Cartoon Auto-painter

CS 698O: Visual Recognition

Mid-term Project Status Report

Anil Kumar Gupta- 14097   Kritagya Dabi - 14326

Group No. - 32
October 2017

1 Brief Description of Project

Realistic image generation from black and white images, sketches or outlines using deep neural networks has become a hot topic recently in visual recognition. Colorful realistic images can be generated at pixel or patch level by training from a large data-set.

The aim of this project is to build a conditional GAN [1][2] based model for translating sketches of cartoons into colorful and appeasing images. The architecture of the model is based the pix2pix model [3]. The model is to be trained by considering different combinations of losses like GAN loss, pixel loss, feature loss and total variance loss. We also aim to compare test results for these models with different losses [4]. For training and testing purpose we have created the data-set of images of minions.

In the next part of the project, we aim at doing the analysis of different conditions on input images, like training on particular cases and observe the change in outputs when trained with particular cases as compared to without them. An example of this would be analysis of the test results on images with missing features in sketches. It includes comparison of results when the model is trained with and without examples with missing features.

2 Work Done

- **Literature Survey :** We went through the following papers related to the project. Image-to-image translation with conditional adversarial networks [3] has the Pix2Pix model which uses conditional GAN along with the U-net architecture and L1 loss function for determining whether an image is real or fake . Auto-painter [4] improves the pix2pix model by taking more losses i.e. feature loss and total variance loss into the picture.

- **Pix2Pix Model Implementation :** We have implementing the code for Pix2Pix from scratch using the code of Pix2Pix implemented in pyTorch by Jun-Yan Zhu as reference. We are training Pix2Pix on this data-set and try to tune it to improve the result obtained. Also we are working on implementing the code of Auto-painter by adding loss functions step by step and observing the variation of the results.

- **Datasets :** We have also manually constructed a small data-set of images of Minions by downloading these images from the internet. For training we need images in pairs of sketch and corresponding cartoon, for this we have written the code for edge detection. Using this edge detection code we generate sketch from the colorful cartoon images.

3 Challenges

- The data-set required is a collection solo images of minions. There is no standard data-set available and we need to manually construct the data-set by scraping images from the internet or by capturing screen-shots from the movies. Also there is no measure of accuracy for comparing results obtained. There is just an aesthetic sense while comparing two images which of the two looks more minion like.

- For the training purposes we were needed pair of images black and white image or sketch and it's corresponding colorful image. But it's hard to find sketch and colorful images in pairs. Also the black and white images obtained through edge detection have either too much of noise or missing features depending on how we set the parameters

4 Future Work

After completing the code for Pix2Pix and the Autopainter, we will fine-tune the results to reach their best possible accuracy. We will then compare the results obtained by the two models. Also we aim at analyzing the effect of varying the conditions on input images on results obtained and how well does it perform.

After completion of training, testing and fine-tuning of the model on cartoon character (minions in our case), we will test the model with specific conditions on training images of minions and compare the results in cases when the model was trained with these particular cases as compared to without them. For eg, we could test the model with images having some missing features and compare results when model was trained with these features vs when model was trained without them.

References

[1] Osindero S. Mirza, M. Conditional generative adversarial nets. *CoRR abs/1411.1784*, 2014.

[2] Akata Z. Yan X. Logeswaran L. Schiele B. Lee H Reed, S. Generative adversarial text to image synthesis. *International Conference on Machine Learning*, 2016.

[3] Tinghui Zhou Alexei A. Efros Phillip Isola, Jun-Yan Zhu. Image-to-image translation with conditional adversarial networks. *Berkeley AI Research (BAIR) Laboratory*, 21 Nov 2016.

[4] Zhenbo Luo Yifan Liu, Zengchang Qin and Hua Wang. Auto-painter: Cartoon image generation from sketch by using conditional generative adversarial networks. *Intelligent Computing and Machine Learning Lab, School of ASEE Beihang University, Beijing 100191, China*, 7 May 2017.