



Everyone could be artist: sketch to style-specified artwork

Fei Jia
Music Science Technology

Qian Dong
Electrical Engineering

Kefan Wang
Electrical Engineering

CS230 Project

Motivation

Sketch-to-image translation is a promising and powerful tool which has the potential to change people's modes of expression. With the power of deep learning network, artwork and creation won't be limited within a small group of artists; on the contrary, everyone has the possibility to fully explore his art potential and be a talented artist. A simple sketch could be magically converted into a masterpiece only with a bit of guidance provided by the deep learning network.

Dataset

Real cat images: cats with different poses and locations, over 9000 images
Van Gogh paintings: 400 images

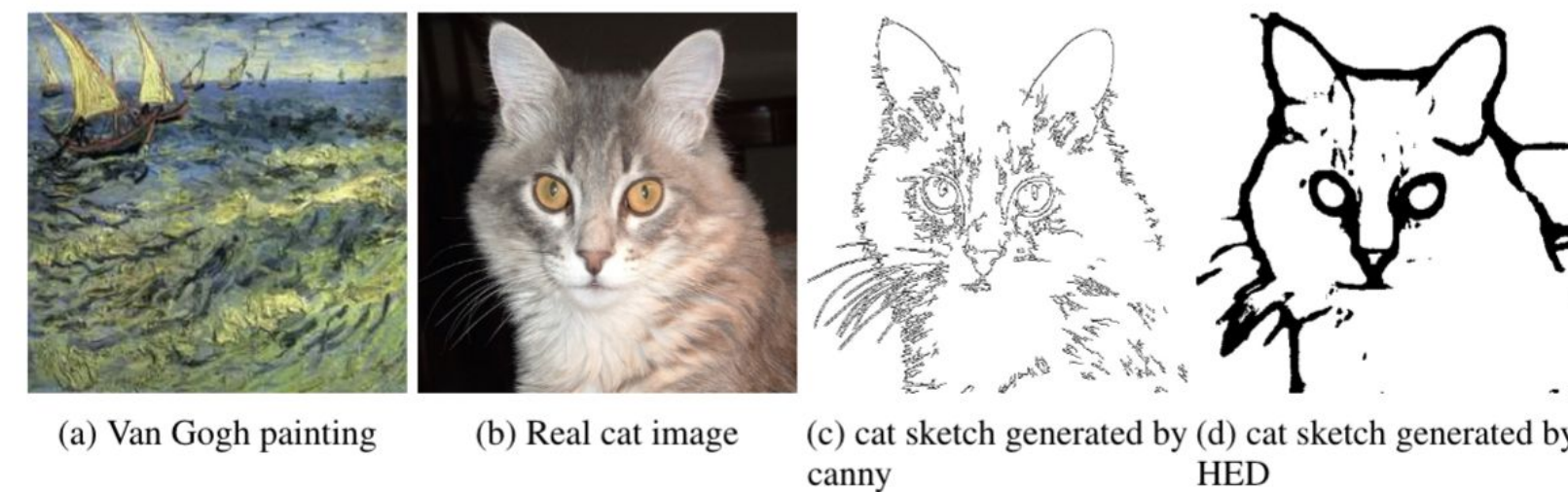


Figure 1: Dataset samples

Data Preprocessing

800 cat images and 400 Van Gogh paintings for training and 50 cat images for testing.

All the images were resized to 256 x 256 and normalized as well.

We used two methods canny and HED to generate cat sketches shown in Fig.1c and Fig. 1d

Methods

Serial cycleGAN structure:

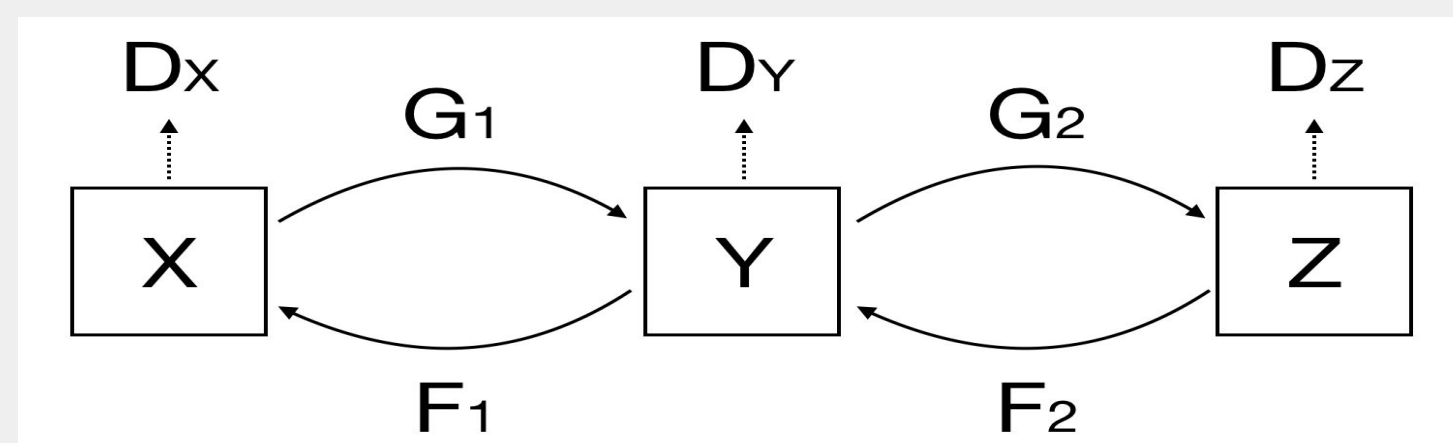


Figure 2: serial cycleGAN structure

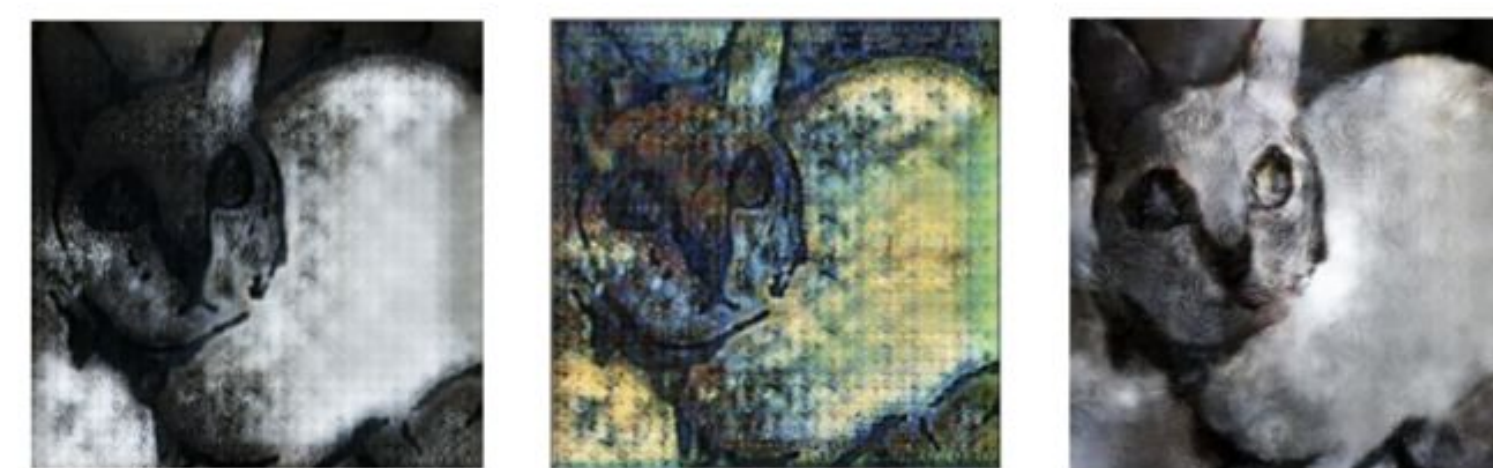
Two cycleGANs are trained together serially with the same number of epochs.

The first cycleGAN generates real cat images from edge.

The second cycleGAN translates the real cat images into images with specified style.

It didn't work well:

1. It is harder to learn to generate real cat images than a image with Van Gogh style.
2. Generating styled cat image is based on the intermediate generated cat image which might have repeated pattern that will be preserved



(a) Serial mid output (b) Serial output (c) Parallel output

Figure 3: Output comparison between serial structure and parallel structure

Parallel cycleGAN structure:

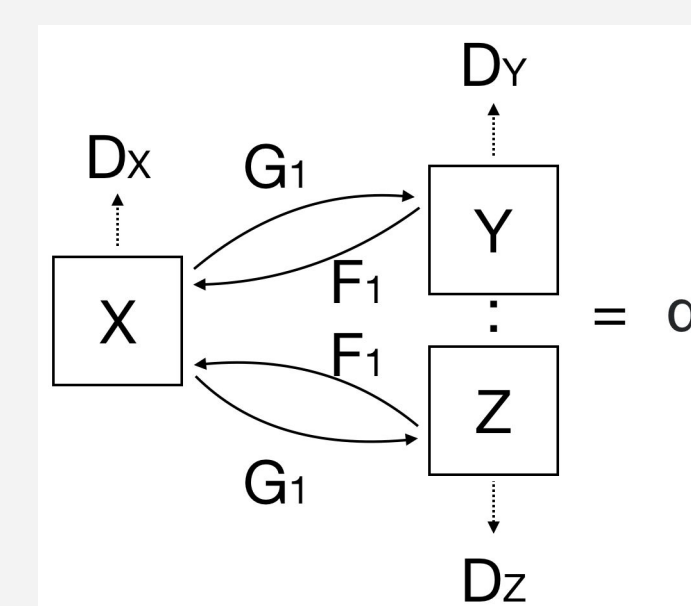


Figure 4: parallel cycleGAN structure

In parallel cycleGAN, we used a single cycleGAN by sharing the same generators and discriminators in both directions.

Considering the different levels of difficulty of two sub-tasks, we proposed loss function to be the adjustable weighted sum of the loss of two sub-tasks.

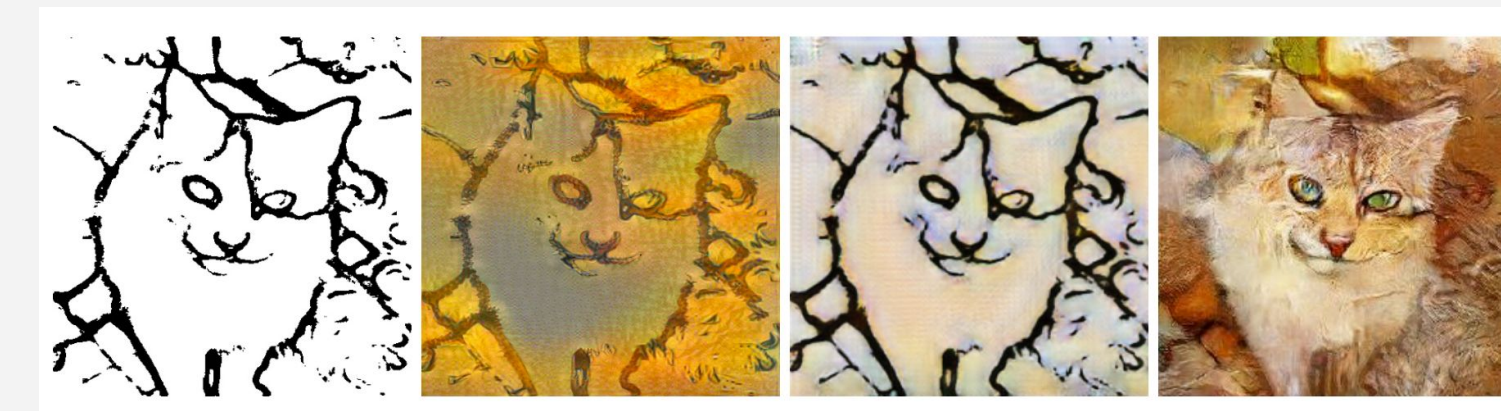
For a harder task (generating real cat images), higher weight will push the model to learn faster. Besides, the result will not be affected by any intermediate generated result.

Baseline



(a) First (b) Second (c) Third (d) Fourth (e) Fifth

Two baseline models: neural style transfer and pretrained cycleGAN. Baseline models simply fills the edge with color from Van Gogh painting, while our parallel cycleGAN outputs detailed and artistic image.



(a) Input edge (b) Neural style output (c) Pretrained cycleGAN output (d) Parallel cycleGAN output

Figure 5: Output comparison between baseline models and parallel cycleGAN structure

Results & Analysis

We collected 2791 responses from 263 participant for 50 generated images and excluded abnormal responses. Each image has been assigned to 48 to 60 different raters. Considering the entire image set,

- Over 70% people believed the images were closer to real cat (cat score ≥ 3)
- 68.3% people considered the images are Van Gogh's work by comparing to several piece of artworks of him (Van Gogh score ≥ 3).
- Surprisingly, 72.6% people indicated that they like our generated artworks (like score ≥ 3)

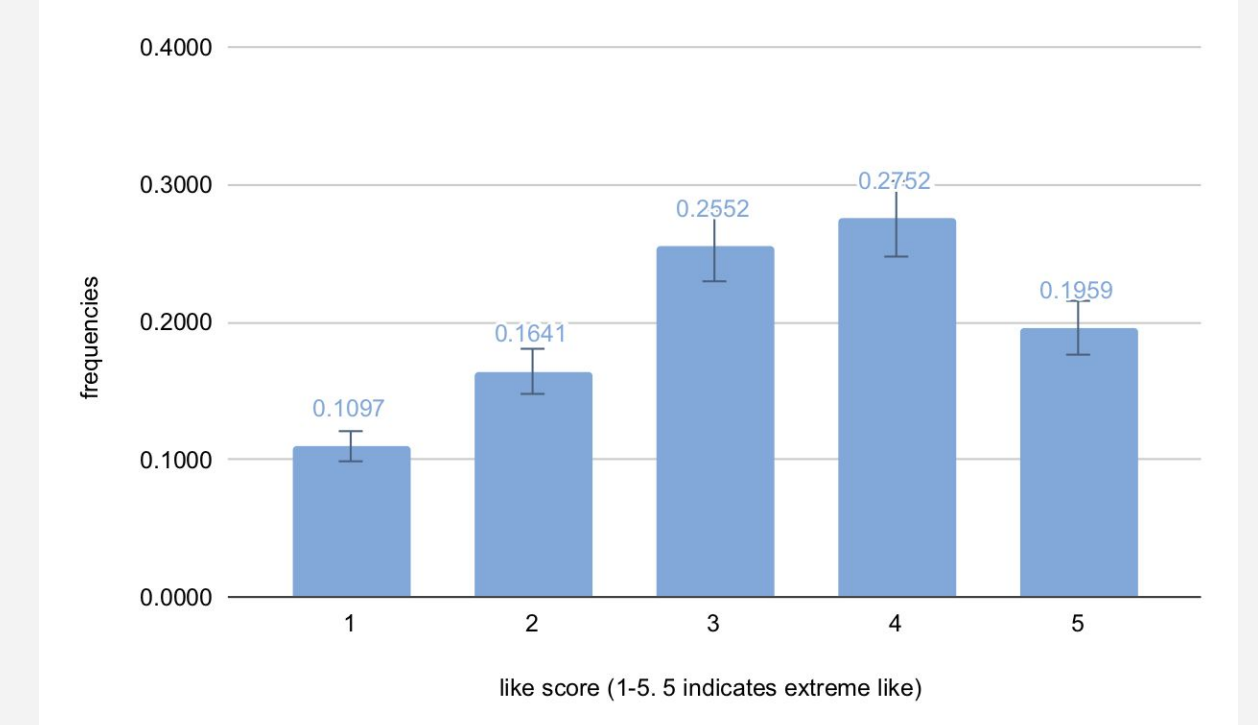
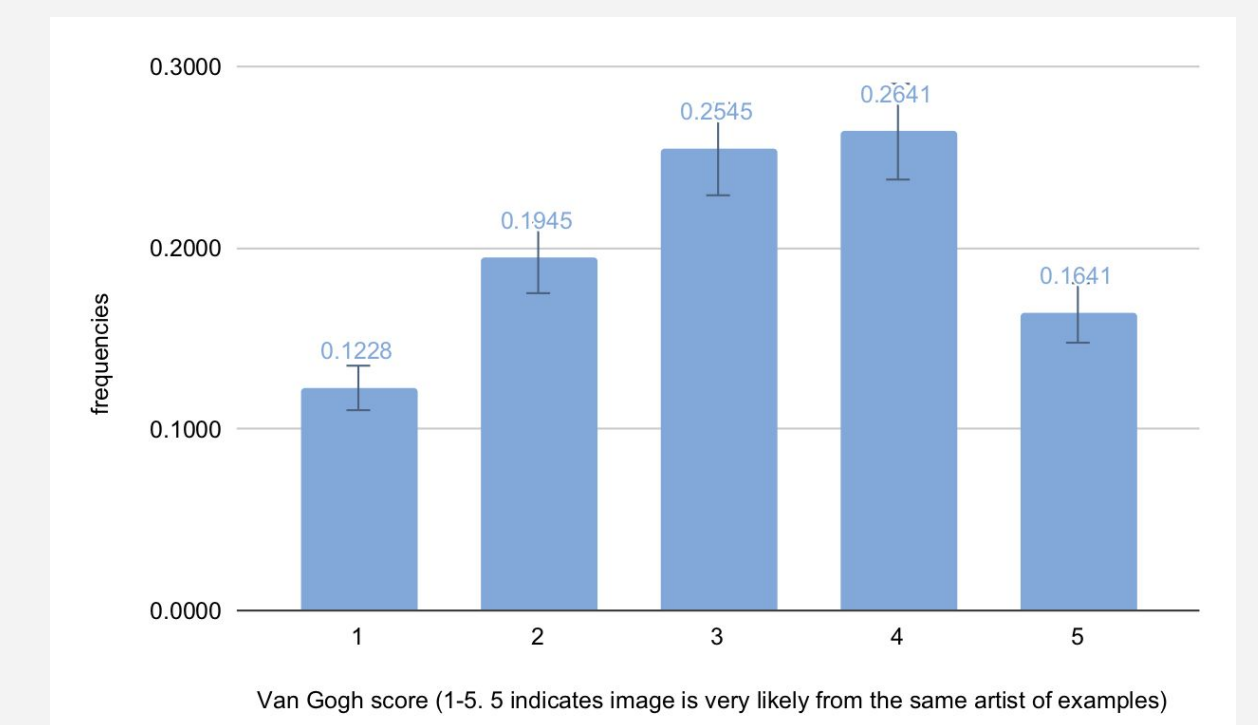
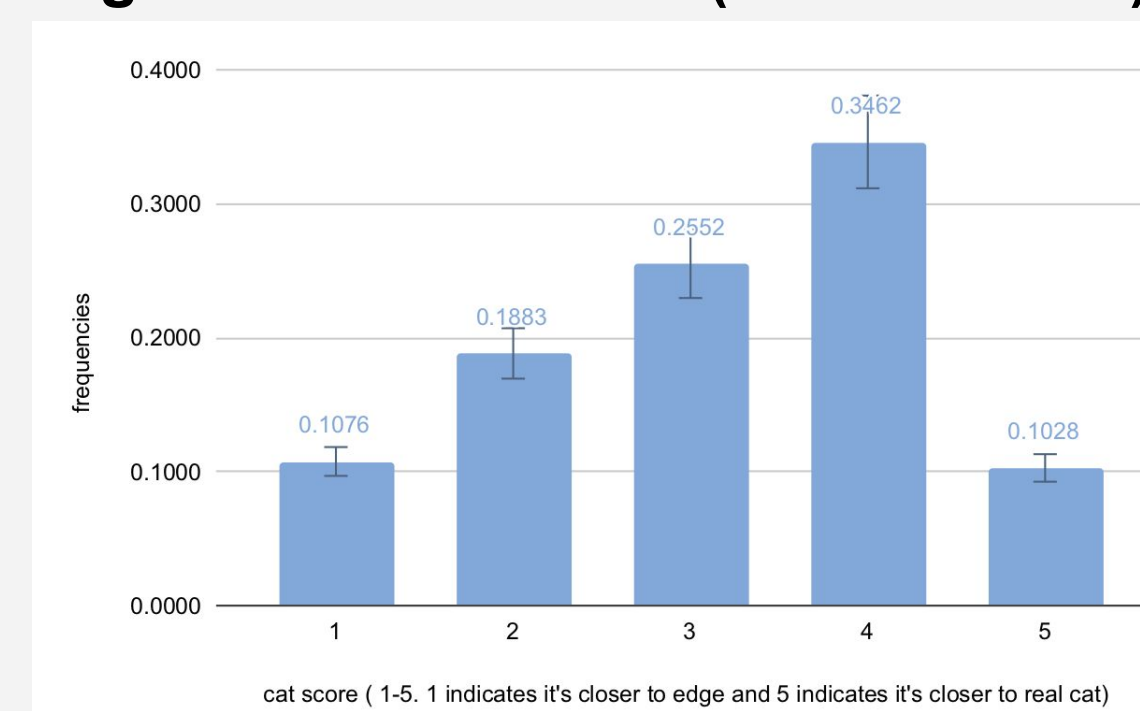


Figure 6: Histogram of responses

Conclusion and Future Work

Two creative models: serial cycleGAN structure and parallel cycleGAN structure.

Parallel cycleGAN structure performs the best to generate images with both real cat details and Van Gogh artistic style.

Based on human evaluation results, the images generated by our model are greatly acknowledged and liked by the majority.

In our future work, we would like to try

- pre-trained VGG16 network
- transfer learning.
- to improve our evaluation strategies
 - quantitative evaluations
 - more participant in AMT perceptual studies.