

Literature Review

Guorui Lu

2018.8.15

1 History and Significance

Synthesizing realistic photo from sketches drawn by human has been a challenging and hot problem for a long time. The research of this technique can be traced back to 2009 [3]. However, the technique of image generation and image-to-image transfer did not rise up until 2014, when the Generative Adversarial Nets(GAN) [7] was proposed. Since then, variety of methods have been proposed to generate a more realistic photo for its broad application prospect. For example, police can use it to catch suspects through synthesised photos based on the drawn sketches, which requires that our photo should be as accurate as possible. As for the entertainment industry, we may need the model is able to generate images of multiple style, which is also a popular but difficult task.

2 Mainstream Method

I will introduce the main method of this project from 2015 to the present because the actual reasearch on sketch-photo synthesis didn't appear until 2014, which I think is largely due to the fact that GAN [7] was present in 2014. Before it, sketches were mainly used to do retrieval work [1, 2, 6, 9–11, 13, 17].

2.1 Generative Adversarial Nets(*GAN*)

2.1.1 Generative Adversarial Nets

The original GAN brought us a completely new idea. It is corresponding a minimax two-player game, which is simple and efficient. But it also has an obvious shortcoming that the constraint is too weak so that we can't control what it will generate.

2.1.2 cGAN and InfoGAN

Soon after the GAN's appearance, conditional GAN [14] was proposed, and then the InfoGAN [5]. These two models are able to partly control the outputs by adding extra codes. However, the pictures they synthesize are still blurry and low-resolution.

2.1.3 Pix2pix and Pix2pixHD

Pix2pix [12] and pix2pixHD [16] are two of the improved versions of the GANs. They took a big step forward to addressing the resolution issues. What they didn't resolve is that only when we have lots of paired images can we train the models.

2.1.4 CycleGAN

Inspired by NLP, Jun-Yan Zhu and Taesung Park et al. came up with a new model named CycleGan [19]. Although it has some drawbacks such as the high computing cost, CycleGAN still works very well for many problems. So I think I can apply it to my project.

2.1.5 *Rescent Work*

Here I would show you some rescent work of sketch-photo in 2017 and 2018.

- **SketchyGAN: Towards Diverse and Realistic Sketch to Image Synthesis [4]:**

A new network structure proposed for generative task. It works better than retrieval model but the result is still blurry and low-resolution. Moreover, it is too faithful to the badly drawn pictures to keep the realism.

- **Facial Attributes Guided Deep Sketch-to-Photo Synthesis [8]:**

This paper puts forward a method that add an auxiliary attribute discriminator to find the false attributes in the output of the generator. However, editing an attribute can cause unwanted structural edition of the image in some area, which is also the weakness of the previous models.

- **TextureGAN: Controlling Deep Image Synthesis with Texture Patches [18]:**

Just as its name implies, TextureGAN adds a texture patch on a sketch at arbitrary locations and scales, which allows it to synthesize high-quality photos. And its novel losses for training deep image synthesis can encourage the generative network to handle new textures never seen on existing objects.

- **Scribbler: Controlling Deep Image Synthesis with Sketch and Color [15]:**

In this research, Pastorn Sangkloy et al. use sketched with sparse color "scribbles" decided by human to generate a higher resolution and more diverse images. However, if users specifies uncommon or even wrong colors and shapes, this model would deem them correct. In addition, sometimes the boundaries between object parts or regions of different colors are likely to become vague.

2.2 Deep Convolutional Neural Networks(DNNs)

Although GANs has been a remarkable success in image generation field, deep convolutional neural network is still also an effective tool.

3 Focus Questions

1. Lack of sketches, especially the paired images and sketches;
2. Uncontrollable results
3. Low-resolution
4. Badly drawn pictures
5. Models can't fit the objects that don't appear during the training procedure

References

- [1] Yang Cao, Changhu Wang, Liqing Zhang, and Lei Zhang. Edgel index for large-scale sketch-based image search. In *Computer Vision and Pattern Recognition*, pages 761–768, 2011. 1
- [2] Yang Cao, Hai Wang, Changhu Wang, Zhiwei Li, Liqing Zhang, and Lei Zhang. Mindfinder:interactive sketch-based image search on millions of images. In *International Conference on Multimedia 2010, Firenze, Italy, October*, pages 1605–1608, 2010. 1
- [3] Tao Chen, Ming Ming Cheng, Ping Tan, Ariel Shamir, and Shi Min Hu. Sketch2photo. *Acm Transactions on Graphics*, 28(5), 2009. 1
- [4] Wengling Chen and James Hays. Sketchygan: Towards diverse and realistic sketch to image synthesis. 2018. 1
- [5] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. 2016. 1
- [6] Mathias Eitz, Kristian Hildebrand, Tamy Boubekeur, and Marc Alexa. An evaluation of descriptors for large-scale image retrieval from sketched feature lines. *Computers Graphics*, 34(5):482–498, 2010. 1
- [7] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *International Conference on Neural Information Processing Systems*, pages 2672–2680, 2014. 1
- [8] A Dabouei S Soleymani NM Nasrabadi H Kazemi, M Iranmanesh. Facial attributes guided deep sketch-to-photo synthesis. *IEEE Winter Applications of Computer Vision Workshops (WACVW)*. 1
- [9] Rui Hu, Mark Barnard, and John Collomosse. Gradient field descriptor for sketch based retrieval and localization. In *IEEE International Conference on Image Processing*, pages 1025–1028, 2010. 1
- [10] Rui Hu and John Collomosse. A performance evaluation of gradient field hog descriptor for sketch based image retrieval . *Computer Vision Image Understanding Cviu*, 117(7):790–806, 2013. 1
- [11] Rui Hu, Tinghuai Wang, and John Collomosse. A bag-of-regions approach to sketch-based image retrieval. In *IEEE International Conference on Image Processing*, pages 3661–3664, 2011. 1
- [12] Phillip Isola, Jun Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5967–5976, 2017. 1

- [13] Yen Liang Lin, Cheng Yu Huang, Hao Jeng Wang, and Wei Chou Hsu. 3d sub-query expansion for improving sketch-based multi-view image retrieval. pages 3495–3502, 2014. [1](#)
- [14] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *Computer Science*, pages 2672–2680, 2014. [1](#)
- [15] Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. Scribbler: Controlling deep image synthesis with sketch and color. pages 6836–6845, 2016. [2](#)
- [16] Jun-Yan Zhu Andrew Tao Jan Kautz Bryan Catanzaro NVIDIA Corporation UC Berkeley Ting-Chun Wang, Ming-Yu Liu. High-resolution image synthesis and semantic manipulation with conditional gans. In *Computer Vision and Pattern Recognition*, 2017. [1](#)
- [17] Changhu Wang, Zhiwei Li, and Lei Zhang. Mindfinder:image search by interactive sketching and tagging. In *International Conference on World Wide Web*, pages 1309–1312, 2010. [1](#)
- [18] Wenqi Xian, Patsorn Sangkloy, Varun Agrawal, Amit Raj, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. Texturegan: Controlling deep image synthesis with texture patches. 2017. [2](#)
- [19] Jun Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. pages 2242–2251, 2017. [1](#)