Proposal for ECE285 project

Hongyuan Du
PID: A53285631
Department of Electrical and Computer Engineering
UC San Diego
`hdu@eng.ucsd.edu`

Abstract

We are going to explore a new way that associate Generative Adversarial Networks (GANs) with style transfer techniques. We expect GANs with enhanced capacity can generate not only target contents, but also specific styles. Our goal is to learn a mapping $G : z \to I$, where z is the domain of samples from low-dimensional latent variables, and I is the domain of target images. The distribution of images from $G(z)$ should be indistinguishable both from the distribution of target content C and the distribution of specific style S.

1 Background

A lot of brilliant works have been done on GANs and style transfer techniques and have achieved many impressive results. One of the most amazing works is CycleGAN proposed by Zhu et al., in which the authors implement a new structure of GANs to solve image-to-image translation problems. Our idea is inspired by CycleGAN, but what we will do is to come back to the origin of GANs — to generate target samples from low-dimensional latent variables. We expect our generate network G can create new images, not translate an image to another. This technique can be very useful in future artistic creation tasks, that just by manipulating low-dimensional parameters we can create artworks with desired contents and styles.

2 Method

In our project, we are going to propose new strategies for training GANs to enhance their capacity in capturing the distributions of contents as well as styles. So far, we have a very simple idea that firstly train the GAN with content images, and then enhance it by training with a style loss network. This idea is inspired by fast-neural-style technique proposed by Johnson et al., in which the authors use a CNN to approximate the style transfer mapping. In our work, we replace the CNN by GAN to achieve our goal. This change leads to different training strategy which we would like to describe it in detail in our final paper. We have done experiments that can generate promising images in 64x64 resolution which attached in appendix. We are still working on the structure of our network and training strategy, and expect to get better results that generate higher resolution images.

We also have a more advanced idea that we want to improve our work by the idea from InfoGAN proposed by Chen et al., which make our latent variable interpretable separately for contents and styles, but we are not sure if we can realize it. We will try our best if there is enough time.

3 Dataset

Right now we are using LSUN dataset to do our experiments, but we won't reject other datasets if we have time to try.

Appendix

The output of the GAN we pre-trained on LSUN church-outdoor class:

Figure 1: Output of our pre-trained GAN on content images.

Continue training the GAN with style loss network by $The\ Scream$ as the target style:
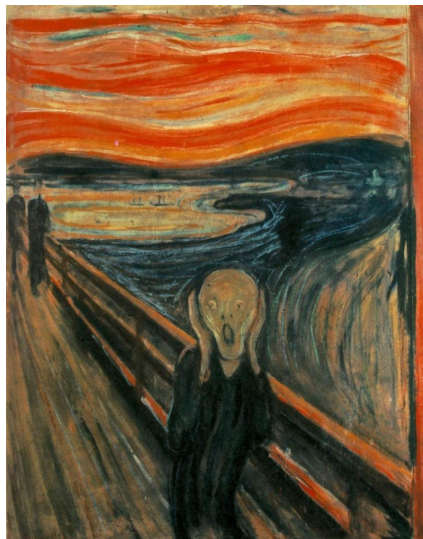

Figure 2: Target style.

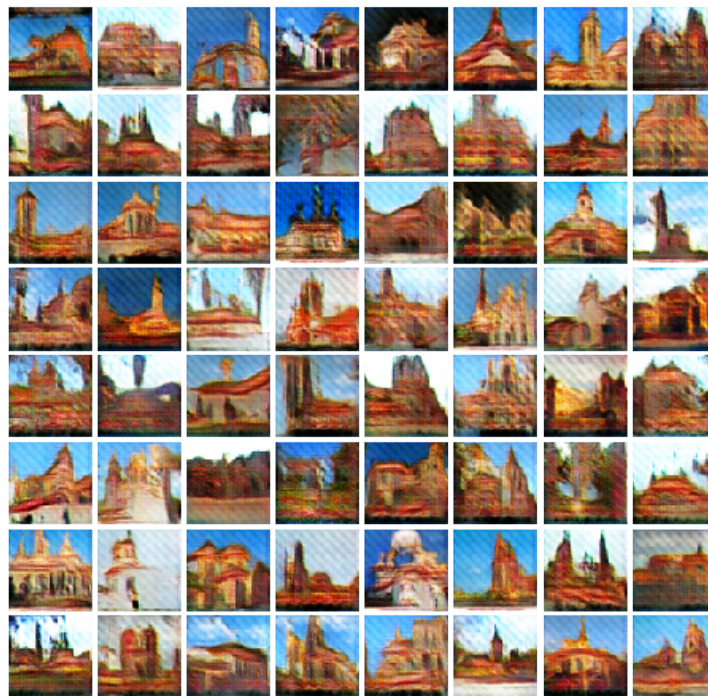Output of our StyleGAN:

Figure 3: Output of StyleGAN (lower learning rate).

Figure 4: Output of StyleGAN (higher learning rate).