaluation metrics.
6. Propose GAN architecture as a benchmark for future research on image reconstruction from brain activity.

3 Review

EEG feature extraction is a crucial step in processing electroencephalogram signals to derive meaningful information for various applications, including image reconstruction. Commonly used techniques can be broadly categorized into time-domain, frequency/spectral domain, and decomposition domain methods. Time-domain features involve direct analysis of the EEG signal's amplitude over time, including measures like mean, variance, and Hjorth parameters. Frequency domain methods transform the signal into the frequency spectrum using techniques like Fourier Transform (FT) or Power Spectral Density (PSD) to extract features related to different frequency bands (e.g., alpha, beta, gamma) (Singh & Krishnan, 2023). Decomposition domain techniques, such as Wavelet Transform (WT) and Empirical Mode Decomposition (EMD), decompose the EEG signal into different components to extract relevant features across various scales and frequencies (Singh & Krishnan, 2023; Shimizu & Srinivasan, 2022). For image reconstruction using Generative Adversarial Networks (GANs), these extracted features serve as conditional information or are mapped to a latent space from which images can be generated (Khaleghi et al., 2022; Kumari et al., 2023). Some studies utilize the raw time samples of EEG channels directly as input to deep networks, bypassing explicit feature extraction (Khaleghi et al., 2022). Furthermore, the conversion of EEG signals into spectrograms using Short-Time Fourier Transform (STFT) creates a time-frequency representation that can be used as input for GAN-based image regeneration (Kumari et al., 2023).

The selection of appropriate EEG feature extraction methods significantly impacts the performance of image reconstruction using GANs. Frequency domain features, by capturing oscillatory brain activity related to visual processing, are often employed when visual stimuli are presented (Kumari et al., 2023; Shimizu & Srinivasan, 2022). For instance, the power in specific frequency bands might correlate with certain image characteristics. Decomposition methods like Wavelet Transform can provide a multi-resolution analysis, potentially capturing transient and non-stationary EEG patterns associated with image perception or imagery (Singh & Krishnan, 2023; Shimizu & Srinivasan, 2022). Studies have also explored using time-frequency representations like spectrograms as input to GANs, allowing the network to learn relevant spatio-temporal features directly from the EEG data (Kumari et al., 2023). This approach can be advantageous as it reduces the need for manual feature engineering. In the context of GANs, these extracted features or representations are often used to condition the generator, guiding the image generation process to produce images that are consistent with the brain activity (Kumari et al., 2023; Khaleghi et al., 2022). For example, features related to visual saliency extracted from EEG could be used to guide the GAN towards reconstructing the most visually prominent aspects of the perceived image (Khaleghi et al., 2022). The effectiveness of these methods can be influenced by factors such as the quality of EEG data (signal-to-noise ratio), the number and placement of EEG electrodes, and the specific visual task or stimuli used in the experiment (Singh & Krishnan, 2023).

While various EEG feature extraction techniques have been applied to image reconstruction using GANs, the field is still evolving, and a universally superior method has not yet emerged. The choice of features often depends on the specific research question, the nature of the visual information being decoded, and the architecture of the GAN being used. The trend towards using time-frequency representations or even raw EEG data as input to deep learning models suggests a shift away from relying solely on handcrafted features. This allows the models to automatically learn the most relevant features for image reconstruction. However, the interpretability of features learned directly by deep networks can be challenging. Evaluating the performance of these methods requires appropriate metrics that assess both the low-level visual similarity (e.g., Structural Similarity Index Measure - SSIM) and the high-level semantic similarity between the original and reconstructed images (Meng & Yang, 2024; Kumari et al., 2023). Future research could focus on developing more sophisticated feature extraction techniques that can better capture the complex and dynamic nature of brain activity related to visual processing. Additionally, exploring methods to enhance the conditioning of GANs with EEG features and improving the training stability of these models are crucial for achieving higher quality and more reliable image reconstructions from EEG signals (Kumari et al., 2023). The combination of advanced signal processing techniques with the generative power of GANs holds significant promise for advancing our understanding of visual perception and developing effective brain-computer interfaces (Khaleghi et al., 2022; Kumari et al., 2023).

GANs have emerged as a powerful tool for image generation and are increasingly being explored for reconstructing visual stimuli from brain activity measured by EEG (Kumari et al., 2023; Shimizu & Srinivasan, 2022). In this context, the goal is to train a generative model that can map patterns of neural activity, captured by EEG, to corresponding visual images. The generator then learns to produce images that are statistically similar to the visual stimuli that evoked the observed EEG activity. The discriminator network, the adversarial counterpart, learns to distinguish between reconstructed images and real images from the target distribution, thereby forcing the generator to produce more realistic and accurate reconstructions (Khaleghi et al., 2022; Kumari et al., 2023). Various GAN architectures, including Deep Convolutional GANs (DCGANs) and more advanced variants like Capsule GANs (CapsGANs), have been adapted for this task (Kumari et al., 2023; Khaleghi et al., 2022). Some approaches also incorporate additional modules, such as geometric deep networks, to first map EEG signals to intermediate representations like visual saliency maps, which then guide the GAN for image reconstruction (Khaleghi et al., 2022).

The application of GANs to EEG-based image reconstruction addresses the inherent challenges of neural decoding, such as the noisy and high-dimensional nature of EEG data and the limited availability of paired EEG-image datasets (Khaleghi et al., 2022; Kumari et al., 2023). GANs, with their ability to learn complex data distributions, offer a potential pathway to generate visually plausible images even from indirect and potentially incomplete neural information. Conditioning the GAN generator on processed EEG signals allows the model to associate specific brain activity patterns with corresponding visual content (Kumari et al., 2023; Shimizu & Srinivasan, 2022). For instance, using EEG spectrograms as input can enable the GAN to learn the temporal dynamics and frequency-specific neural responses related to visual perception or imagery (Kumari et al., 2023). The adversarial training process encourages the generator to refine its output, leading to reconstructions that are not merely average representations but can capture some level of detail and variability present in the original stimuli (Khaleghi et al., 2022). However, the quality of the reconstructed images is heavily dependent on the effectiveness of the EEG signal processing, the capacity of the GAN architecture, and the strength of the learned association between EEG and visual features. Furthermore, the interpretability of the mapping learned by the GAN remains a challenge, making it difficult to understand precisely how specific EEG patterns contribute to the generated image features.

The research on using GANs for image reconstruction from EEG signals demonstrates promising initial results, particularly in reconstructing basic outlines or salient features of visual stimuli (Khaleghi et al., 2022; Kumari et al., 2023; Shimizu & Srinivasan, 2022). Studies employing advanced GAN architectures like CapsGAN have reported improvements in image quality metrics such as the Structural Similarity Index Measure (SSIM), suggesting a closer resemblance to the original images (Kumari et al., 2023). Approaches that incorporate intermediate steps, like saliency map prediction, can also provide a more guided reconstruction process (Khaleghi et al., 2022). Nevertheless, the field still faces significant challenges. The reconstructed images often lack fine-grained details and can be blurry or semantically inaccurate. The limited size of EEG-image datasets necessitates the development of robust training strategies and potentially the use of transfer learning or data augmentation techniques. Evaluating the fidelity of the reconstructions requires a combination of objective metrics (e.g., SSIM, FSIM) and subjective human evaluations to assess both perceptual quality and semantic consistency (Kumari et al., 2023). Future research should focus on developing more sophisticated GAN models that can better leverage the information present in EEG signals, exploring novel ways to condition GANs on brain activity, and addressing the challenges of data scarcity and model interpretability to achieve higher quality and more reliable image reconstructions from EEG (Khaleghi et al., 2022; Kumari et al., 2023; Shimizu & Srinivasan, 2022).

4 Proposed Research Design

4.1 Qualitative Data Pathway

4.1.1 Data Requirements:

- Existing literature reviews related to GAN-based image reconstruction from brain activity.
- Detailed technical reports on various GAN architectures as documented in the literature.
- Expert evaluations regarding their performance.

4.1.2 Data Collection Methods:

- A systematic literature review will be undertaken to source peer-reviewed articles, technical reports, and industry white papers addressing GAN applications in neural decoding.
- Structured interviews and email surveys with academic supervisors and computer vision specialists, to discuss technical feasibility and challenges associated with different GAN approaches.
- Secondary data will be sourced from public repositories and research databases.

4.1.3 Data Analysis:

Objective: To identify and synthesize the key factors that influence the success of GAN-based image reconstruction from brain activity.

Methodology: Thematic analysis will be applied to any interview transcripts and survey responses, with particular attention to emergent themes such as image realism and the perceived practical utility of the reconstruction methods. This analysis will provide valuable context for interpreting the quantitative results and help inform subsequent modelling decisions.

4.2 Quantitative Pathway

4.2.1 Objective:

The quantitative pathway is designed to model and compare multiple GAN architectures with their ability to reconstruct images from brain activity data.

4.2.2 Methodology:

- Quantitative parameters will be derived from the literature review and qualitative insights.
- High-resolution brain imaging datasets (EEG recordings) paired with corresponding visual stimuli will be employed as training and testing data. Open-source datasets will be utilised as well as datasets recorded by the university under specific conditions.
- The study will implement and compare various GAN models using simulation frameworks in Matlab/Python.
- Performance indicators will include objective metrics such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and mean squared error (MSE), along with measures of computational efficiency and convergence behaviour under different operational scenarios.

4.2.3 Expected Outcomes:

This computational evaluation will generate robust, data-driven evidence regarding the relative performance of the GAN architectures under study. Quantitative findings are expected to either validate or challenge the qualitative insights gathered from expert interviews, thereby pinpointing optimal model configurations for accurate and efficient image reconstruction. Ultimately, the integration of qualitative and quantitative results will provide a comprehensive understanding of the feasibility and potential impact of deploying GAN-based methods in neural decoding, informing both future research directions and potential clinical applications.

5 Project Risk Assessment

Risk	**Likelihood**	**Impact**	**Mitigation Strategy**
Failure of lab equipment/laptop	Moderate	Moderate	Use GitHub and regularly commit work to back up completed work
Technical Complexity and Implementation	Moderate	High	Seek provisional guidance and leverage current research to identify proven architectures.
Data availability and data quality	Moderate	High	Identify and access to data sources early, including public dataset. Preprocess to improve data quality

Uncertain outcomes and reproducibility	Moderate	High	Design experiments with clear, quantifiable metrics and include cross-validation techniques.
Computational Cost	High	Moderate	Use specialist libraries and tools to optimise algorithms. Make use of GPUs and HPC to improve calculation time.

6 Ethics

Human participants will be included for surveys. Approval through application via the Plymouth Ethics Online System (PEOS) is required to ensure compliance with university and national research ethics policies. Data confidentiality and participant consent will be strictly maintained. The application for ethical approval will be submitted by 28th March 2025.

7 Impact

The proposed research has the potential to:

1. Scientific and Technological Advancement: Specifically, both neuroscience and artificial intelligence by investigating the potential of GANs for decoding brain activity. The potential of enhancing our understanding of neural representations, providing empirical evidence on the relationship between brain signals and visual stimuli. Furthermore, by rigorously comparing multiple GAN architectures, the study fosters innovation in algorithm development and model optimization, thereby contributing to methodological improvements.
2. Social and Health Related Benefits: Potential to yield substantial benefits in the realms of social welfare and healthcare by informing the development of advanced brain–computer interfaces. Such interfaces could facilitate improved communication and control mechanisms for individuals with severe motor impairments, thereby enhancing their quality of life. Moreover, the research may contribute to the creation of novel diagnostic tools for neurological disorders, supporting personalized medicine approaches and offering clinicians new pathways to assess and treat complex brain-related conditions.
3. Economic Growth and Commercialization: The development of novel software tools and devices based on the project’s outcomes could pave the way for commercially exploitable products. By fostering industry-academic collaborations and establishing intellectual property, the research not only supports economic growth but also encourages the translation of scientific discoveries into market-ready applications, thereby bridging the gap between cutting-edge research and practical, economically viable solutions.
4. UN’s Sustainable Goals: Aligns with several of the United Nations Sustainable Development Goals by addressing critical issues in health, innovation, and education. It has the potential to improve healthcare outcomes (SDG 3) through enhanced diagnostic and therapeutic tools, foster industrial innovation and infrastructure development (SDG 9) via cutting-edge AI and neurotechnology research, and contribute to quality education (SDG 4) by disseminating advanced scientific knowledge.

8 List of References

Kumari, N., Anwar, S., Bhattacharjee, V. and Sahana, S.K. (2023). Visually evoked brain signals guided image regeneration using GAN variants. *Multimedia Tools and Applications*. doi:https://doi.org/10.1007/s11042-023-14769-4.

Khaleghi, N., Rezaii, T.Y., Beheshti, S., Meshgini, S., Sheykhivand, S. and Danishvar, S. (2022). Visual Saliency and Image Reconstruction from EEG Signals via an Effective Geometric Deep Network-Based Generative Adversarial Network. *Electronics*, 11(21), p.3637. doi:https://doi.org/10.3390/electronics11213637.

Shimizu, H. and Srinivasan, R. (2022). Improving classification and reconstruction of imagined images from EEG signals. *PLOS ONE*, 17(9), p.e0274847. doi:https://doi.org/10.1371/journal.pone.0274847.

Meng, L. and Yang, C. (2024). Semantics-Guided Hierarchical Feature Encoding Generative Adversarial Network for Visual Image Reconstruction From Brain Activity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32, pp.1267–1283. doi:https://doi.org/10.1109/tnsre.2024.3377698.

Shen, G., Dwivedi, K., Majima, K., Horikawa, T. and Kamitani, Y. (2019). End-to-End Deep Image Reconstruction From Human Brain Activity. *Frontiers in Computational Neuroscience*, 13. doi:https://doi.org/10.3389/fncom.2019.00021.

Ghislain St-Yves and Naselaris, T. (2018). Generative Adversarial Networks Conditioned on Brain Activity Reconstruct Seen Images. *bioRxiv (Cold Spring Harbor Laboratory)*. doi:https://doi.org/10.1101/304774.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative Adversarial Nets. [online] Available at: https://papers.nips.cc/paper_files/paper/2014/file/f033ed80deb0234979a61f95710dbe25-Paper.pdf.

Singh, A. K., & Krishnan, S. (2023). Trends in EEG signal feature extraction applications. Frontiers in Artificial Intelligence, 5, 1072801. doi:https://doi.org/10.3389/frai.2022.1072801.

9 Bibliography

Ren, Z., Li, J., Xue, X., Li, X., Yang, F., Jiao, Z. and Gao, X. (2021). Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning. *NeuroImage*, 228, pp.117602–117602. doi:https://doi.org/10.1016/j.neuroimage.2020.117602.

Seeliger, K., Güçlü, U., Ambrogioni, L., Güçlütürk, Y. and van Gerven, M.A.J. (2018). Generative adversarial networks for reconstructing natural images from brain activity. *NeuroImage*, 181, pp.775–785. doi:https://doi.org/10.1016/j.neuroimage.2018.07.043.

Luo, T., Fan, Y., Chen, L., Guo, G. and Zhou, C. (2020). EEG Signal Reconstruction Using a Generative Adversarial Network With Wasserstein Distance and Temporal-Spatial-Frequency Loss. *Frontiers in Neuroinformatics*, 14. doi:https://doi.org/10.3389/fninf.2020.00015.

Jaiswal, A., AbdAlmageed, W., Wu, Y. and Natarajan, P. (2018). CapsuleGAN: Generative Adversarial Capsule Network. arXiv (Cornell University). [online] doi:https://doi.org/10.48550/arxiv.1802.06167.