Computational Creativity Assignment: Generating Abstract Masterpieces using StyleGAN3 and StyleTransfer

Saiham Rahman, ID: 210769593

School of EECS
Queen Mary University of London, UK
ec21286@qmul.ac.uk

Abstract. In this paper, we present the flexibility of using limited data to generate portraits similar to the metropolitan art museum using StyleGAN3 and by using abstract paintings for neural style transfer on the generated images; we a explore a new perspective of artistically stylized abstract GAN (Generative Adversarial Networks) portrait images. To the best of our knowledge, this is the sole contributor to the idea of generating abstract portrait images trained on the museums' portrait masterpieces. In this study, we will dive into the details of the architectures of both the networks, evaluate the performances and results. We will also cover information about the museum data and abstract paintings used. Finally, we will cover some of the philosophical issues of using such system which will bring into light a higher-level computational creativity issue for us to discuss.

1 Introduction

Generative Adversarial Networks or, GAN's have made a major impact in the and has gained much popularity in the past couple of years. The main issue with GAN's is that they only produce very good results with large datasets. To produce a high-quality with very limited data has been described to be a problem in some surveys. To solve this NVIDIA has provided a novel way to train GAN's with limited datasets and has proven to be a success in their method. In this study, we utilize this architecture to generate some museum portrait images.

Neural Style Transfer of images has also gained much popularity. We pass our generated images from the GAN network to this system to generate new abstract paintings. This system takes input and style images to form a combined style-transfer image.

In this paper, we will utilize both of these systems to generate 'oil-on-canvas' style portrait images of historic museum masterpieces. We first briefly introduce the systems and go into the experimentation part where we will be able to see some of the outputs from both the systems. Lastly, we will also discuss some higher level computational creativity issues and explore how this work could be extended in future.

2 Background

Generative Adversarial Networks(GAN) have become a popular tool to generate artwork and images in the last couple of years [1]. In this creative domain of visual arts, incorporated with Artificial Intelligence, it become a continuing contribution in the field of digital art. Generative artworks are now welcomed by the public, we can see the rise of NFT's[2] [3] and how they have become a valuable asset like the artworks found in the real world. More tailored and better GAN's were created and artworks like these [4][5] have gained much value. Recently, NVIDIA[6] has introduced StyleGAN3[7] which is an iteration of StyleGAN[8] and StyleGAN2[9] with certain improvements. Utilizing StyleGAN's architecture researches have generated images[10] which amazed this field of research to an extent that, people are ready to commercialize these methods. StyletTransfer[11] has also gained much recognition[12]. This system has also contributed much in this domain. Now we can see neural style transfer images from research fields[13] to a simple filter on a selfie[14]. In this paper, we intend to fill the gap between generating portrait images and style transfer with the help of the well-established system of StyleGAN3 and StyleTransfer.

3 Dataset

Initially we thought of collecting digital versions of museum portrait artworks from the national history museum of England. Unfortunately, due to licensing issues, we have considered another dataset. The current dataset is based on the Metropolitan Museum of Art in New York[15]. They have provided open access to the public of their digital images of the portraits we require. Although we have acquired a dataset more suited to our needs provided by NVIDIA as the MetFaces dataset. [16]

This new dataset, MetFaces, is an extracted version of human faces from images from the Metropolitan Museum of Art online collection. They have used these search terms to acquire these images, 'paintings', 'watercolor' and 'oil on canvas', and were downloaded via the https://metmuseum.github.io/ API which resulted in a set of source images that depicted paintings, drawings, and statues. Although they have provided the source code to extract the faces from these images. The provided software extracts these faces via face detection and image quality metrics etc. to find the set of images which only contain human faces. To filter out low-quality photographs missed by automated filtering, a manual selection pass was done on the remaining images. Finally, the faces were cropped and aligned to create 1,336 high-quality photos with a resolution of 1024x1024 pixels. The whole dataset, including the unprocessed images, is provided here[17].

Another dataset we used for transfer learning was the FFHQ dataset[18]. The dataset contains 70,000 high-quality PNG photos with a resolution of 1024x1024 pixels, with a wide range of age, ethnicity, and image background. It also includes eyeglasses, sunglasses, caps, and other accessories well. The images were extracted from Flickr and automatically aligned and cropped.

For both the datasets, we have used the pretrained networks[19] for our training and testing cases in StyleGAN3. The FFHQ pretrained model was used for the transfer learning for the MetFaces dataset. The MetFaces pretrained network was used for generating the images we require for this project.

4 System Description

We have used StyleGAN3 for our project, which is based on Generative Adversarial Networks. Generative adversarial networks (GANs) introduced in [1] is based on adversarial game theory which is a framework created by generative models trained in a minimax zero-sum game approach [20]. The architecture described in the paper composes of two deep networks. These are,
Generator: A deep neural network which generates samples by transforming a noise variable.
Discriminator: An adversarial deep neural network, which distinguishes the generated samples from the real data distribution.
The architecture is depicted as below,

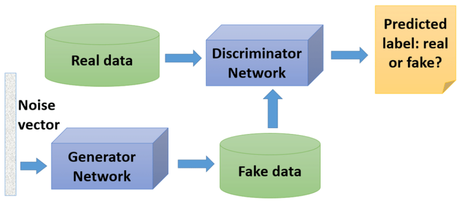

Fig 1: GAN Architecture

StyleGAN on the other hand, is based on a progressive GAN[21] architecture. The difference from that architecture is the initial 4x4x512 learned vector, and a style mapping is used to generate the latent vector, which is then fed into the generator via an AdaIN layer. The discriminator does not vary in terms of the architecture. AdaIN is a normalization method for style transfer proposed in [22]. There many improvements done over the original paper, StyleGAN2 and StyleGAN3. the general architecture of a StyleGAN is shown below,

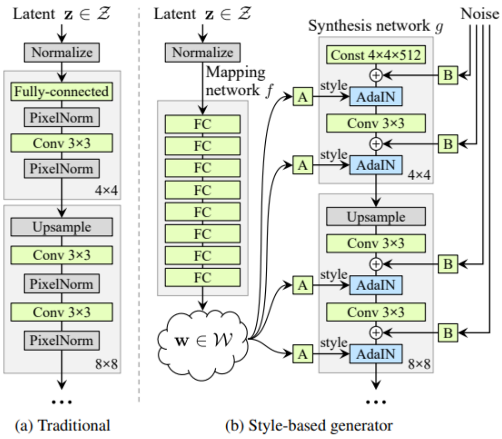

Fig 2: StyleGAN Architecture

After generating an image, this image is passed to the StyleTransfer model. The general architecture of a style transfer network consists of Convolutional Neural Networks(CNN)[23]. Using a training network VGG-16, which is a convolutional network formed of 16 layers, the style and content images are optimized upto a certain number of iterations based on a loss function. Here, in fig(3) Image x is first estimated by passing input image y through the CNN after adding an amount of white noise. Then, with the CNN weights fixed, we backpropagate this loss through the network in order to update the pixels of x. After thousands of training epochs, an x should emerge that (ideally) fits the style of y_s and the content of y_c. The architecture of the neural style transfer network is shown below,

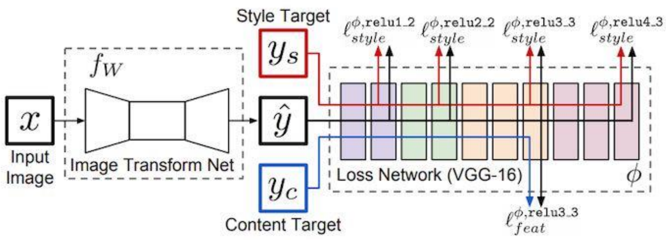

Fig 3: StyleTransfer Architecture

Other than this architecture, we have also used pretrained neural style transfer models provided by TensorFlow Hub Module[24] to generate our outputs. Combining both the systems, we are able to retrieve our desired generative style transferred output.

5 Experiments and Results

The main goal in the experimentation with StyleGAN3 model is to generate portrait images. Initially for training the data, we have used one of two different approaches for generating the model. Firstly, we trained the data without any previously trained data. Secondly, we used the FFHQ dataset as a means for transfer learning for the MetFaces dataset. To evaluate the model performances, the paper used Fréchet inception distance(FID) and two equivariance metrics EQ-T and EQ-R. Each of the equivariance metrics is expressed using the peak signal-to-noise ratio (PSNR) between two sets of images, where decibels(dB) is the unit of measurement. The PSNR between I(image reference) and K(noisy approximation) is defined by the mean squared error(MSE) as below:

$$\mathrm{MSE}_{\mathcal{D}}(I, K) = \frac{1}{\|\mathcal{D}\|} \sum_{i \in \mathcal{D}} \left(I[i] - K[i]\right)^2,$$

$$\mathrm{PSNR}_{\mathcal{D}}(I, K) = 10 \cdot \log_{10} \left(\frac{I_{max}^2}{\mathrm{MSE}_{\mathcal{D}}(I, K)} \right),$$

We trained our model on google colab using 1x NVIDIA T4 GPU or K80 GPU based on which was available at that time. Due to computational limitations, we were continuously disconnected from the server and tried to resume training from it was left off. After a few hours of training we have consumed the quota for GPU's in colab for which our trained model performances were not great. So, we opted for other pretrained networks. For this report, we have only used the pretrained network provided by NVIDIA as the MetFaces1024-r network. The images generated from that network are shown below,

Fig 4: Generated METFaces images by StyleGAN3

The images generated from the StyleGAN3 network are then passed to the StyleTransfer model. For our experiment we have only used one style image which is a painting by Wassily Kandinsky - Composition VI (1913) as shown below,

Fig 5: Style Image, Wassily Kandinsky - Composition VI (1913)

For artistic neural transfer, it is done in this way to provide consistent visual results in the specific domain of abstract artwork for portrait images. For the StyleTransfer model, we have also used evaluation metrics to calculate the loss between the reference style image, GAN image and the output image. his by calculating the mean square error for your image's output relative to each target, then take the weighted sum of these losses as the following,

$$\mathcal{L}_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{content}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{style}(\vec{a}, \vec{x})$$

Where, p, a, x are content image, style image and combined image respectively. The results generated are GAN images from random seeds from the GAN network, then passed through the StyleTransfer optimization loop for upto 500 iterations. The mean computed loss at 500 iterations for the images is 1735.24. The images are as below,

Fig 6: Style Transfer at 500 iterations

Apart from this, we have used the TF-HUB library to generate some style transfer images faster. Although the model is based in a different network configuration and provides different results, but according to the desired outcome in mind, this provides the slightly accurate representation of the desired outcome of this project.

Fig 7: Style Transfer using Tensorflow hub

6 Discussion and a Higher Level Computational Creativity Issue

In our method, to generate an expected artwork requires a lot of computationally intensive algorithms. For which, considering training in GAN and optimization in StyleTransfer, to get a desired output requires a lot of time. We also had to deviate from the original proposed dataset of the national history museum of England, due to licensing issues. Training the model was an issue for us due to computational limitations. Apart from all the limitations, we have managed to generate somewhat of the expected output. To our surprise, the generated portraits were almost like the real thing. It seemed like it was from the museum itself. After applying a style transfer on the images, the outputs we have found were astonishing. It felt like a modernized oil-on-canvas portrait of historic portraits of people from that era.

In a world where digital artwork is getting highly reputed, there are also issues following it. The first issue might be the fact that it could have an impact on how general people perceive artworks and paintings. Artists around the world and through time have made huge contributions to how we perceive artworks. They have set an example and new artists have followed into that passion. But what will happen when we put a generative artwork side-by-side an original art piece drawn by hand? There could be a chance that the generative systems set an unimaginable boundary to artists if people perceive those generative artworks to be superior in any case. Another issue which might rise is the originality of the work, generative systems and style transfer are trained/optimized on the inputs they are given. As the systems try to mimic those artworks, will there not be a question for originality of the work?

Perhaps, these issues can be countered considering the generative systems as nothing but a collaborative way of generating digital artwork. Generative arts when displayed to any public domain can include the details and the descriptions of the artworks it was based upon. This is will of course be a discussion of the arts society, but maybe something to be considered about, as these artworks are becoming more public.

7 Conclusions and Future Work

Our method of utilizing both these models to generate these masterpieces can be considered a contribution towards the domain of visual arts. In this project, we take the advantage of two systems, StyleGAN3 and StyleTransfer. Generative systems has contributed much in the past couple of years in the field of digital artwork. Now utilizing style-transfer we have added a new way of generating abstract portrait paintings of museum images. We have also covered a higher level computational creativity issue in this domain and set an example for newer works to come, which would utilize these systems together. For our future work, we are looking into the evaluation of these generated style transferred outputs apart from the original metrics both quantitively and qualitatively. We plan to merge both these systems into one generic system to reduce some of the complexity of the algorithms. We will also generate more images using different parameters values and fine-tuning for specific exhibitions. Finally, we will explore more datasets other than museum images for generating artwork.

References

1. Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). Generative Adversarial Nets (PDF). Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014). pp. 2672–2680.
2. CNN Business, Jazmin Goodwin. “What Is NFT? Non-Fungible Tokens Explained - CNN.” *CNN*, edition.cnn.com, 17 Mar. 2021, https://edition.cnn.com/2021/03/17/business/what-is-nft-meaning-fe-series/index.html.
3. Nadini, M., Alessandretti, L., Di Giacinto, F. et al. Mapping the NFT revolution: market trends, trade networks, and visual features. Sci Rep 11, 20902 (2021). https://doi.org/10.1038/s41598-021-00053-8
4. Kastrenakes, Jacob. “Beeple Sold an NFT for $69 Million - The Verge.” The Verge, www.theverge.com, 11 Mar. 2021, https://www.theverge.com/2021/3/11/22325054/beeple-christies-nft-sale-cost-everydays-69-million.
5. Kastrenakes, Jacob. “Who’ll Make the next Million-Dollar NFT? - The Verge.” The Verge, www.theverge.com, 26 Aug. 2021, https://www.theverge.com/2021/8/26/22639353/cool-cats-bayc-generative-nft-profile-avatar.
6. “NVIDIA Research Homepage.” Research, Publications & Journals | NVIDIA, www.nvidia.com, https://www.nvidia.com/en-us/research/. Accessed 20 Apr. 2022.
7. Karras, Tero, et al. “Alias-Free Generative Adversarial Networks.” ArXiv:2106.12423 [Cs, Stat], Oct. 2021. arXiv.org, http://arxiv.org/abs/2106.12423.
8. Karras, Tero, et al. “A Style-Based Generator Architecture for Generative Adversarial Networks.” ArXiv:1812.04948 [Cs, Stat], 3, Mar. 2019. arXiv.org, http://arxiv.org/abs/1812.04948.
9. Karras, Tero, et al. “Analyzing and Improving the Image Quality of StyleGAN.” ArXiv:1912.04958 [Cs, Eess, Stat], Mar. 2020. arXiv.org, http://arxiv.org/abs/1912.04958.
10. Chong, Min Jin, et al. “StyleGAN of All Trades: Image Manipulation with Only Pretrained StyleGAN.” ArXiv:2111.01619 [Cs], Nov. 2021. arXiv.org, http://arxiv.org/abs/2111.01619.

11. Gatys, Leon A., et al. “A Neural Algorithm of Artistic Style.” ArXiv:1508.06576 [Cs, q-Bio], Sept. 2015. arXiv.org, http://arxiv.org/abs/1508.06576
12. “Neural Style Transfer - Wikipedia.” Neural Style Transfer - Wikipedia, en.wikipedia.org, 13 Feb. 2019, https://en.wikipedia.org/wiki/Neural_style_transfer.
13. Jing, Yongcheng, et al. “Neural Style Transfer: A Review.” ArXiv:1705.04058 [Cs, Eess, Stat], Oct. 2018. arXiv.org, http://arxiv.org/abs/1705.04058.
14. Jeroen. “Selfie2art—Neural Style Transfer.” Medium, blogemtech.medium.com, 30 Apr. 2019, https://blogemtech.medium.com/selfie2art-neural-style-transfer-6af144cc493.
15. Metmuseum.org. n.d. [online] Available at: <https://www.metmuseum.org/about-the-met/policies-and-documents/image-resources> [Accessed 20 April 2022].
16. Karras, Tero, et al. “Training Generative Adversarial Networks with Limited Data.” ArXiv:2006.06676 [Cs, Stat], Oct. 2020. arXiv.org, http://arxiv.org/abs/2006.06676.
17. NVlabs. “GitHub - NVlabs/Metfaces-Dataset.” GitHub, github.com, 16 Sept. 2021, https://github.com/NVlabs/metfaces-dataset.
18. NVlabs. “GitHub - NVlabs/Ffhq-Dataset: Flickr-Faces-HQ Dataset (FFHQ).” GitHub, github.com, 8 Mar. 2022, https://github.com/NVlabs/ffhq-dataset.
19. “StyleGAN3 Pretrained Models | NVIDIA NGC.” NVIDIA NGC Catalog, catalog.ngc.nvidia.com, https://catalog.ngc.nvidia.com/orgs/nvidia/teams/research/models/stylegan3. Accessed 20 Apr. 2022.
20. Osborne, Martin J., and Ariel Rubinstein. A Course in Game Theory. Cambridge, MA: MIT, 1994. Print. ISBN 9780262150415.
21. Progressive Growing of GANs for Improved Quality, Stability, and Variation(https://arxiv.org/abs/1710.10196
22. Xun Huang, Serge Belongie, Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization,ICCV2017
23. Jing, Yongcheng, et al. “Neural Style Transfer: A Review.” ArXiv:1705.04058 [Cs, Eess, Stat], 7, Oct. 2018. arXiv.org, http://arxiv.org/abs/1705.04058.
24. Golnaz Ghiasi, Honglak Lee, Manjunath Kudlur, Vincent Dumoulin, Jonathon Shlens. Exploring the structure of a real-time, arbitrary neural artistic stylization network. Proceedings of the British Machine Vision Conference (BMVC), 2017.