Impressionist StyleGAN

Alexander Johannes Grabner
Technische Universität Wien
e12222190@student.tuwien.ac.at

1 INTRODUCTION

This project focuses on creating a diverse, multi-purpose dataset of Impressionist artworks. Although several Impressionist art datasets already exist (1; 2), they often contain relatively few images and frequently provide only the images themselves, without relevant metadata. This limits their usefulness for tasks beyond image generation, such as predicting the year a painting was created.

To address this, I curated a dataset of 23,238 Impressionist artworks, with each entry accompanied by metadata. Using this dataset, I implemented *StyleGAN* (3) and trained it on the collected artworks.

This report is organized as follows. Section 2 introduces Impressionism and explains why it was chosen for this project. Section 3 describes the dataset and its metadata. Section 4 presents the *StyleGAN* implementation and the training procedure. Section 5 summarizes the main takeaways from the project and outlines what I would do differently in hindsight. Finally, Section 6 provides a breakdown of the time spent on the individual tasks.

2 BACKGROUND

Impressionism was a 19th-century art movement that originated in France. Early Impressionists, such as Claude Monet and Pierre-Auguste Renoir, were regarded as radicals by the French academic painting community, since Impressionism violated many rules of classical academic painting. Impressionists focused on capturing movement and light, often painting outdoors and not emphasizing the precision of their brushstrokes, rather than aiming to paint their scenes with photographic precision. (4)

Today, the founders of the Impressionist movement are long dead, and many other styles have since emerged. Nevertheless, its characteristics make Impressionism an intriguing choice for training a modern deep learning architecture. Modern image generation models such as *StyleGAN* (3) and diffusion models (5) are known for their ability to generate images in different styles. If trained long enough and on sufficient data, they can produce images that are almost indistinguishable from the real images they were trained on. However, such training is time-consuming and often requires multiple state-of-the-art machine learning GPUs. Therefore, I hypothesized that training becomes easier if the dataset contains content that is inherently “imprecise.” Impressionism fits this property for the reasons discussed above. In addition, many Impressionist artworks exist, and many of them are visually similar, which can further help the models. For example, Claude Monet painted many similar versions of his *Water Lilies* series. This makes a dataset of Impressionist artworks well-suited for training an image generation model. Another advantage of Impressionism for such a project is that most works stem from the late 19th century, meaning many are in the public domain and therefore easy to use in a machine learning dataset.

3 DATASET

I created a dataset of 23,238 Impressionist artworks, with each entry containing metadata about the artwork. The dataset was acquired from public-domain sources such as *WikiArt* (6). WikiArt offers convenient functionality for gathering large datasets of images belonging to a specific art style such as Impressionism, since it provides text lists of URLs linking to artworks of the chosen style.

The final dataset contains images as well as metadata, such as the genre of the painting (e.g., still life, landscape) and a `CLUSTER_IND` column. During early training runs, I noticed that the model had difficulty generalizing to the full dataset. I hypothesized that this was because it contains images from vastly different genres; for example, the dataset includes both portraits and landscape paintings, while StyleGAN architectures are most often trained on homogeneous datasets, such as *CelebA-HQ* (7).

Therefore, I applied a K-means clustering preprocessing step using embeddings generated by OpenAI's CLIP image encoder (8). This subdivided the dataset into six clusters, where each cluster contains images that are visually and semantically similar to one another. The final model was trained on the cluster with ID 0, which contains 6,251 images. Figure 1 shows a t-SNE visualization of the six clusters, and Figure 2 shows samples from cluster 0.

The final dataset contains the following columns:

Id Unique identifier of the dataset entry.
Author Artist/author of the artwork.
Style Artistic style (e.g., Impressionism, Post-Impressionism).
Title Title of the artwork.
Date Estimated year (or date range) when the artwork was created.
Genre Subject/theme (e.g., still life, portrait, landscape).
Image_urls List of one or more URLs pointing to the artwork image file(s).
URL Source webpage URL the record was collected from.
Cluster_ind Cluster ID assigned by K-Means clustering.

4 IMPRESSIONIST STYLEGAN

To generate images in the style of Impressionism, I built my own implementation of the StyleGAN architecture and trained it on my self-curated dataset. The final model achieved an FID6k value of 40.38. Figure 3 shows a grid of images generated by this final model.

4.1 Model Architecture

The architecture of my StyleGAN model largely follows the original design described by Karras et al. (3). However, I incorporated some elements from later StyleGAN variants. After partitioning the dataset using K-means clustering, I encountered issues with discriminator overfitting. I mitigated this by incorporating *Adaptive Discriminator Augmentation (ADA)*, as described in StyleGAN2-ADA (9).

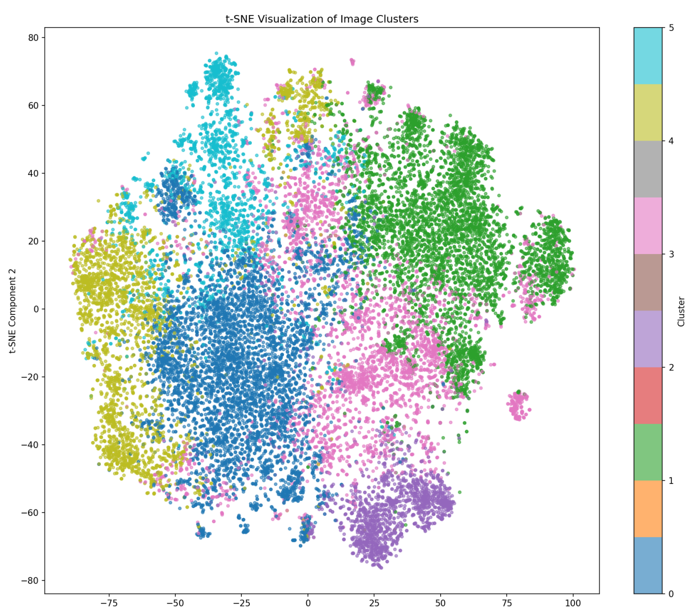

Figure 1. t-SNE Visualization of Image Clusters

Figure 2. Cluster 0: Image Samples

Figure 3. Images Generated by Final Model

Table 1: Model hyperparameters.

Component	**Parameter**	**Value**
Latent spaces	Z dim	512
	W dim	512
Mapping network	Layers	8
Optimization	Adam (β_1, β_2)	$(0.0,\ 0.99)$
	Discriminator steps	1
	Style-mixing prob.	0.5
Learning rates	Generator/Discriminator base LR	0.002
	Mapping network LR	$0.01 \times 0.002 = 2 \times 10^{-5}$
	EMA Generator Weight Decay	0.999
Architecture	Kernel size	3
	Base channels	512
	Floor channels	128
	Max output resolution	128×128
	Upsampling interpolation	bilinear
ADA	Target	0.6

4.2 Training Procedure

The model was trained on a subset of 6,251 images from the full dataset using the following per-resolution settings. Due to GPU constraints, training had to be stopped early, at the end of the 64×64 resolution phase.

Table 2: Progressive growing schedule and per-resolution settings.

Res.	**Channels**	**Batch**	**Images**	$\mathbf{LR}_{\text{model}}$	$\mathbf{LR}_{\text{map}}$
4	512	128	300 000	0.002	0.0002
8	512	128	350 000	0.002	0.0002
16	512	64	400 000	0.002	0.0002
32	512	32	450 000	0.002	0.0002
64	256	32	600 000	0.002	0.0002
128	128	16	800 000	0.002	0.0002

Fade-in uses 50% of the images at each resolution.

5 TAKEAWAYS

Below are some of the key takeaways and insights I gathered while working on this project.

1. **FID is sensitive to dataset size.** In early runs after partitioning my dataset, I observed higher FID values even though the subjective image quality improved. This was due to FID's sensitivity to dataset size.
2. **Dataset homogeneity strongly affects StyleGAN generator quality.** In early training runs on the full dataset, the model struggled to generate high-quality images. This was most likely due to the heterogeneity of the dataset in genre, composition, and other factors. After partitioning the dataset with K-means clustering, image quality improved dramatically.
3. **ADA works well with the original StyleGAN architecture.** After partitioning my dataset and encountering discriminator overfitting, I was initially skeptical about how effective ADA would be in combination with the original StyleGAN architecture, since ADA was introduced in later versions. However, it turned out to be very effective.

4. The number of channels strongly affects output quality. For most training runs, I aggressively reduced the number of channels used in each resolution layer. In my final run, I kept the number of channels at 512 for the first four layers, which significantly improved both the perceived quality and the FID score of the generator output.

5.1 If I could do it again, what would I do differently?

I plan to continue the *Impressionist StyleGAN* project after acquiring a stronger GPU, allowing training at higher resolutions (up to 256×256) and on additional clusters of the dataset. To do this effectively, I will use an evaluation metric that is less dependent on dataset size, such as *Kernel Inception Distance (KID)* (10). This will allow me to compare the quality of images generated by models trained on differently sized clusters more reliably.

Another change I would make is to implement a later StyleGAN variant, such as StyleGAN2-ADA (9), instead of relying on the original architecture. StyleGAN2-ADA is simpler due to its elimination of progressive growing and typically produces higher-quality outputs.

6 WORK BREAKDOWN

Below is a breakdown of how much time was spent on each task in this project. Building the model took significantly longer than my original estimate of 10 hours, since I went through many iterative steps to progressively improve it:

Table 3: Time spent per project task.

Task	**Time**
Researching project ideas	2 h
Planning project	2 h
Finding dataset sources	6 h
Writing code for dataset collection	5 h
Collecting all dataset entries	10 h
Refining dataset	4 h
Building model	40 h
Training the model	25 h
Building an application to present the results	10 h
Preparing the final report and presentation	10 h
Total	**114 h**

REFERENCES

[1] Panchajanya Banerjee. Impressionist classifier data. Accessed: 16.12.2025.

[2] robgonsalves. Impressionist-landscapes-paintings. Accessed: 16.12.2025.

[3] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410, 2019.

[4] Wikipedia Contributors. Impressionism. Accessed: 16.12.2025.

[5] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[6] WikiArt Team. Wikiart. Accessed: 16.12.2025.

[7] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.

[8] OpenAI. Clip: Connecting text and images. Accessed: 16.12.2025.

[9] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. *Advances in neural information processing systems*, 33:12104–12114, 2020.

[10] Mikołaj Bińkowski, Danica J Sutherland, Michael Arbel, and Arthur Gretton. Demystifying mmd gans. *arXiv preprint arXiv:1801.01401*, 2018.