Final Project Proposal

Group 3
Nov. 11th

1. Introduction

We are all amazed by the artist's unique thinking and perfect paintings. I often feel that their talent is a gift from god. But modern science and technology allow each of us to become a very powerful artist.
Thanks to Generative Adversarial Networks (GANs), we can use algorithms to imitate the "je ne sais quoi" of artists like Claude Monet -- In the Kaggle competition "I'm Something of a Painter Myself", it requires bringing that style to photos or recreate the style from scratch.

2. Dataset

The data set is from Kaggle, which contains four directories: monet_tfrec, photo_tfrec, monet_jpg, and photo_jpg. The monet_tfrec and monet_jpg directories contain the same painting images, and the photo_tfrec and photo_jpg directories contain the same photos. There're 300 Monet paintings and 7028 other photos, each sized 256*256. The output should be 7028 images with Monet style.

Data link: https://www.kaggle.com/c/gan-getting-started/data

3. Method

3.1 Network

A GAN consists of at least two neural networks: a generator model and a discriminator model. The generator is a neural network that creates the images. This generator is trained using a discriminator.
The two models will work against each other, with the generator trying to trick the discriminator, and the discriminator trying to accurately classify the real vs. generated images.

3.2 Evaluation:

Results are evaluated on **MiFID** (Memorization-informed Fréchet Inception Distance), which is a modification from Fréchet Inception Distance (**FID**). The smaller MiFID is, the better generated images are.
FID is calculated by computing the Fréchet distance between two Gaussians fitted to feature representations of the Inception network. In addition to **FID**, Kaggle takes training sample memorization into account.
The memorization distance is defined as the minimum cosine distance of all training samples in the feature space, averaged across all user generated image samples. This distance is thresholded,

and it's assigned to 1.0 if the distance exceeds a pre-defined epsilon.

4. Related technology

The CycleGAN is a technique that involves the automatic training of image-to-image translation models without paired examples. The models are trained in an unsupervised manner using a collection of images from the source and target domain that do not need to be related in any way. It is an approach for learning to translate an image from a source domain X to a target domain Y in the absence of paired examples.

5. Schedule

Task	Time
Familiar with the data	2days
Data processing	4days
Familiar with GAN	3days
Build a GAN	1-2weeks
Evaluation and improvement	1week

* If there is time left, proceed with project expansion