Full-body Anime Character Generation using Generative Adversarial Networks

Qiucan Huang
HKUST
Clear Water Bay, Hong Kong
qhuangag@connect.ust.hk

Sida Chen
HKUST
Clear Water Bay, Hong Kong
schenci@connect.ust.hk

Yifei Lin
HKUST
Clear Water Bay, Hong Kong
ylinbz@connect.ust.hk

Introduction

Nowadays, Generative Adversarial Networks (GAN) [1] by Goodfellow et al. is more and more popular and widely used in many areas. People from NVIDIA also made great improvements in GAN by introducing better architectures like Progressive GAN [3] and style-based generator [5], by improving the quality of the generated image [2]. Besides these real faces generation, a lot of people are also using it for entertainment like anime character generation [6, 7]. However, all of them want to be as general as possible, and the generated images have big randomness, and they also only generate avatar. What we want is a generator that could generate full-body anime characters based on everyone's special favors. We plan to utilize the method as mentioned by Y. Jin et.al [6] that requires user inputs to define certain features. We also plan to use the idea of progressive GAN [3] but generating in an order of line sketch, colorization, and ornaments.

1.   Problem Statement

Our goal is to use Generative Adversarial Nets [1] to generate full-body images of anime-like characters based on the text description. We have found some public datasets [8, 9, 10] and prepared some full-body figures collected from the internet and mobile games (Arknights, Azur Lane, Fate Grand Order, Horcrux college, Sdorica -sunset) as a supplement.

The expected result is to generate figures that are similar to real anime characters and consistent with the description. It will be evaluated mostly by similarity, rationality, and consistency.

2.   Technical Approach

As mentioned above, we want to address the problem of generating full-body anime characters with specific feature input. First of all, since this is a problem of generation, we need to build a generative model. Based on our research, GAN [1] is a very popular model that is being used in generation problems in the past few years. Besides GAN, there are other variations of GAN that accomplish similar problems in our problem statement, like StyleGAN [5] and progressive GAN [3]. We will use a mixture of these two GAN networks to make our own network.

Our approach is first to learn the features in the datasets, like StyleGAN. Then, we will have three generative networks to generate line sketch, colorization of line sketch, and ornaments separately. We will also have three discriminative networks correspondingly, just like progressive GAN. The datasets of the discriminative networks at intermediate states will come from original images with ornaments removed and color removed by openCV. However, a problem at this stage is that we might not be able to remove ornaments properly. For example, the hair may be removed when we are removing hair accessories. Depending on the results, we might remove the stage of ornament.

Learning features and generating features based on input could be straight forward, as shown by Y. Jin et al. [6]. However, different from the progressive GAN that generates images from 8x8, to 16x16, to 32x32, and finally reaching 1024x1024, we would like to generate line sketches at 256x256 directly. This could reduce the complexity and dimensionality of the network hugely, making the result more tangible.

3.   Intermediate/Preliminary Results

To train a good generator, we need to make a dataset first. According to [6], the style of images is changing with the publication time changing, the image before 2005 may not fit people's aesthetics today. We find some public datasets [8, 9, 10] and check some images in datasets randomly to evaluate the quality of those datasets. The Manga109 [8] contains tons of images of different ages. But most of them are collected from cartoons that were created before 2005 and lots of images are drawing only. Danbooru 2019[10] provides anime images too, but the styles of images are not consistent. To use Danbooru 2019 as our dataset, we need to check over 10G data one by one by ourselves.
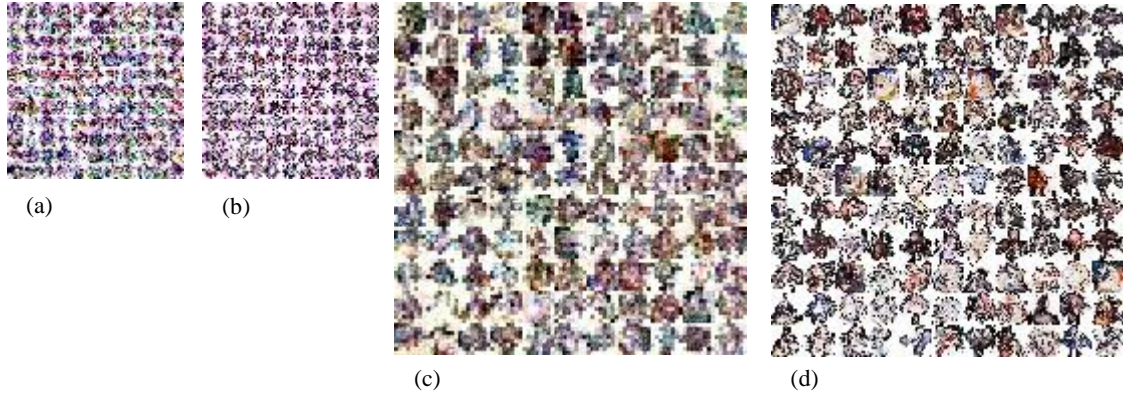
Fig 1: From Figure 1a to Figure 1b. We can find there are more details in b, even though we cannot identify an anime character. As the training goes, we can get more details in Figure 1c and Figure 1d. Until Figure 1d, we can roughly identify the contour of an anime character in some images which are generated by the generator.

To solve dataset problems, we filter the image in the dataset by some properties, size etc. and check the rest of the images one by one. Then we collect images from some games and as supplements. Up to now, we had collected 1000+ images from two different games and around 400 images from different datasets.

Nvidia publishes StyleGAN++ [4] recently. StyleGAN [5] proposes a non-linear mapping network to map the input latent into different layers and using noise to control the detail. StyleGAN++ [4] analyses the SytelGAN and tries to reconstruct the details of StyleGAN. We have tried StyleGAN architecture on a dataset with 919 images which was collected from a game. The network didn't finish training due to some technical issues that we are still trying to deal with. But we got some images that presented the progress of training. See Figure 1.

References

[1] I. J. Goodfellow et al., "Generative Adversarial Networks," arXiv [stat.ML], 2014.

[2] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," arXiv [cs.CV], 2019.

[3] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," arXiv [cs.NE], 2017.

[4] H. Zhang et al., "StackGAN++: Realistic image synthesis with Stacked Generative Adversarial Networks," arXiv [cs.CV], 2017.

[5] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," arXiv [cs.NE], 2018.

[6] Y. Jin, J. Zhang, M. Li, Y. Tian, H. Zhu, and Z. Fang, "Towards the automatic Anime characters creation with generative adversarial networks," arXiv [cs.CV], 2017.

[7] K. Hamada, K. Tachibana, T. Li, H. Honda, and Y. Uchida, "Full-body high-resolution Anime generation with Progressive Structure-conditional Generative Adversarial Networks," arXiv [cs.CV], 2018.

[8] "Manga109," Manga109.org. [Online]. Available: http://www.manga109.org/en/. [Accessed: 08-Nov-2020].

[9] "- nico-opendata," Nico-opendata.jp. [Online]. Available: https://nico-opendata.jp/en/seigadata/index.html. [Accessed: 08-Nov-2020].

[10] Anonymous, The Danbooru Community, and &. G. Branwen, "Danbooru2019." 13-Jan-2020.