Datasets

- Input domain: Danbooru Sketch dataset (Figure 1), Quick Draw (Figure 2)
- Output domain: Danbooru dataset 2019 (Figure 3), Cartoon Set (Figure 4)

Motivation & Objectives

Human stick figures are always very intuitive and abstract. We found solving the problem of transforming human face sketches into anime avatars would be an interesting area. One application could be developing a kid friendly app that records and converts a kid’s hand draw sketches into anime or cartoon images. In this project, the input to our algorithm is an image of a human face sketch, and the output would be an anime picture of the face. We use GAN as our algorithms to generate the output. The ultimate goal is that output avatars should look similar with input features and meanwhile look authentic.

Models

we have explored CycleGAN[1], DiscoGAN[2], and Unit. We will just demo CycleGAN here

In the CycleGAN, we have the Adversarial Loss from X to Y domain. We also have the same loss from another direction.

$$L_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{x \sim p_{data}(x)} [\log D_Y(y)] + \mathbb{E}_{y \sim p_{data}(y)} [\log (1 - D_Y(G(x)))]$$

We also use Cycle Consistency Loss to implies that generators should be able to bring x or y back to the original input,

$$L_{cy}(G, F) = \mathbb{E}_{x \sim p_{data}(x)} \| F(G(x)) - x \|_1 + \mathbb{E}_{y \sim p_{data}(y)} \| G(F(y)) - y \|_1$$

Then we can get the full loss function,

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{adv}(F, D_X, Y, X) + \lambda L_{cy}(G, F)$$

The optimization goal is,

$$G^*, F^* = \min_{G, F} \max_{D_X, D_Y} L(G, F, D_X, D_Y)$$

We use the implementation provided by Erik Lindner-Norën. In this implementation, the generators are ResNet. PatchGAN is used in Discriminators.

Experiment 1

- **Cycle GAN** [1] (lr =0.0002 , adam_optimizer, n_residual_blocks=9, batch_size = 64, Figure 5 at 20 epochs about 14700 iteration)

Experiment 2

- **Cycle GAN** (lr =0.0002 , adam_optimizer, n_residual_blocks=9, batch_size = 64, Figure 6 and Figure 7 at around ~40 epochs)

References
