Cycle Generative Adversarial Nets for Style Transfer of Monet Paintings

Solmaz Mohammadi
solmaz.mohammadi@studenti.unipd.it

Nicla Faccioli
nicla.faccioli@studenti.unipd.it

Abstract

This article aims at learning the style of the impressionist painter, Claude Monet. Since the generative adversarial model has its limitations when it comes to more complex computer vision problems, we will be using a variation of generative adversarial models, cycle consistent GAN to capture the "impression" of an artist from their surroundings. CycleGAN is an unpaired image to image translation using cycle consistent adversarial nets.

1. Introduction

The focus of this paper is to reproduce the style of Claude Monet, the influential impressionist painter in the nineteenth century. Claude Monet's style is unique and recognizable by his short and thick strokes that capture the essence of the subject. Monet's work is particularly interesting because of the strong logarithmic correlations within his paintings that makes it possible to generate new Monet-like images. [3] This project attempts to transfer the painter's style into existing real-life photos. Style transfer is a computer vision technique that aims to combine two images together. In particular, given a content image (i.e. is interesting for its content) and a style image (i.e. is interesting for its style), the task is to blend them together in order to obtain as output a third image which will represent the content of the first image in the style of the second one. Therefore, we aim to learn a mapping function G: X⟶Y such that the distribution of G(X) is inseparable from the disribution of Y using the adversarial loss (i.e. the discriminator). To further constraint the generator, we pair it with an inverse map F:Y⟶X and use cycle consistency loss to enforce F(G(X))∼X. [9]

2. Related Work

In **Generative Adversarial Nets (GANs)**, two models are trained simultaneously: a generative model G the captures the data distribution, and a discriminative model D that estimates the probability that a sample is from the original data rather than generated by G [1]. The training procedure for G is to maximize the probability of D making a mistake. Therefore, it corresponds to a 2-player minimax problem in which generator tries to "trick" the discriminator. This framework is then the core idea of many generative model applications including cycle consistent GANs. However, the simple GAN is under-constrained for style transfer problem [9]. We apply a cycle consistency loss objective to further condition GAN.

Paired Image-to-Image Translation: In some applications the aim is not to generate the image from scratch but rather, "translate" an image from a domain X, to another domain Y. This can be done by conditioning GAN to learn a dataset of paired images and the function "translating" them. This approach introduces cGAN, where the given the input, the output is conditoned by a set of constraints. The conditional GANs learn a mapping from observed data x and random noise vector z, to the output y, G:x,z→y [2].
cGANs in the context of image translation can be generalized to cover various problems. Image prediction from a map [8], future frame prediction [5], and image generation from sparse annotations are some of the examples [7]. Image-to-Image mappings are also the core idea for style transfer problem, but with GAN applied unconditionally [4].

3. Dataset

The dataset was collected from Kaggle, a platform in which competitions and challenges are proposed and where we got the idea for this project. The dataset is composed by:

- A set of 300 Monet paintings sized 256x256 in both JPEG and TFRecord format;
- A set of 7028 photos sized 256x256 in both JPEG and TFRecord format.

We used the Monet paintings to train our model so that it could learn Monet style and then we used the photos to apply the learned style.

4. Method

In our problem given the real-life images $x_i \in X$ and monet paintings $x_i \in Y$ we aim to learn mapping functions between the two domains X and Y, donating the data distribution as $x \sim p_{data}(x)$ and $y \sim p_{data}(y)$. Our model captures the mapping between domains using two conditionally generative [2] models G : X $\longrightarrow$ Y and F : Y $\longrightarrow$ X. Additionally, we also include two adversarial discriminators D_X and D_Y where task of D_X is to discriminate between $\{$x$\}$ and translated images $\{$F(y)$\}$. Similarly, D_Y aims to distinguish between $\{$y$\}$ and $\{$G(x)$\}$ [9]. The objective of our work is to minimize two losses: *adversarial loss* [1] which helps to close the gap between the data distribution of target domain and the generated images; and *cycle consistency loss* to keep the results of the two mapping functions consistent.

4.1. Adversarial Loss

To learn the distribution of $x \sim p_{data}(x)$, we first define a prior on input noise variables $p_z(z)$ and define a mapping G:z $\longrightarrow$ that can be trained using a convolutional neural network [6]. We also define a discriminator D (also trained using a CNN) that instead of a single scalar [1], outputs an image with decreased dimentionality where pixels that are more likely to be in data distribution $p_{data}(x)$ have higher values. We train both models simultaneously in a two-player minimax approach that can be expressed as: [formula1] The adversarial loss, maximizes the probablity of discriminator D differentiating training examples and generated samples

4.2. Cycle Consistency Loss

4.3. Full Objective

4.4. Implementation

5. Experiments

References

[1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2016.

[3] Jaron Kent-Dobias. Log-correlated color in monet's paintings, 2022.

[4] Chuan Li and Michael Wand. Precomputed real-time texture synthesis with markovian generative adversarial networks, 2016.

[5] Michael Mathieu, Camille Couprie, and Yann LeCun. Deep multi-scale video prediction beyond mean square error, 2015.

[6] Keiron O'Shea and Ryan Nash. An introduction to convolutional neural networks, 2015.

[7] Scott Reed, Zeynep Akata, Santosh Mohan, Samuel Tenka, Bernt Schiele, and Honglak Lee. Learning what and where to draw, 2016.

[8] Xiaolong Wang and Abhinav Gupta. Generative image modeling using style and structure adversarial networks, 2016.

[9] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks, 2017.