GENERATING MONET STYLE PAINTINGS

EE259-Spring 2021

Rojin Zandi
0014491256

Contents

1. Data Analysis

The development of machine learning techniques has attracted attention of scientists and artists to art area, including painting, music, poetry, etc. Their goal is to implement algorithms for generating new masterpieces or improving the quality of old and original ones. Also, there are high quality datasets related to art, such as Art Images: Drawing/ Painting/ Sculptures/ Engravings, Best Artworks of all time, and Chinese Paintings. The Best Artwork of all time includes paintings of 50 famous painters such as Salvador Dali, Vincent Van Gogh, and William Turner. The Chinese Painting dataset have been created by Guanyang Wang, Ying Chen, and Yuan Chen who have applied WGANs on their dataset to generate Chinese paintings.

The dataset of *I'm something of a painter myself* challenge includes 300 paintings of Oscar-Claude Monet, who was a French painter. There are also 7028 photos that we aim to add Monet Style (French Impressionism) to these images and generate new paintings while keeping the originality of the image. As shown in the figure, the size of images and paintings are 256 x 256 x 3.

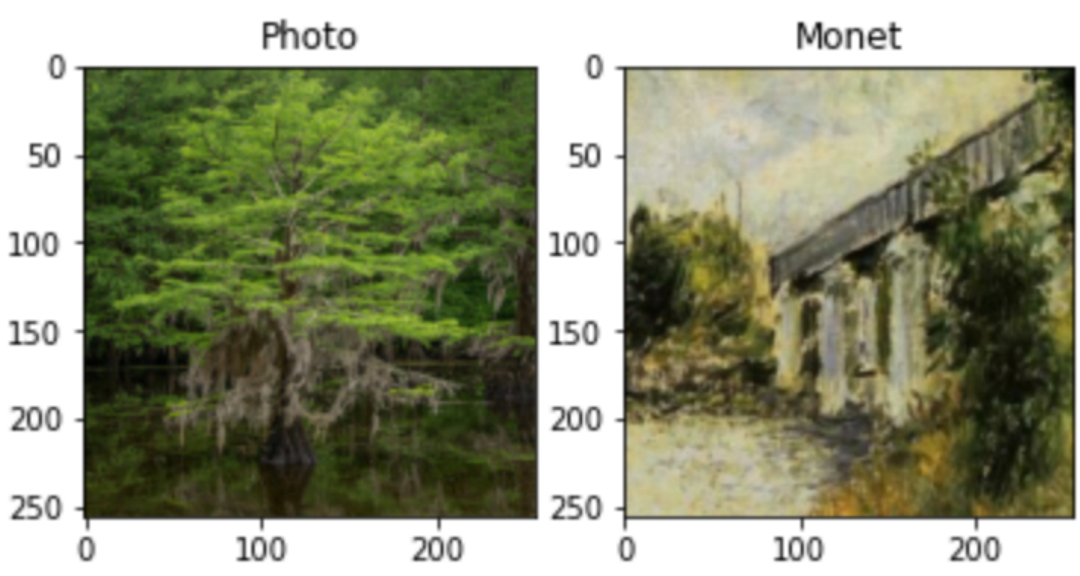


Figure 1.1: A photo and a painting from the Monet dataset.

As mentioned above, the images and paintings in the dataset contain three spectral layers, known as RGB. In order to study the dataset, we have plotted histogram of each color channel and visualized individually (Figure 1.2).

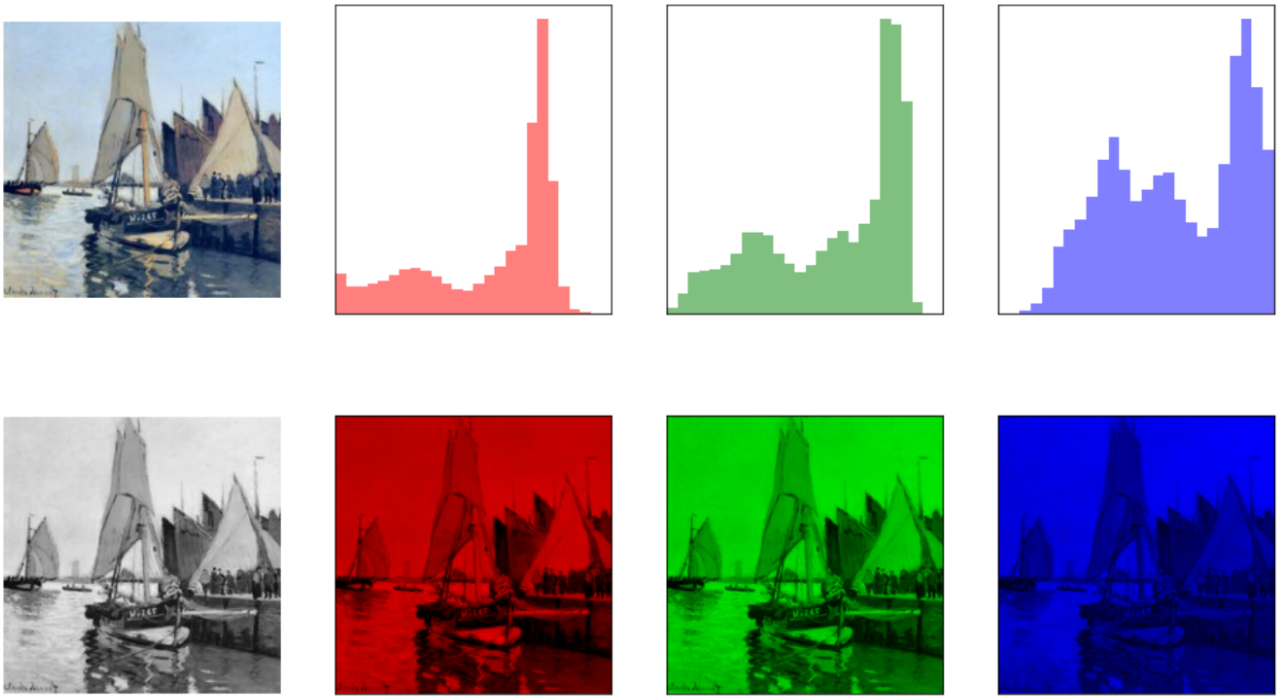


Figure 1.2: Individual channel visualization of a Monet painting

2. Related Works

Using Generative learning, researchers have generated variety of works of art. In 2017, Tikhonov and Yamshchikov [1] applied Variational Autoencoders (VAE) to generate music. They combined recurrent neural networks and VAEs to generate monotonic music. The structure of VAEs encodes the data and then using Gaussian distribution, it can generate new artwork. Generative learning methods have been also applied to paintings. Mao Li and Jiancheng Lv [2] used K-means algorithm and collection of color blocks to generate Abstract paintings. Another area of interest in applying machine learning to art is transferring painting styles to natural images. The difference between transferring and generating is saving the originality of the natural image. In the previous method [2], the painting had to follow the aesthetic rules, but in style transferring the generated painting must be as similar as possible to the original image. In area of transferring style to a new set of data, Generative Adversarial Networks (GANs) are popular.

GAN is a neural network class that can generate new data from the prior given data to the network. To generate the new data, there are two neural networks known as Generator (G) and Discriminator (D), which contest each other to win a game [3]. The G endeavors to produce fake data, and the D efforts to distinguish the generated data. This method's significance is the dynamic update and improvement of the G and D in the contest.

As mentioned above, saving the originality of the image is as important as high-quality style transferring. The problem with GANs is that they have only one loss function which focusses on style transferring, so we need a generative network with two loss functions. Cycle-Consistent GAN [4] answers this need by converting back the generated painting to the original image. The key point of this method is the cycle consistency loss which pushes network to reverse the mapping.

In section 4, we will study the application of some of mentioned methods.

3. Preprocessing

Preprocessing is a set of activities done on dataset in order to provide a clean and suitable data for the learning algorithm. These activities include normalization, reduction, and feature selection. In this project, we have different models, and the preprocessing are also different. We have used normalization and feature selection. Also, in the last model, the data is augmented by resizing, cropping and rotating.

4. Implemented Models and Improvements

In this section, we will review the generative methods that we have used to generate Monet style paintings and show the advantage of each model. As we move further, the loss value decreases and the Kaggle ranking gets better.

4.1 Variational Autoencoder (VAE)

Variational autoencoders aim to approximate the data density and produce a mean μ and a standard deviation σ. After finding μ and σ, it generates a Gaussian distribution and takes samples from the distribution. Then the samples are given as input to the decoder.

The encoder and decoder in our VAE use ReLU activation function and it has been optimized by Adam method with 1×10^{-4} as learning rate and 1×10^{-5} as the decay. After 20 epochs the reconstruction error is 1260262.57, which is extremely high.

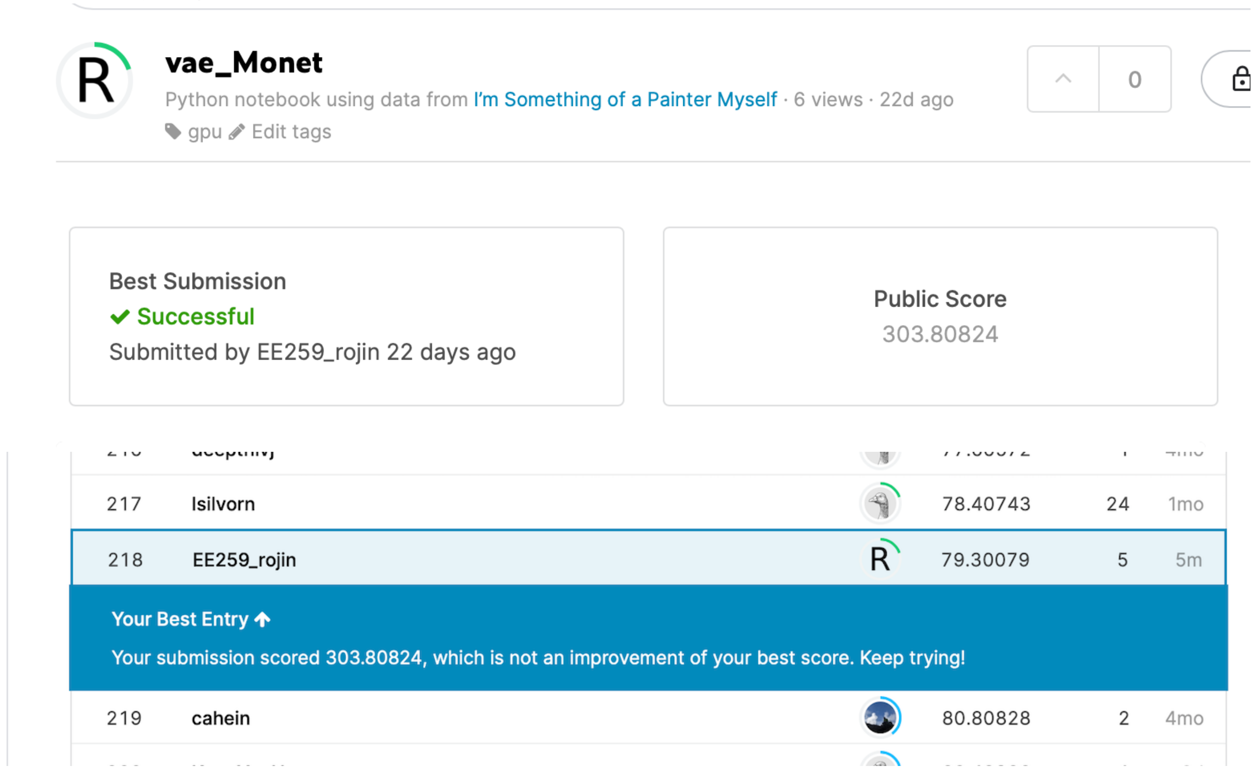


Figure 4.1.1: Kaggle score of VAE submission. Unfortunately, there was no ranking because I had submitted a better model with better score, so the leaderboard showed me my previous ranking. I have also mentioned it my submissions.

4.2 Generative Adversarial Network

As mentioned in section 2, GANs are made by two neural networks: Generator (G) and Discriminator (D) [3]. In this project, G takes the Monet paintings and aims to apply the style to the natural images, on the other hand D, which is a binary classifier, tries to classify the painting and the fake ones. There is a match between two networks that improves the performance of both. Considering that each network is being optimized to win this match, the loss function is a minmax game:

$$\min_G \max_D V(G,D) = \min_{\theta_g} \max_{\theta_d} [E_{x\sim p_{data}} logD_{\theta_d}(x) + E_{x\sim p_{(z)}} \log\left(1 - D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)],$$

where θ_g and θ_d are parameters of generator and discriminator, respectively. Notice that this dynamic architecture brings some challenges such as mode collapse and oscillating parameters. The generator in our model contains 8 down-sampling and 7 up-sampling layers, a random initializer and the last layer which is the convolutional layer with 2x2 strides and tanh activation function. The discriminator also contains the normal initializer for weights and uses batch normalization and leaky ReLU for classification. The optimizer in G and D is Adam with $2\text{x}10^{-4}$ learning rate and 0.5 decay. The model has been trained for 10 epochs. Figure 4.2.1 shows the result of this model.

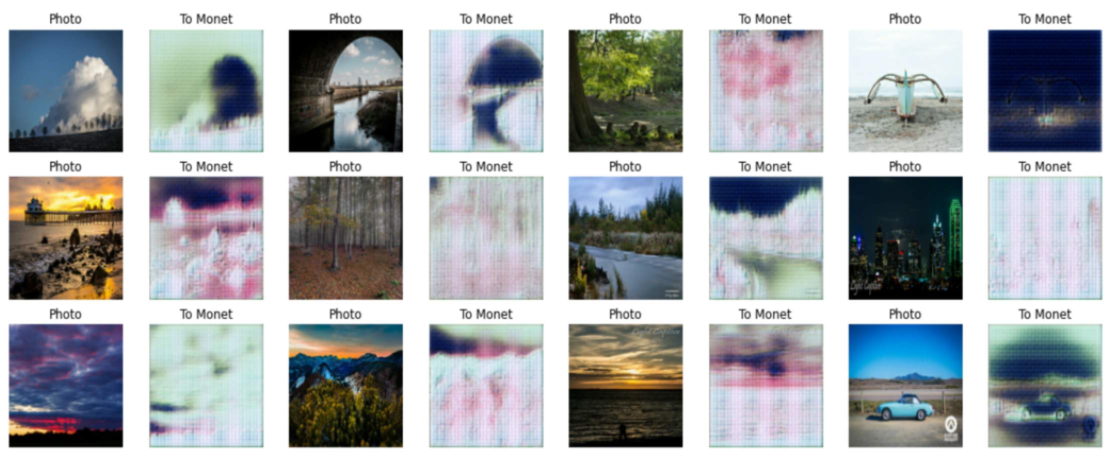


Figure 4.2.1: The original photo and generated Monet Style with GAN

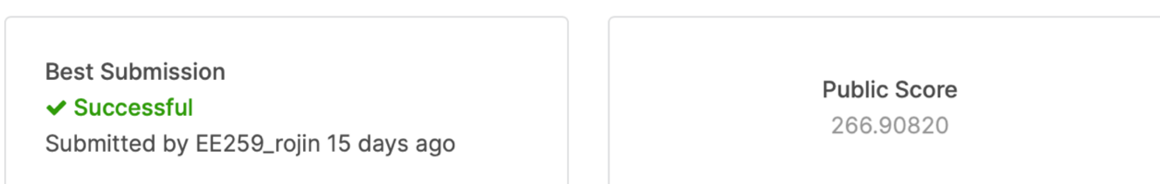


Figure 4.2.2: Kaggle score of GAN submission. Unfortunately, the same issue with this model, but the score has improved.

4.3 Cycle-consistency GAN

In 2017, Cycle-consistency GAN was introduced for unpaired image to image translation [4]. The significant point about this network is recycling the generated painting to the original image. We mentioned there are two important factors in this project:

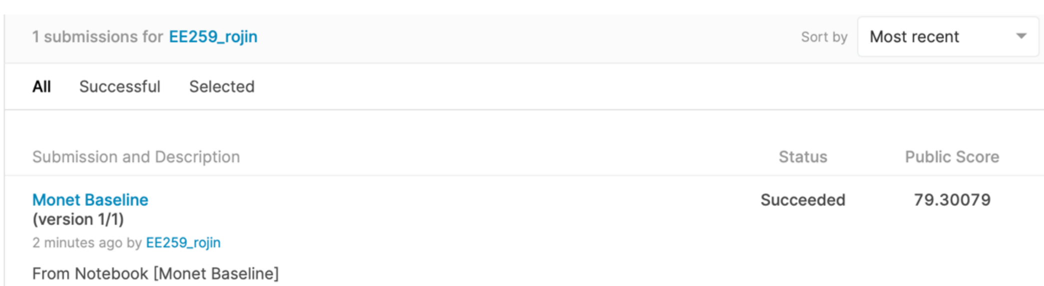
- Applying Monet style to photos
- Keeping the originality of the image

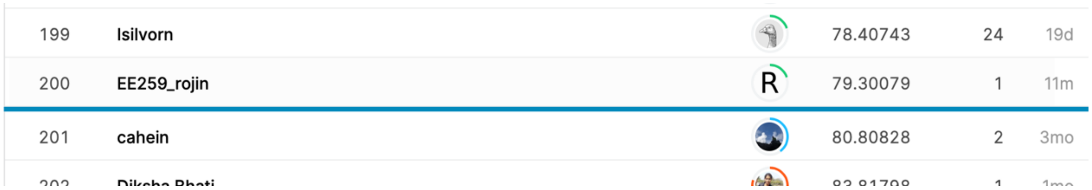
But in previous methods, there was not any control on the second factor. The cycle GAN adds another loss function which is similar to the GAN loss function, but the new function focusses on the original images and translated images from the fake paintings. So, there are two generators and two discriminators, and each one of them return one loss value.

The generator architecture is similar to the GAN generator in previous model, but the discriminator contains instance normalization by gamma initializer and the loss is Binary Cross Entropy [5]. The general GAN model contains 4 models with 4 optimizers, and they are Adam with same values in GAN model. In the first try, there was 5 epochs, and the loss values of this model are shown in table 4.3.

	Monet Gen.	**Monet Disc.**	**Photo Gen.**	**Photo Disc.**
Loss	3.376	0.629	3.66	0.561

Table 1: Loss values of cycle GAN 1





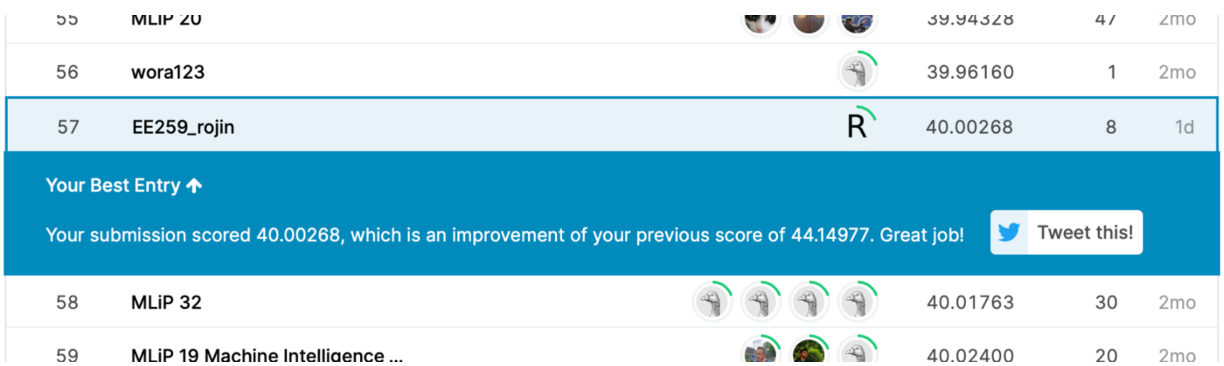
199	Isilvorn	78.40743	24	19d
200	EE259_rojin	79.30079	1	11m
201	cahein	80.80828	2	3mo
202	Diksha Bhati	[illegible]	1	[illegible]

Figure 4.3.1: Kaggle score and ranking of cycle GAN 1 submission.

Figure 4.3.2: The original photo (left) and generated painting (right)

Considering that the dataset contains only 300 paintings, we have augmented the data to improve the performance of the cycle GAN. Notice that small dataset may cause overfitting of the discriminator. The images have been resized, randomly cropped and randomly flipped. We have also increased the number of epochs to 30. The ranking in leaderboard proves the improvement of the new model.

**The architecture of generators and discriminators of last model are shown in the Appendix.*



55	MLiP 20	39.94328	47	2mo
56	wora123	39.96160	1	2mo
57	EE259_rojin	40.00268	8	1d
58	MLiP 32	40.01763	30	2mo
59	MLiP 19 Machine Intelligence ...	40.02400	20	2mo

Figure 4.3.3: Kaggle score and ranking of cycle GAN 2 submission.

5. Results and Conclusion

In this project, we have applied three different generative learning methods to generate Monet style painting. The first model is a variational autoencoder which performed weakly and reconstruction error was high, subsequently the Kaggle ranking, and score were not satisfying. In the next step, we used GAN for painting generation, which was an improvement over the previous model, but as shown in figure 4.2.1, the generated paintings were not similar to the original photos. To solve this issue, cycle-consistency GAN was applied, and the performance improvement was dramatic. The Kaggle score changed by 187 points and the ranking was 218. In order to decrease the loss values, we augmented the input data and increased the number of epochs, and as anticipated the Kaggle ranking and score changed to 57 and 40, respectively.

Finally, after comparing different implemented models in previous sections we have observed:

- Cycle-consistency GAN can control the similarity of the generated data to the original one, which is an important factor in this study
- Generative adversarial networks are more common in image-generation based problems compared to autoencoders
- Small datasets in GANs may cause overfitting in discriminators
- In the case of small dataset, data augmentation in neural networks can improve the model efficiency

For further study, we suggest Cycle-GAN with two-objective dual head discriminator [6], which prevents the overfitting risk.

Appendix

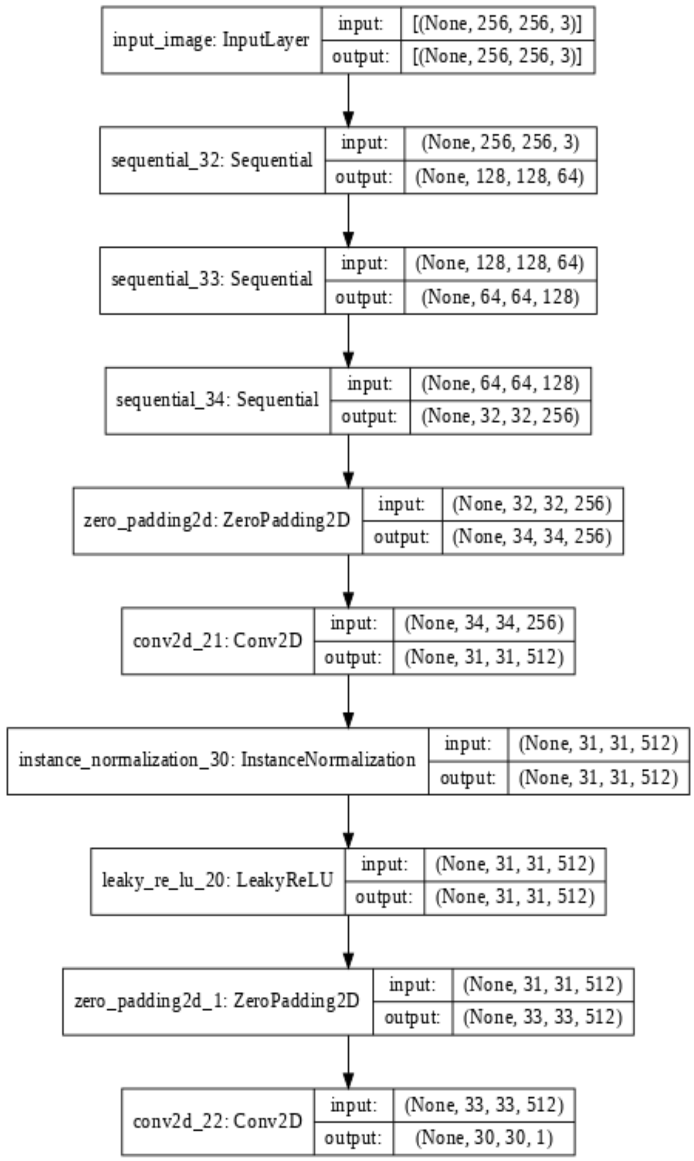


Figure 1: Architecture of discriminator of Monet and photo

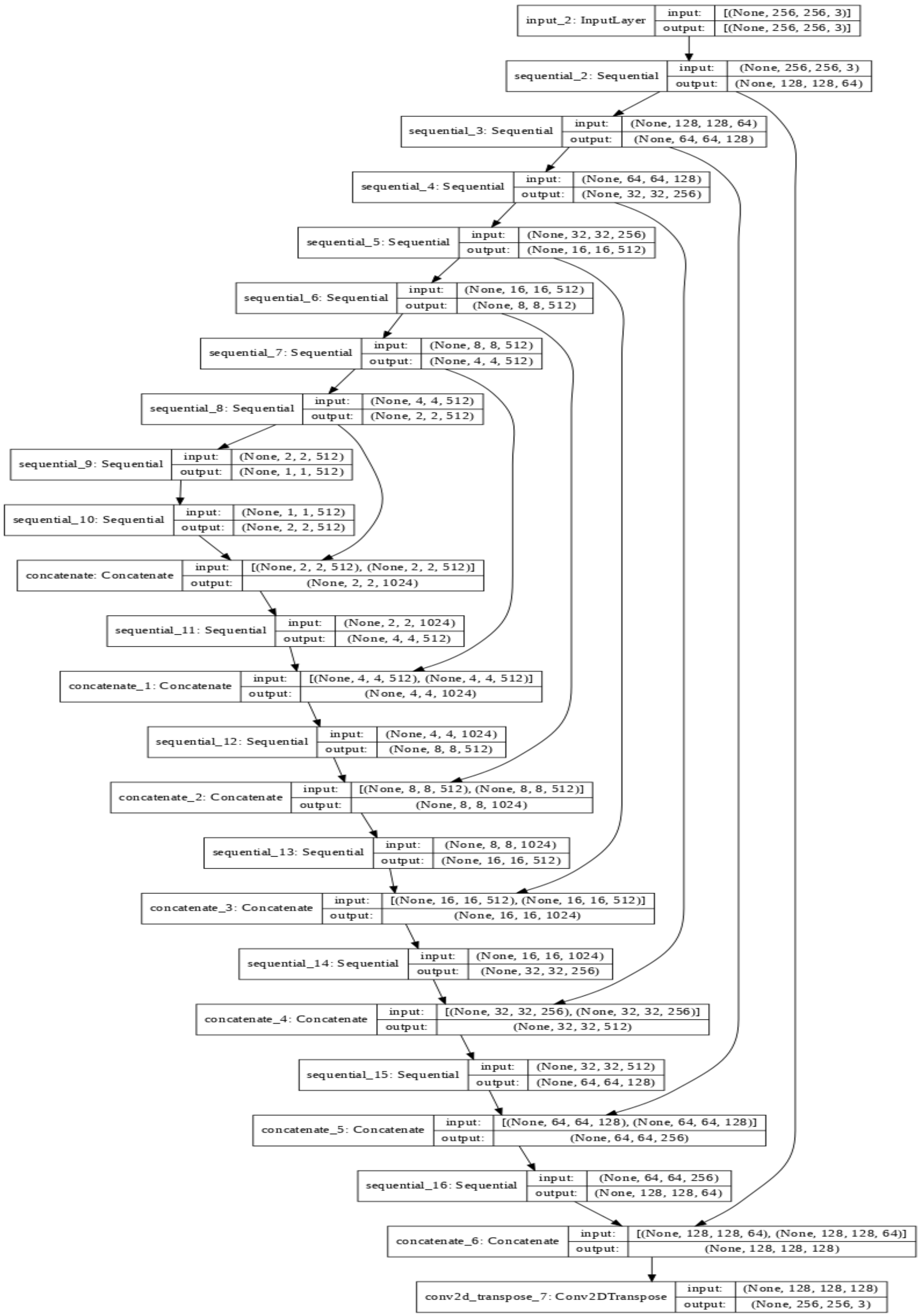


Figure 2: Architecture of generator of Monet and photo

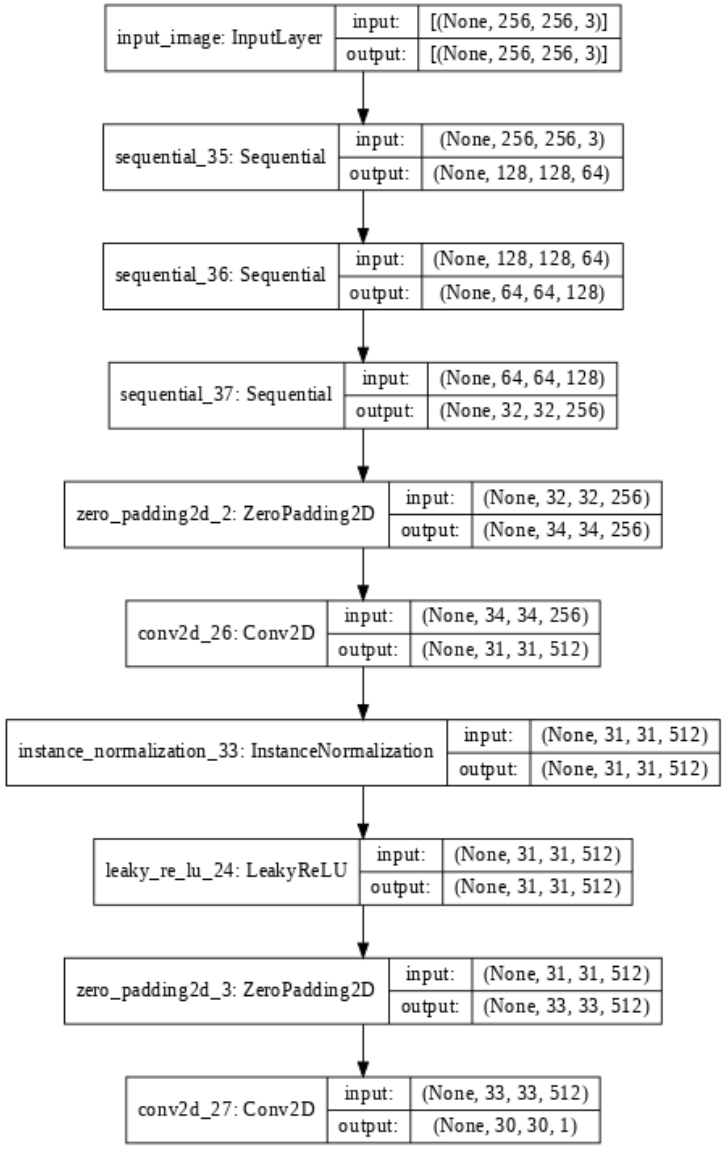


Figure 3: Architecture of discriminator of generated photo and photo

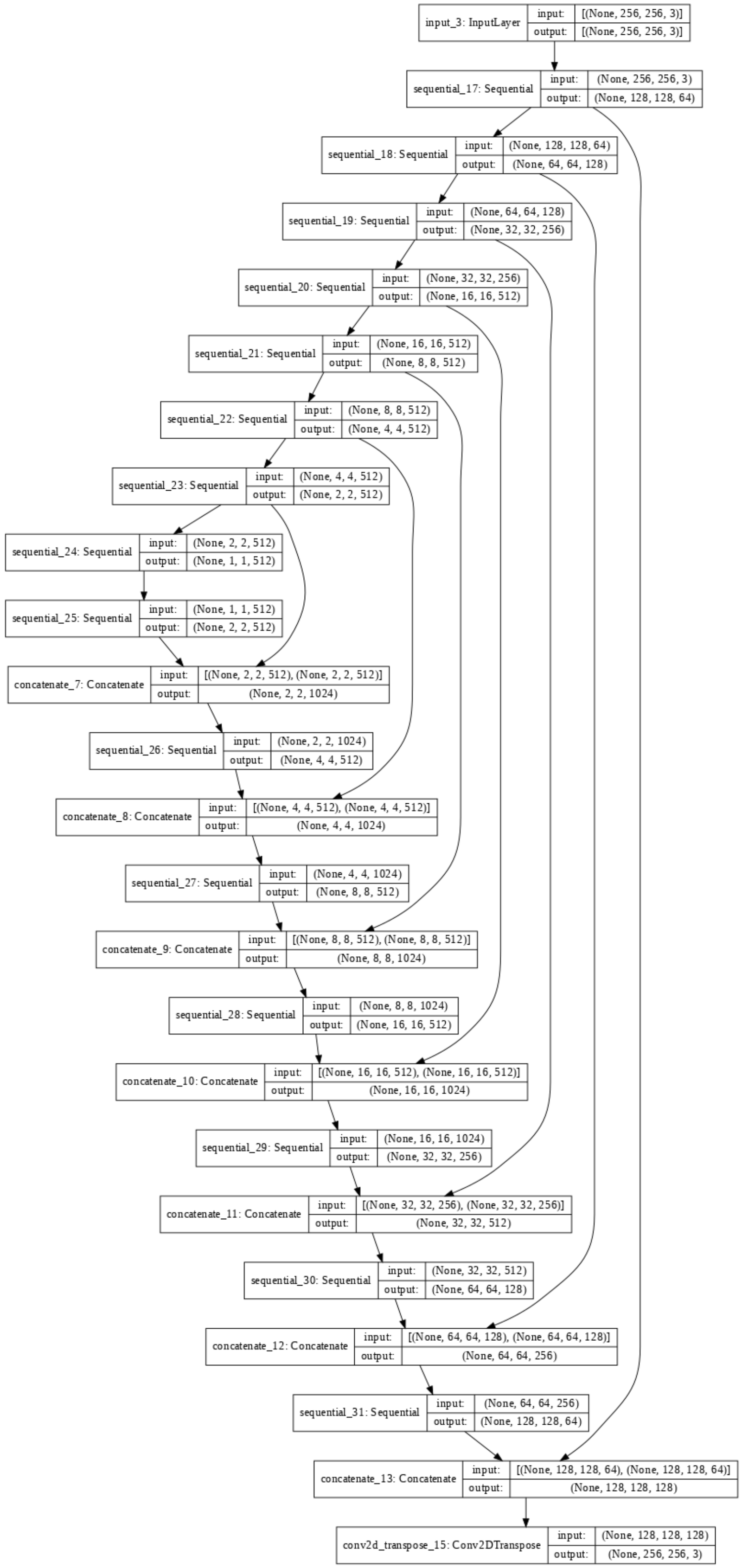


Figure 4: Architecture of generator of generated photo and photo

References:

[1] Alexey Tikhonov and P. Yamshchikov, Ivan, "Music generation with variational recurrent autoencoder supported by history," *arXiv preprint,* 2017.

[2] Mao Li1, Jiancheng Lv, Jian Wang, Yongsheng Sang,, "An Abstract Painting Generation Method Based on Deep Generative Model," *Springer Nature,* 2019.

[3] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, "Generative Adversarial Networks," *Advances in Neural Information Processing Systems 27 (NIPS),* 2014.

[4] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *Proceedings of the IEEE International Conference on Computer Vision (ICCV),* pp. 2223-2232, 2017.

[5] "CycleGAN is obsolete," May 2021. [Online]. Available: https://www.kaggle.com/unfriendlyai/cyclegan-is-obsolete.

[6] "Two-objective discriminator," 2021. [Online]. Available: https://www.kaggle.com/unfriendlyai/two-objective-discriminator.