Line Drawing Auto-Colorization

Tianchi Liu
Stanford University
kitliu5@stanford.edu

Yuyan Wang
Stanford University
thea0718@stanford.edu

1. Introduction

Colorization of line drawing is often skipped by artists due to its time-consuming nature. Automating this process could help artists to save time and effort while enriching content of art pieces. In this work, we develop a system that automatically colorizes line drawings using deep learning methodologies.

The goal of this project is to build a model that converts sketches to colored drawings. A number of deep neural networks, such as deep convolutional neural network [4] and U-Net [2], have proven to be effective in colorizing grayscale real-life images. Inspired by these works, we hypothesize that sketch colorization can also be accomplished by CNN and GAN models. In this project, we will implement CNN and GAN architectures to automate sketch colorization.

2. Problem Statement

2.1. Dataset Aggregation

The main focus of this project is to investigate the implementation of CNN and GAN models to automate sketch colorization. Unfortunately, as most of computer vision research today focuses on real-life images instead of drawn illustrations, there are no publicly available datasets that only contain illustrations. Therefore, we created a dataset by retrieving hand-drawn images from *Pixiv.net*, a Japanese online platform for sharing artwork.

So far we have crawled 1,500 images from *Pixiv.net* based on user ID. The monochrome images are removed from the dataset. Then a series of computer vision techniques(OpenCV) are applied to convert colored illustrations into sketch. During data preprocessing, the images are also reformatted to 256px by 256px, the input size of the network. The examples of the resulting dataset images can be seen in Figure 1. The dataset was then split into 80/10/10 into training set, test set, and the validation set.

2.2. Evaluation

We expect the results from our model to be similar to the original colored image in color, shading and texture.

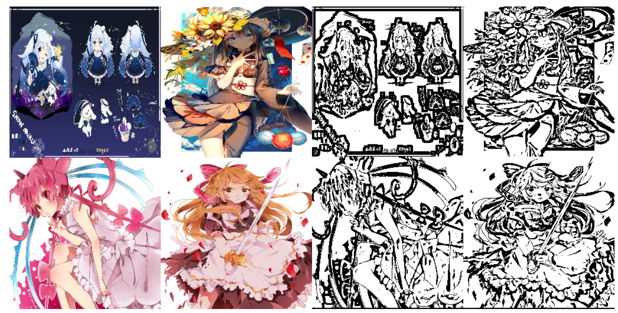

Figure 1. Example Dataset images

We propose two types of metrics to evaluate our model. The first is a comparison quantified by MSE based on CIE-LCH color space. That is the lightness, chroma(saturation level) and hue of an output image will all be compared to the corresponding original colored image. The second is a user score defined as a measurement of aesthetics regarding human perception. Since the aesthetic standard of artistic stylization is highly subjective, we hope that a user study among students can help to access the performance of our models more precisely from an aesthetic and intuitive perspective.

3. Technical Approach

3.1. U-Net

We first perform adversarial training with an U-Net style generator, as shown in Figure 2, as the baseline model. We adapted the GAN architecture from those in Frans [1] and both generator and discriminator have Convolution-BatchNorm-ReLu modules. We expect this encoder-decoder architecture with skip connections to be powerful enough to learn a mapping from gracyscale image to colored image.

3.2. CycleGAN

CycleGAN by *Zhu et al* [5] performs well at various image-to-image translation tasks, such as style transfer and photo enhancement. It contains two generator-discriminator pairs that performs forward and reverse mapping, respec-

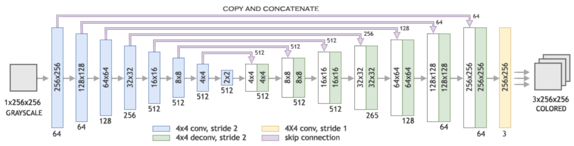

Figure 2. U-Net as generator [3]

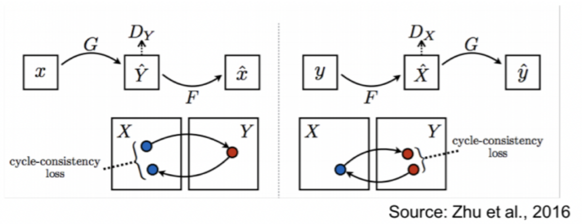

Figure 3. Forward and Reverse Mapping of CycleGAN

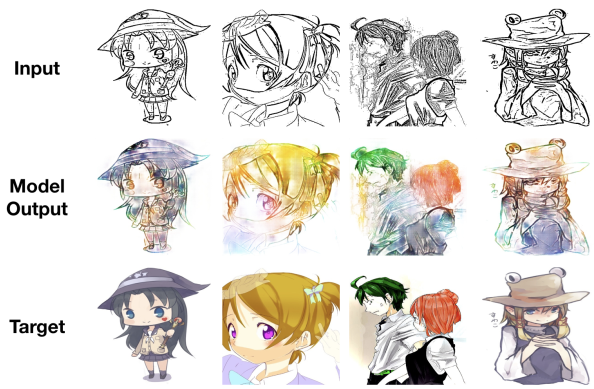

Figure 4. U-Net model results

tively, as shown in Figure 3. The training loss of this model also has an extra cycle-consistency term to promote learning of reversible mapping. We utilize the pre-trained weights provided by the authors to perform transfer learning to see if the model performs well in colorizing line drawing. As a sanity check, output of the generator that performs reverse mapping from originally colored images to sketches, will also be visualized to see if the model performs well at generating sketches, which is a completely different mechanism compared to the preprocessing described in Section 2.1.

4. Preliminary Results

4.1. U-Net

The sample output of U-Net model is shown in Figure 4. The color of the output images is analogous to the color of original images because of the implementation of color hints. As shown in Figure 4, the results from U-Net have some undesired blurry pattern in the output. This may be caused by insufficient training data.

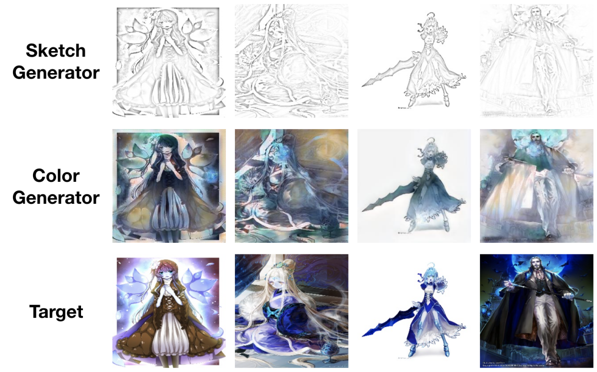

Figure 5. CycleGAN model results

4.2. CycleGAN

The sample output from the sketch generator and color generator in CycleGAN model is shown in Figure 5. Output from color generator is bluish because most of the training images have a blue tone. Compared with the output from U-Net, CycleGAN considers a more fine grained color application to images from sketch generator.

5. Future Work

So far we have established a data processing pipeline and two running models. For the rest of the project, we plan to improve our model based on the following approaches:

1. Build a larger illustration dataset. Currently we are only training model based on pretrained weights and 1,500 line drawing images, and it is obvious that the size of the data is too small. We will try to experimenting different models with a dataset containing more than 50,000 images. Also we will crawl images from various painters in order to avoid the similarities in tone.

2. Apply the idea of color hints into our model. Contrary to real objects, the color of illustrations vary among different painters and styles. Inspired by Frans [1], We can give the network another input image containing the colors of the original image achieved by applying a large 100px radius blur and passing that as our color hint.

3. Perform quantitative evaluation on further improved results using metrics described in Section 2.2.

References

[1] K. Frans. Deepcolor: automatic coloring and shading of manga-style lineart, 2017.

[2] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[3] K. Nazeri, E. Ng, and M. Ebrahimi. Image colorization using generative adversarial networks. In *International conference on articulated motion and deformable objects*, pages 85–94. Springer, 2018.

[4] P. I. Zhang Richard and A. A. Efros. Colorful image colorization. *European coference on computer vision*, 2016.

[5] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.