

DELFT UNIVERSITY OF TECHNOLOGY

DEEP LEARNING

CS4180

cGANs for cartoon to real-life images

Authors (Group 25):

Kanya Satis, 4777336

Pranjal Singh Rajput, 4788087

Sonnya Dellarosa, 4783344

Wenxuan Huang, 4925777

Obinna Agba, 4793765

April 28, 2019



1 Problem description

We came across paper [1] that detailed Pix2Pix and decided that it was interesting to work with the proposed framework. It uses a cGAN (Conditional Generative Adversarial Network) to convert drawings or black/white (BW) images to real-life images. We believe it will be intriguing to convert cartoons to real-life facial images. How would your cartoon super hero look like in real life? We try to answer this question through this project.

At this moment, Pix2Pix has already been used by various graphics practitioners and visual artists on a variety of novel image-to-image translation tasks such as sketch-to-portrait and sketch-to-pokemon, and have shown that the network is a generic commodity tool for image-to-image translation problems. Thus, hypothetically, it should also work for *cartoonized* human faces as well.

2 Data collection

After a rigorous search for existing cartoon or sketched images of human faces, we found that most of the existing datasets found either have too few images, which would necessitate augmentation, or are behind some paywall. We also considered the possibility of mixing and matching images from different datasets, but this would most likely produce a non i.i.d dataset which might, in turn, affect results. To this end, we plan to start with a dataset of human faces. We will be using an existing code to generate a cartoon version of the images and use them as our base dataset for training.

We have contacted several agencies (FERET, SCFace and more) that offer facial image data-sets. The response and the availability of these database is pending. These database include tagged human facial image from different racial backgrounds, camera angles and facial expressions. The diversity of the database and samples would enable us to examine and evaluate the robustness of the Pix2Pix [1] algorithm.

3 Method and implementation

The code is based on a Pix2Pix [1] implementation. This will be used for our application instance. There were no prior similar application instances that we came across using a similar network. This project would use the same model architecture as the original Pix2Pix implementation.

4 Preliminary material

The project is based on [1]. We will be going through [2] and [3] for a more in-depth understanding.

5 Hypothesis

The research questions that we aim to answer with this experiment are listed below:

1. Can the same set of hyper-parameters (used in the original implementation) be used on a different type of image transformation chosen for this experiment?
2. Does it properly recreate facial expressions and different postures?

6 Evaluation

For qualitative evaluation, we will compare the generated image with the actual image that it should have generated. For qualitative analysis, we are planning to use the FCN-score as discussed in [1].

7 Timeline management

Week	Task
1	Collect image data-set
1	Generate a cartoon version of the data-set
2-4	Train Pix2Pix mode using default parameters on our data-set
5-7	Evaluate FCN-score and perform qualitative analysis
5-7	Refine parameters and model
8	Write report

8 Feasibility

The code for the architecture already exists and we are adapting it for different application instance. Hardware limitation might be an issue due to limited computation resources (low-performance CPU and GPU) that may delay the training period. So it should be feasible within the timeline.

References

- [1] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [2] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *CoRR*, vol. abs/1411.1784, 2014.

- [3] J. Gauthier, “Conditional generative adversarial nets for convolutional face generation,” 2015.