Brushstrokes of the Machine: Generating Van Gogh-Style Artworks with GAN-Enhanced ResNet and U-Net Models

Ayesha Siddiqa Kanza Rizwan Navroop Chandwani

University of Toronto
ayesha.siddiqa@mail.utoronto.ca
kanza.rizwan@mail.utoronto.ca
navroop.chandwani@mail.utoronto.ca

Abstract

In the pursuit of automating artistic creativity, this study addresses the problem of non-paired image-to-image translation with the goal of transforming modern photographs to emulate the impressionistic style of Vincent van Gogh. Utilizing a CycleGAN framework, we evaluate the performance of two distinct generative architectures: a baseline U-Net model and a modified ResNet model. Our methodology involves training these networks to internalize Van Gogh's stylistic nuances from an unpaired dataset and subsequently applying this learned style to modern imagery. The models' effectiveness is measured not only through the discriminative accuracy of a PatchGAN classifier but also through qualitative assessment of the stylized output images. The dataset includes 2025 images sourced from Van Gogh's art works alongside 7038 contemporary photographs. Preliminary results indicate the superior capability of the ResNet architecture over the U-Net baseline in generating stylized images, showcasing its potential in capturing the intricacies of impressionist art. This approach offers a pathway for further exploration into the domain of artistic intelligence in machine learning.

1 Introduction

The use of image-enhancing filters on social media platforms has demonstrated the potential of machine learning technologies to transform simple photographs into visually appealing artworks. Our research project builds upon this foundation by aiming to convert conventional images into artworks that emulate the distinctive impressionistic style of Vincent van Gogh. The central challenge we address is the transformation of standard photographs into Van Gogh-styled paintings without relying on paired training examples.

We approach this challenge by employing a CycleGAN framework, a type of generative adversarial network that is adept at capturing the essence of one image domain and transferring it to another. Our model architecture incorporates two generator networks, U-Net and ResNet, to evaluate their performance and optimize the quality of the artistic transformation. These generators are tasked with the bidirectional conversion of images, from the domain of standard photography to that reflective of Van Gogh's paintings and vice versa. In parallel, discriminator networks are trained to distinguish between the transformed images and their authentic counterparts. This adversarial process, fine-tuned through a cycle-consistency loss function, is designed to guide the generators toward producing images that convincingly replicate Van Gogh's signature style.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

CycleGAN was chosen for its efficacy in learning from unpaired datasets and its robustness in smaller data environments, which is particularly beneficial given the challenge of acquiring extensive sets of paired images for training. Beyond creating art, this project attempts to solve the challenge of utilizing machine learning frameworks to replicate complex art styles. By developing an architecture that makes it possible to generate intricate paintings from regular photos, we take a step toward understanding the potential of artificial intelligence in creative fields.

2 Background and Related Work

A U-Net generator is used in our model to map the input image to a Van Gogh-style output image, with some shared underlying content but a different style. U-Net's architecture consists of a series of convolutional and pooling layers that progressively down-sample the input image. This contractive path captures increasingly abstract representations of the input image data. As the data moves across the layers, the spatial dimensions are reduced but the number of channels increases. The bottleneck layer (bottom curve of 'U') is where the deepest and most compressed representation is formed. Gradually upsampling occurs through transposed convolutions as the feature representations revert to their original image size.

A Generative Adversarial Network (GAN) consists of two neural networks, the generator and the discriminator, which are trained simultaneously through adversarial processes. At a low level, the generator tries to produce fake data instances that are good enough to fool the discriminator and the discriminator tries to get better at distinguishing real data from the fake ones.

CycleGAN is a specific kind of GAN that is used for unpaired image-to-image translations. In our model, the input and output images don't have the exact before-and-after examples of what the transformation is supposed to look like.

CycleGAN uses a PatchGAN discriminator, which is a neural network architecture, used for tasks like style transfer and image generation. It ensures that Van Gogh's style transfer is effectively applied in a locally coherent way across the input image. PatchGAN assesses the authenticity of patches in an image, instead of classifying the entire image as real or fake (as done by traditional GAN). In our model, PatchGAN outputs a matrix of values where each value checks the corresponding patch of the input image. Values that are closer to 1 are more likely to be classified as real and values closer to 0 are more likely classified as fake.

Van Gogh's technique, characterized by relatively small, thin, yet visible brush strokes, open composition, emphasis on accurate depiction of light in its changing qualities, and ordinary subject matter, presents a unique challenge to the field of neural style transfer. Recent advancements in deep learning have seen attempts to replicate such styles algorithmically, with varying degrees of success in capturing the subtleties of color interplay and brushstroke texture.
Several papers have addressed the problem of style transfer and image-to-image translation using deep learning:

1. Goodfellow et al. [3] proposed a new framework for estimating generative models via adversarial nets, that uses a generative model G for capturing the data distribution, and a discriminative model D to the probability of a sample belonging to the training set, both running simultaneously.
2. The influential paper on CycleGAN paper by Zhu et al. [5] demonstrates unsupervised image-to-image translation, utilizing cycle-consistency loss to maintain key attributes between source and target domains without paired examples. In [5] Jun Yan Zhu wt al. present "an approach for learning to translate an image from a source domain X to a target domain Y in the absence of paired examples." A cycle consistency loss which maps Y->X is introduced because the mapping from X->Y is highly constrained.
3. Leon A. et al. [2] introduced the concept of neural style transfer, where the style of a famous painting can be applied to a photograph.

4. A more recent and successful approach to learn transformations between image distribution comes from the work done by Chu et al. [1]. The papers explores how CycleGANs can be utilized to embed information within images in a way that is not detectable by human observers, demonstrating the model's capability for data hiding alongside its traditional use in image-to-image translation.

Our model and approach diverge from the above-mentioned methodologies by integrating a modified loss function and a U-Net generator that better captures the distinctive elements of Van Gogh's style while preserving the original structural contents of the image.

3 Data

Our study utilizes datasets sourced from Kaggle, a popular platform where data scientists and machine learning enthusiasts share and explore datasets, and participate in analytical competitions. The first dataset, titled "Van Gogh Paintings," comprises various categories of Van Gogh's artwork [6]. To closely align with his distinctive painting style, we specifically chose 308 images from folders named 'Aries,' 'Auvers sur Oise,' 'Saint Remy,' and 'Nuenen.' These folders were selected for their consistent representation of Van Gogh's painted works, distinct from his sketches and drawings.

Additionally, we used the "I'm Something of a Painter Myself" dataset from Kaggle for modern digital images [4]. Out of 7,038 available images, we selected 2,000 that varied widely in subject and setting. This selection process aimed to provide a diverse range of training examples for the model. All images used in our study are in PNG format. This uniformity simplifies the data preprocessing stage, where we standardize image resolution, color, and size. These steps ensure that our model receives consistent and high-quality input.

By combining these carefully curated datasets, our model is trained to transform modern photographs into artworks that mirror Van Gogh's iconic painting style. This approach bridges historical art with contemporary imagery, creating a unique blend of past and present visual expressions.

4 Model Architecture

As previously mentioned, our model aims to generate images reminiscent of Van Gogh's artistic style using real photographs as input. Essentially, it will replicate the typical features of Van Gogh's paintings and apply them to transform the original photograph into a possible painting with similar characteristics. Particularly, we want to train an unsupervised image translation model where our data set includes unpaired samples of actual photographs and Van Gogh-style paintings.

Since our model will be creating new data instances, we will be using GANs, **General Adversarial Models**, model architecture which allows the use of generative models to create paintings that would follow Van Gogh's style. GAN is a class of machine learning framework that makes use of two different types of models,

Generator model:
This aims to generate new data instances. We intend to train our model using two distinct generator models: **UNet** and **ResNet**. These will serve as our hyperparameters, and the generator model type that yields superior accuracy is selected for the final implementation, as explained in the 'Results' section.
The **UNet** generator is a convolutional neural network comprising an encoder and a decoder. The encoder is responsible for down-sampling the input data, and extracting high-level features. Subsequently, the decoder up-samples this processed information, filtering intermediate features to generate the final output.
The **ResNet** generator is a convolutional neural network design featuring residual blocks, which addresses the challenges of training deep networks. It includes skip connections that enable the flow of information directly from one layer to another. This facilitates the learning of residual features,

allowing the model to effectively capture intricate details and nuances in the input data during both down-sampling and up-sampling processes.

Discriminator model:
In our model, we specifically employ a **PatchGAN** discriminator, a choice well-suited for GAN models. This discriminator classifies the input image by breaking it down into local patches, with each patch categorized as either real or generated data. The discriminator's output is derived by averaging responses across all patches throughout the entire image.

As previously mentioned, our model confronts the challenge of lacking paired data – distinct datasets for genuine photographs and Van Gogh-style paintings. This departure from conventional input/label supervised training data steers our image translation model towards a specialized variant of GAN architecture, referred to as **CycleGAN**. This architectural choice ensures that our output image (Van Gogh painting) is an authentic transformation of the input image (actual photograph).

To implement this cycle for our unpaired data, our CycleGAN model features two generators and two discriminator models. This configuration contributes to the robustness and efficacy of the image translation process.

Let the data set containing Van Gogh style paintings be $\mathbf{V}$ and the data set containing photographs be $\mathbf{P}$, where $\mathbf{p} \in \mathbf{P}$ and $\mathbf{v} \in \mathbf{V}$.

4.1 Model Components

The model can be categorized into the following components:

- **Van Gogh Generator Model** $Generator_{P \to V}$: Transforms a Photograph into a Van Gogh-style Painting
- **Photograph Generator Model** $Generator_{V \to P}$: Transforms a Van Gogh style Painting into a Photograph
- **Van Gogh Discriminator Model** $Discriminator_V$: Distinguishes between a generated Van Gogh Painting and a real one
- **Photograph Discriminator Model** $Discriminator_P$: Distinguishes between a generated photograph and a real one

4.2 Cycle Consistency

To enable a Cycle GAN model to successfully convert a real image into a Van Gogh-style image while maintaining consistency and coherence in the transformation process, it is crucial to establish cycle consistency. This ensures that the outputs generated by the two involved generators harmonize seamlessly without contradicting one another.

The primary objective is to achieve cycle consistency in both directions: Generator A should take an input "a" and produce an output, which, when fed into Generator B, results in an output almost identical to the original input "a." This reciprocal transformation establishes a bijection relationship between the two generator models, guaranteeing a robust and reliable mapping between the source and target domains.

4.3 Loss functions

Three primary loss categories are essential for training our model, which includes the following:

4.3.1 Adversarial Loss

$Generator_{P \to V}$ aims to generate an output that resembles a Van Gogh-style image v'.
$Discriminator_V$ aims to classify the output v' of $Generator_{P \to V}$ as a fake painting.
Similarly, $Generator_{V \to P}$ aims to generate an output that resembles an actual image p'.
$Discriminator_P$ aims to classify the output p' of $Generator_{V \to P}$ as a fake photograph.

Discriminator Loss: Penalizes the discriminator for incorrectly classifying a real image as fake or a fake as real.

Generator Loss: Penalizes the generator for producing a sample that the discriminator network classifies as fake.

Therefore, we can notice the adversarial relationship between the generator and discriminator models, which renders it challenging to train them concurrently, and they should be trained alternatively.

4.3.2 Cycle Consistency loss

The **Forward cycle** in CycleGAN involves transforming a Photograph-style image (p) into Van Gogh style (v') using $Generator_{P \to V}$, and then reverting it back to Photograph style (p') using $Generator_{V \to P}$. The goal is to enforce similarity between the original image (p) and the final output (p'), ensuring content preservation through the style transformation cycle.

Forward cycle Enforce $p \approx p'$
$Generator_{P \to V}$: Input $p \to$ Van Gogh Generator $\to$ Output v'
$Generator_{V \to P}$: Input $v' \to$ Photograph Generator $\to$ Output p'
$p \to Generator_{P \to V}(p) \to (Generator_{V \to P}(Generator_{P \to V}(p)) \approx p$

The **Backward cycle** in CycleGAN ensures similarity between a Van Gogh-style image (v) and the image obtained after a backward cycle transformation (v'). This involves transforming v to the Photograph style (p') using $Generator_{V \to P}$, and then reverting it back to Van Gogh style (v') using $Generator_{P \to V}$. The objective is to enforce $v \approx v'$, validating the consistency of the style transformation cycle and maintaining the fidelity of the original Van Gogh-style image.

Backward cycle Enforce $v \approx v'$
$Generator_{V \to P}$: Input $v \to$ Photograph Generator $\to$ Output p'
$Generator_{P \to V}$: Input $p' \to$ Van Gogh Generator $\to$ Output v'
$v \to Generator_{V \to P}(v) \to (Generator_{P \to V}(Generator_{V \to P}(v)) \approx v$

4.3.3 Identity Loss

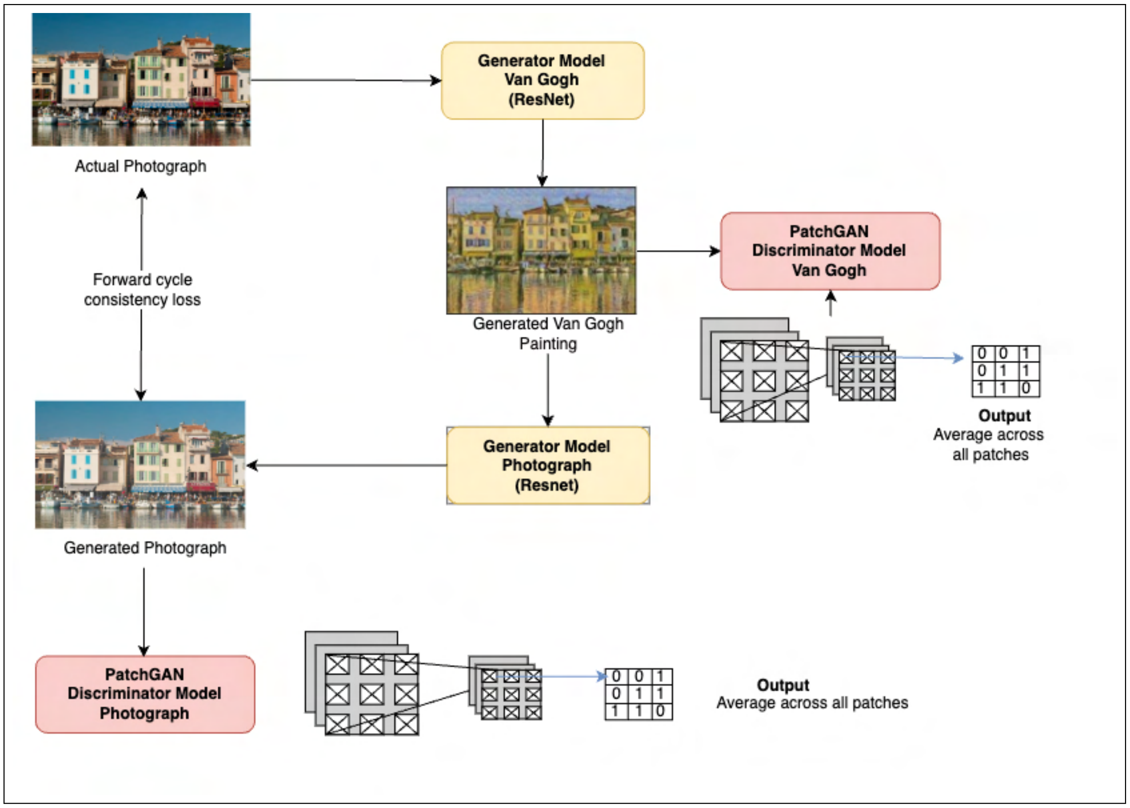
When a Van Gogh-style painting is given as an input to the Van Gogh generator, we should get back the same painting since nothing needs to be transformed. Similarly, when we input a photograph into the Photograph generator, we should get back the same photograph.

$Generator_{P \to V}$: Input $p \to$ Photograph Generator $\to$ Output p
$Generator_{V \to P}$: Input $v \to$ Van Gogh Generator $\to$ Output v

4.4 Training the model

To train the model, we backpropagate through both the discriminator and generator to obtain gradients. Since the trainable weights of the generator model are dependent upon the discriminator's weight, we start backpropagation at the discriminator which updates its weights as per the discriminator loss. It then flows back through the discriminator to the generator adjusting the weights as per the generator loss.

5 Model Architecture Figure



Model Architecture Figure: The image is a flowchart depicting a cycle-consistent Generative Adversarial Network (CycleGAN) process. It shows two paths, one starting with an "Actual Photograph" and the other with a "Generated Photograph." The first path uses a "Generator Model Van Gogh (ResNet)" to convert the photograph into a "Generated Van Gogh Painting," which is then evaluated by a "PatchGAN Discriminator Model Van Gogh." The second path uses a "Generator Model Photograph (ResNet)" to convert the generated painting back into a photo, which is then evaluated by a "PatchGAN Discriminator Model Photograph." Both discriminator models output an average across all patches, which is represented by binary numbers. The diagram also indicates a "Forward cycle consistency loss" between the actual and generated photographs, which is part of the cycle consistency that CycleGANs aim to achieve.

6 Results

6.1 Results of the U-Net Baseline Model

We detail the performance of our baseline model, which employs a U-Net architecture, in synthesizing Van Gogh-esque images from modern photographs. Over the course of 18 epochs, our model was subjected to 2000 photographic inputs and 308 authentic Van Gogh paintings to learn the characteristic style.

The model's training performance is summarized in Table 1. Notably, there was a consistent decline in generator loss, indicating a growing proficiency in creating stylized images. Correspondingly, discriminator losses for both Van Gogh-styled (disc_loss_M) and photographic (disc_loss_P) images showed a downward trend, suggesting an increased ability of the discriminator to discern between generated and real images.

Table 1: Training Performance of U-Net Baseline Model

Epoch	Learning Rate (lr)	Generator Loss (gen_loss)	Discriminator Loss Van Gogh (disc_loss_M)	Discriminator Loss Photos (disc_loss_P)
1	0.00020	7.73720	0.28478	0.28568
2	0.00020	6.10436	0.22484	0.22015
3	0.00020	5.45786	0.20315	0.20717
4	0.00020	5.26526	0.19405	0.17802
5	0.00020	5.14508	0.18903	0.13715
6	0.00019	5.02569	0.18807	0.09208
7	0.00017	4.95498	0.15957	0.07022
8	0.00016	4.96295	0.13159	0.03873
9	0.00014	4.95668	0.11565	0.03190
10	0.00013	4.84305	0.10524	0.02684
11	0.00011	4.77320	0.09958	0.01721
12	0.00010	4.73151	0.08700	0.01510
13	0.00009	4.70024	0.06692	0.01220
14	0.00007	4.65973	0.04999	0.02062
15	0.00006	4.49177	0.07534	0.01641
16	0.00004	4.35653	0.09286	0.01547
17	0.00003	4.26153	0.09589	0.02630
18	0.00001	4.07192	0.11613	0.03746

Visual outcomes, crucial for evaluating the success of style transfer models, are depicted in Figures 1-5. These figures provide a compelling side-by-side comparison of original photos (top rows) and their Van Gogh-styled counterparts (bottom rows), produced at the end of the training period.

Figure 1: Sample 1 - Photo-to-Van Gogh Translation

Figure 2: Sample 2 - Photo-to-Van Gogh Translation

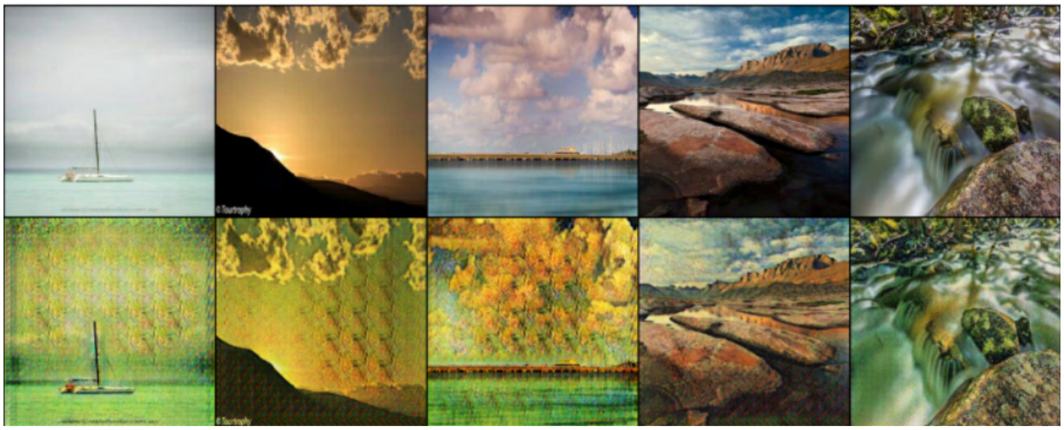


Figure 3: Sample 3 - Photo-to-Van Gogh Translation

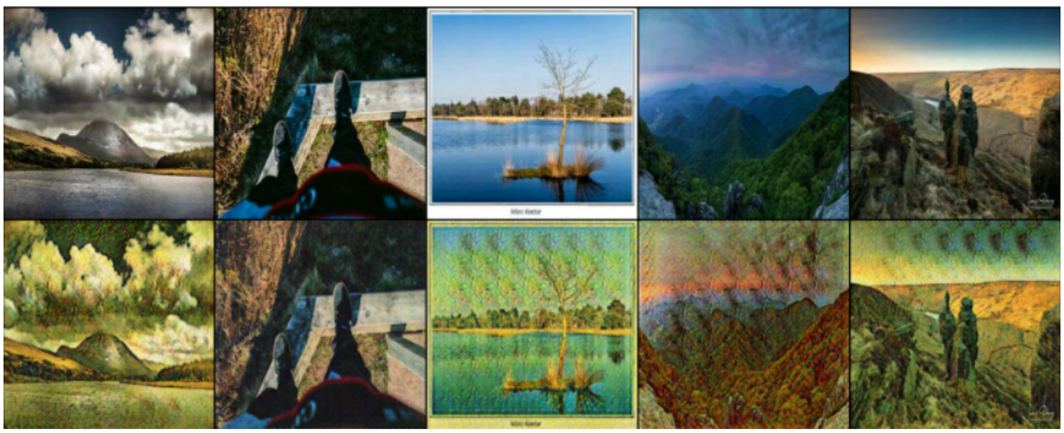


Figure 4: Sample 4 - Photo-to-Van Gogh Translation

Figure 5: Sample 5 - Photo-to-Van Gogh Translation

While the U-Net model achieved notable success in emulating Van Gogh's distinctive textures and vibrant color schemes, certain limitations were apparent. Some images demonstrate a tendency towards over-saturation and a loss of fine detail, diverging from the subtleties present in Van Gogh's original work. This was particularly evident in complex scenes where the model struggled to maintain structural integrity and color balance.

6.2 Results of the ResNet Modified Model

Over the course of 18 epochs, our modified ResNet model was subjected to 2000 photographic inputs and 308 authentic Van Gogh paintings to learn the characteristic style. After training the model, we detail the performance of our modified model in creating Van Gogh paintings from modern photos.

Table 2: Training Performance of ResNet Modified Model

Epoch	Learning Rate (lr)	Generator Loss (gen_loss)	Discriminator Loss Van Gogh (disc_loss_M)	Discriminator Loss Photos (disc_loss_P)
1	0.00020	9.21105	0.28212	0.28655
2	0.00020	7.86094	0.21818	0.21815
3	0.00020	7.28435	0.21456	0.21446
4	0.00020	7.00394	0.20824	0.20865
5	0.00020	6.76260	0.20534	0.20815
6	0.00019	6.43239	0.19604	0.20338
7	0.00017	6.24632	0.19110	0.19990
8	0.00016	6.05427	0.18666	0.19465
9	0.00014	5.85949	0.18192	0.19049
10	0.00013	5.70987	0.18201	0.18251
11	0.00011	5.59886	0.17725	0.18081
12	0.00010	5.40949	0.17662	0.17424
13	0.00009	5.29747	0.17000	0.17097
14	0.00007	5.18645	0.16274	0.16516
15	0.00006	5.10430	0.16201	0.16223
16	0.00004	4.98971	0.15814	0.15880
17	0.00003	4.89230	0.15556	0.15399
18	0.00001	4.80997	0.15138	0.15158

The training performance of the ResNet model, as detailed in Table 2, shows a steady decrease in generator loss, reflecting enhanced capability in generating images with Van Gogh-like stylization. The discriminator losses, for both the Van Gogh style (disc_loss_M) and the photographic authenticity (disc_loss_P), also decreased over time. This indicates improved discrimination by the model, enhancing its ability to differentiate between generated and real images with better accuracy.

The results of the style transfer done by the ResNet model are depicted in Figures 6-10. These figures show the original photos in the top rows and their Van Gogh-styled counterparts in the bottom rows. These images are produced after the training has finished.

Figure 6: Sample 1 - Photo-to-Van Gogh Translation

Figure 7: Sample 2 - Photo-to-Van Gogh Translation

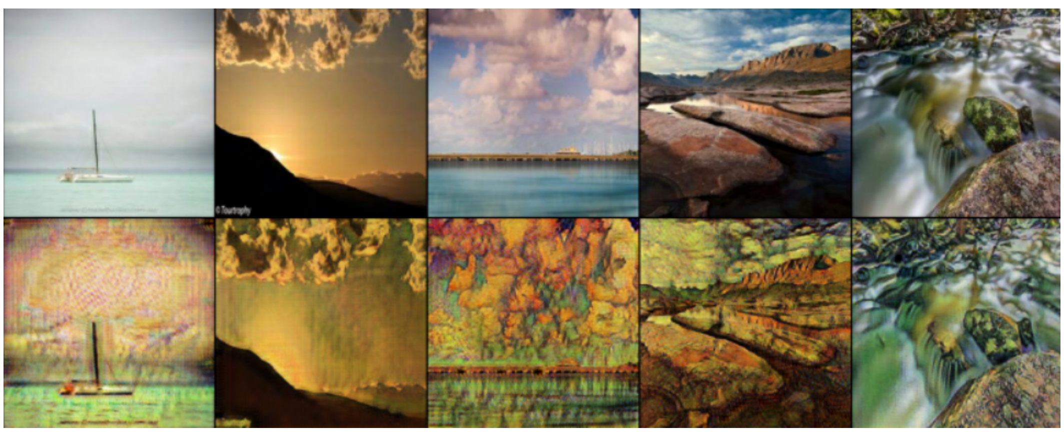


Figure 8: Sample 3 - Photo-to-Van Gogh Translation

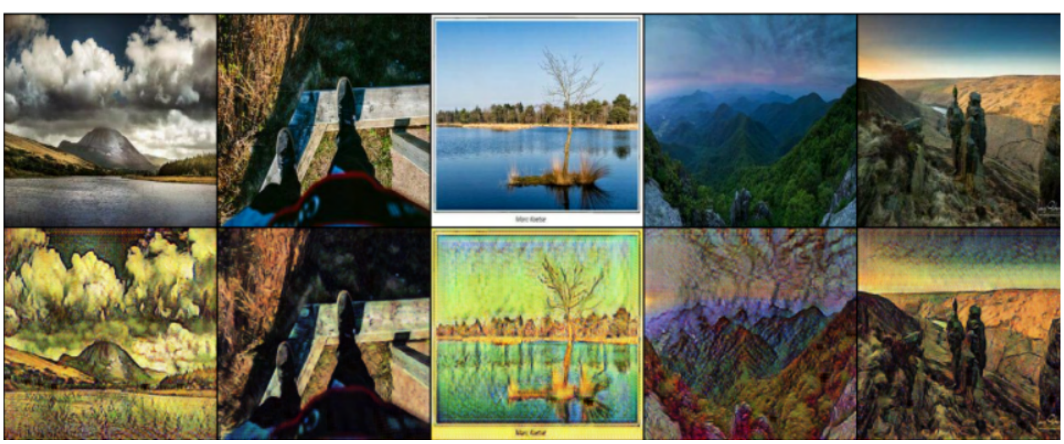


Figure 9: Sample 4 - Photo-to-Van Gogh Translation

Figure 10: Sample 5 - Photo-to-Van Gogh Translation

The ResNet model's translations appear more defined, preserving structural details with clarity closer to Van Gogh's technique. Its layers likely enable a nuanced blend of style and content, maintaining the integrity of the original images while effectively integrating Van Gogh's distinct brushstrokes and color palette.

Compared to U-Net, the ResNet model demonstrates a subtler transformation. Where U-Net tends towards less controlled coloration, potentially overshadowing finer details, ResNet maintains content fidelity alongside stylistic adaptation. This results in translations that respect the balance of form and hue characteristic of Van Gogh's paintings. Additionally, the training data indicates a discernible difference in performance. The ResNet model began with a higher generator loss (9.21105) but exhibited a steady and significant decline across epochs, indicating a consistent improvement in generating stylized images. In contrast, the U-Net model started with a lower initial generator loss (7.73720) but showed a less pronounced rate of improvement. Notably, the discriminator losses for both Van Gogh-styled (disc_loss_M) and photographic (disc_loss_P) images decreased more substantially for the U-Net model, suggesting a better ability to differentiate between real and generated images. However, the overall lower generator loss in later epochs for the ResNet model, coupled with its consistent improvement, suggests that it was more effective in closely emulating Van Gogh's distinct style. This is further supported by qualitative assessments, where ResNet-generated images displayed a greater similarity to the intricate and characteristic features of Van Gogh's artwork, achieving a balance between artistic style and photographic content.

7 Discussion

We employed a CycleGAN architecture, integrating both the U-Net generator (as our baseline model) and the ResNet model. The training process involved careful tuning of hyperparameters, including learning rates, number of residual blocks, and optimization strategies. Our dataset comprised a curated collection of Van Gogh paintings, and we utilized adversarial and cycle-consistency losses to guide the training process.

Given the generative nature of our model, we conducted a thorough analysis of the qualitative results of the generated paintings produced by the two generators. Notably, the ResNet model exhibited significantly superior performance. It showcased a heightened fidelity to Van Gogh's artistic style, accurately capturing intricate details and brush strokes in a more faithful manner compared to the U-Net generator. Upon closer examination, the ResNet-generated paintings displayed improved details and a heightened sense of realism in contrast to those generated by the U-Net model.

Van Gogh's artistic repertoire is characterized by intricate patterns and textures, and the ResNet model demonstrated a noteworthy proficiency in handling these complexities. This capability contributed to a more nuanced and richer representation of the chosen artistic style. The observed enhancements in fidelity, detail, and realism underscore the effectiveness of the ResNet model in the context of generating Van Gogh-style paintings, providing valuable insights for future developments in artistic style transfer.

In our particular scenario, the base model, U-Net, possesses a U-shaped architecture with skip connections, primarily crafted for tasks involving semantic segmentation. Nevertheless, its architectural configuration may lack the necessary depth and capacity to intricately capture the nuanced details present in Van Gogh's paintings, such as delicate brushstrokes, textures, and elaborate patterns. In contrast to ResNet, U-Net might not incorporate the depth and mechanisms for feature reuse that prove advantageous in capturing hierarchical representations and spatial information. This deficiency can lead to the compromise of fine details, resulting in an inability to faithfully reproduce Van Gogh's distinctive brushwork.

U-Net architectures may encounter challenges in generalizing effectively across a spectrum of artistic styles, constraining their adaptability to the specific characteristics intrinsic to Van Gogh's paintings. On the other hand, ResNet's demonstrated capability to generalize across diverse styles and datasets emerges as a contributing factor to its heightened performance in this context. The broader generalization scope of ResNet aids in effectively adapting to the intricacies of Van Gogh's unique artistic style, further underscoring its superiority in our comparative evaluation.

8 Limitations

The model's understanding of Van Gogh's style is limited to the data it was trained on. If the training dataset lacked diversity in terms of Van Gogh's range of styles, the model did not accurately replicate all aspects of his artistry. Data images containing drawings and sketches were omitted from the Van Gogh dataset since it differed from his usual painting style. There is always an inherent challenge to capture an artist's work within a finite dataset. Integrating techniques from 3D modeling and texture analysis could further enhance the model's ability to replicate brush strokes and texture. Also, the model's focus on Van Gogh's style might limit its versatility in adapting to other artistic styles without extensive retraining. To address these limitations and extend the capabilities of the model, NLP techniques can be incorporated to analyze the context of the scenes in the input images, enabling the model to better interpret and replicate the emotional and thematic elements in the style of Van Gogh. To accurately mimic the color and light dynamics used by Van Gogh in his paintings, more sophisticated color theory algorithms and lightning models can be used. Through transfer learning or more modular neural network designs, a more flexible architecture can adapt to various artistic styles and not just Van Gogh's.

9 Ethical Considerations

The Generative Adversarial Networks(GAN) model is used to imitate the style of the artist Claude Van Gogh. While this technology can be used creatively and provide various benefits, there are ethical concerns that may arise from its development and application:

- Intellectual Property Rights The transformation of input images to replicate the painting style of Van Gogh may lead to the infringement of Intellectual Property (IP) Rights to his work. For living artists and those whose work is under copyright, imitation of their style can lead to legal issues.
- Authenticity and Deception There's a risk that the images generated by our model could be passed off as original Van Gogh paintings or represented as new recently discovered work. This could also lead to fraudulent cases where fake art is produced and sold in markets, misleading the buyers.
- Cultural and Historical Misinterpretation If the images generated by our model are not clearly labelled as AI-generated then there is a risk of misinterpreting the historical context and significance of Van Gogh's work. The model may also unintentionally overshadow the styles of other lesser-known artists, by replicating the styles of Van Gogh.
- The GAN technology can be repurposed to generate deepfakes and be used in other unethical contexts such as creating morphed, counterfeit images.

To mitigate some of the issues related to ethical consider the following steps can be taken:

- All output images are clearly labelled as AI-generated and provide information on the model's capabilities and shortcomings. This results in clear licensing and attribution of the work of the original artist.
- To ensure the model doesn't reinforce biases, a diverse set of training data should be used.
- Guidelines for responsible use of our technology and discourage its misuse should be communicated clearly to all the users of the model.

10 Conclusion

In conclusion, this research demonstrates the potential of ResNet models in transforming images into Van Gogh-style paintings and its superior performance over the U-Net baseline in generating more defined and authentic Van Gogh-style paintings. The descending trend in generator and discriminator losses reflects the ResNet model's growing sophistication in producing images that closely mirror the intricate balance of color and form characteristic of Van Gogh's work. This advancement underscores the potential of AI models in the realm of artistry and opens pathways for future exploration. Further studies could enhance the robustness of these models, delve into the ethical implications of AI-generated art, and refine the training datasets to encapsulate a broader spectrum of artistic styles with greater precision.

References

[1] Casey Chu, Andrey Zhmoginov, and Mark Sandler. Cyclegan, a master of steganography. *arXiv preprint arXiv:1712.02950*, 2017.

[2] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. A neural algorithm of artistic style, 2015.

[3] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[4] Amy Jang, Ana Sofia Uzsoy, and Phil Culliton. I'm something of a painter myself. Kaggle, 2020.

[5] Phillip Isola Alexei A. Efros Jun Yan Zhu, Taesung Park. Unpaired image-to-image translation using cycle-consistent adversarial networks. *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[6] Kaggle. Van gogh paintings. `https://www.kaggle.com/datasets/ipythonx/van-gogh-paintings/`, 2022.

[7] Eishkaran Singh. Cyclegan implementation. Kaggle, 2023.