



CartoonGAN: Generative Adversarial Networks for transfer human face

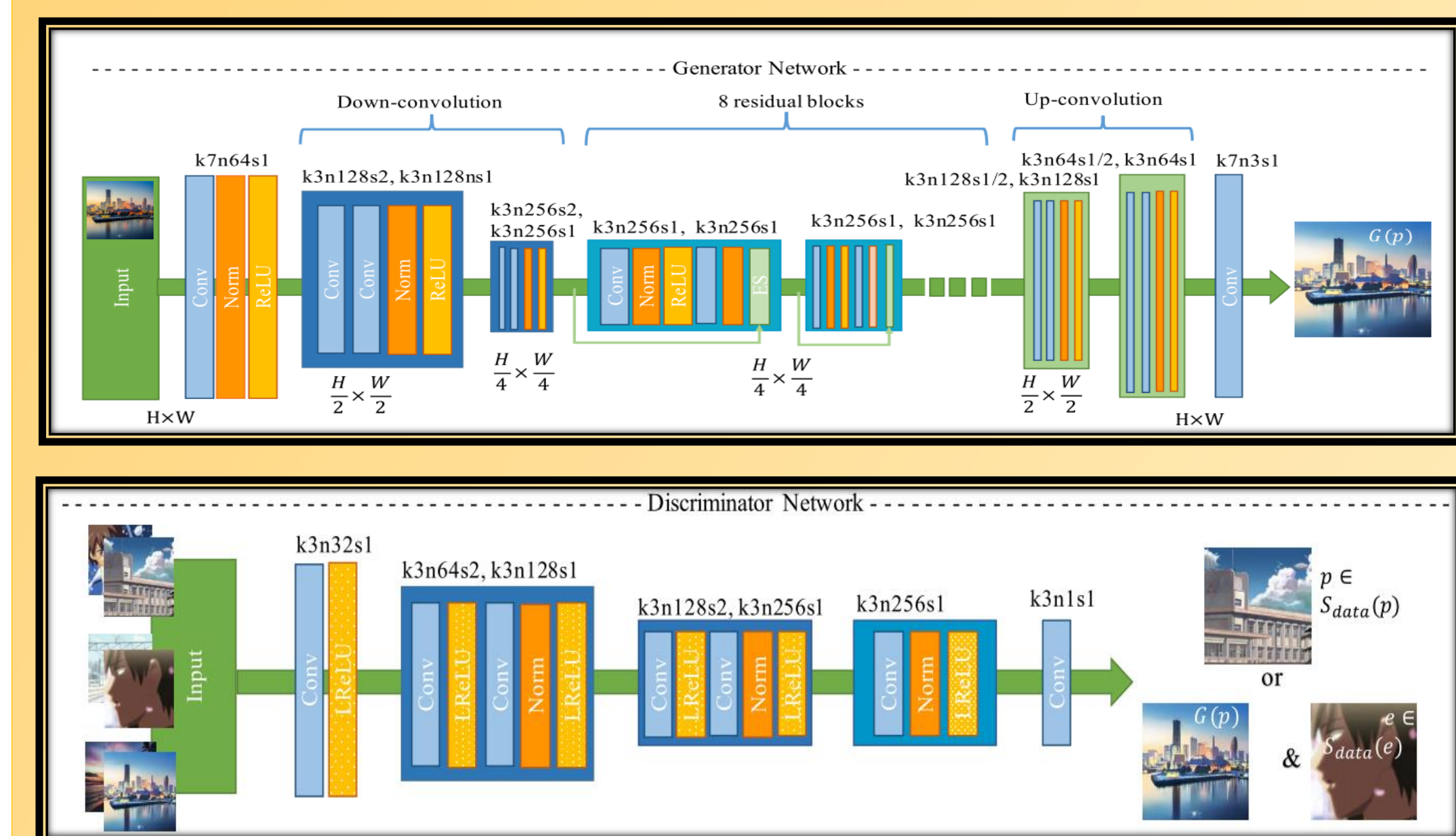
Po-Yuan Lin
National Tsinghua University
Taiwan

Chien-Li Lu
National Chiao Tung University
Taiwan

Introduction

CartoonGAN propose a solution to transforming photos of real-world scenes into cartoon style images, which is valuable and challenging in computer vision and computer graphics. Cartoons are an artistic form widely used in our daily life. In addition to artistic interests, their applications range from publication in printed media to storytelling for children's education. Like other forms of artworks, many famous cartoon images were created based on real-world scenes.

Network architecture



Loss function

$$\mathcal{L}(G, D) = \mathcal{L}_{adv}(G, D) + \omega \mathcal{L}_{con}(G, D)$$

Adversarial loss:

To drives the generator network to achieve the desired manifold transformation.

$$\mathcal{L}_{adv}(G, D) = \mathbb{E}_{c_i \sim S_{data}(e)} [\log D(c_i)] + \mathbb{E}_{e_j \sim S_{data}(e)} [\log(1 - D(e_j))] + \mathbb{E}_{p_k \sim S_{data}(p)} [\log(1 - D(G(p_k)))]$$

Cartoon image data
Cartoon image data with edges smoothed out
Generated image form G

content loss:

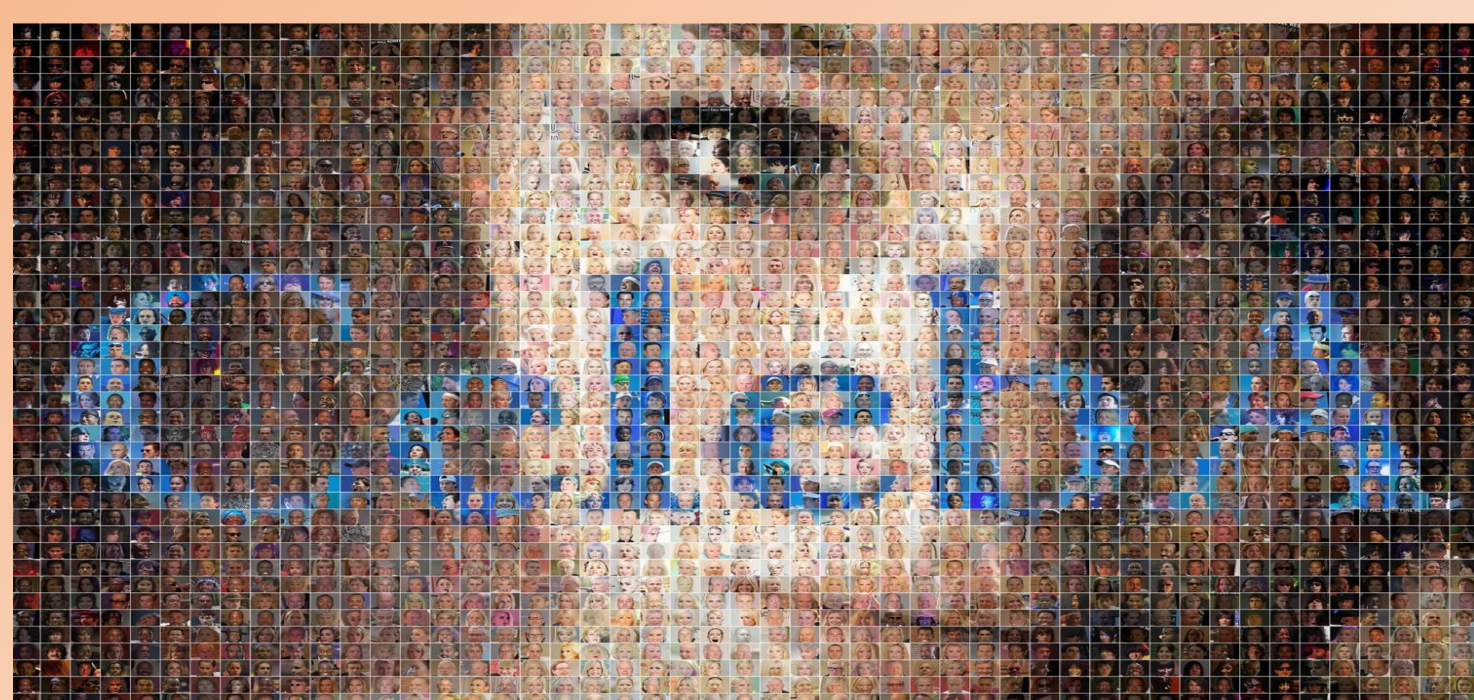
To preserves the image content during cartoon stylization

$$\mathcal{L}_{con}(G, D) = \mathbb{E}_{p_i \sim S_{data}(p)} [\|VGG_l(G(p_i)) - VGG_l(p_i)\|_1]$$

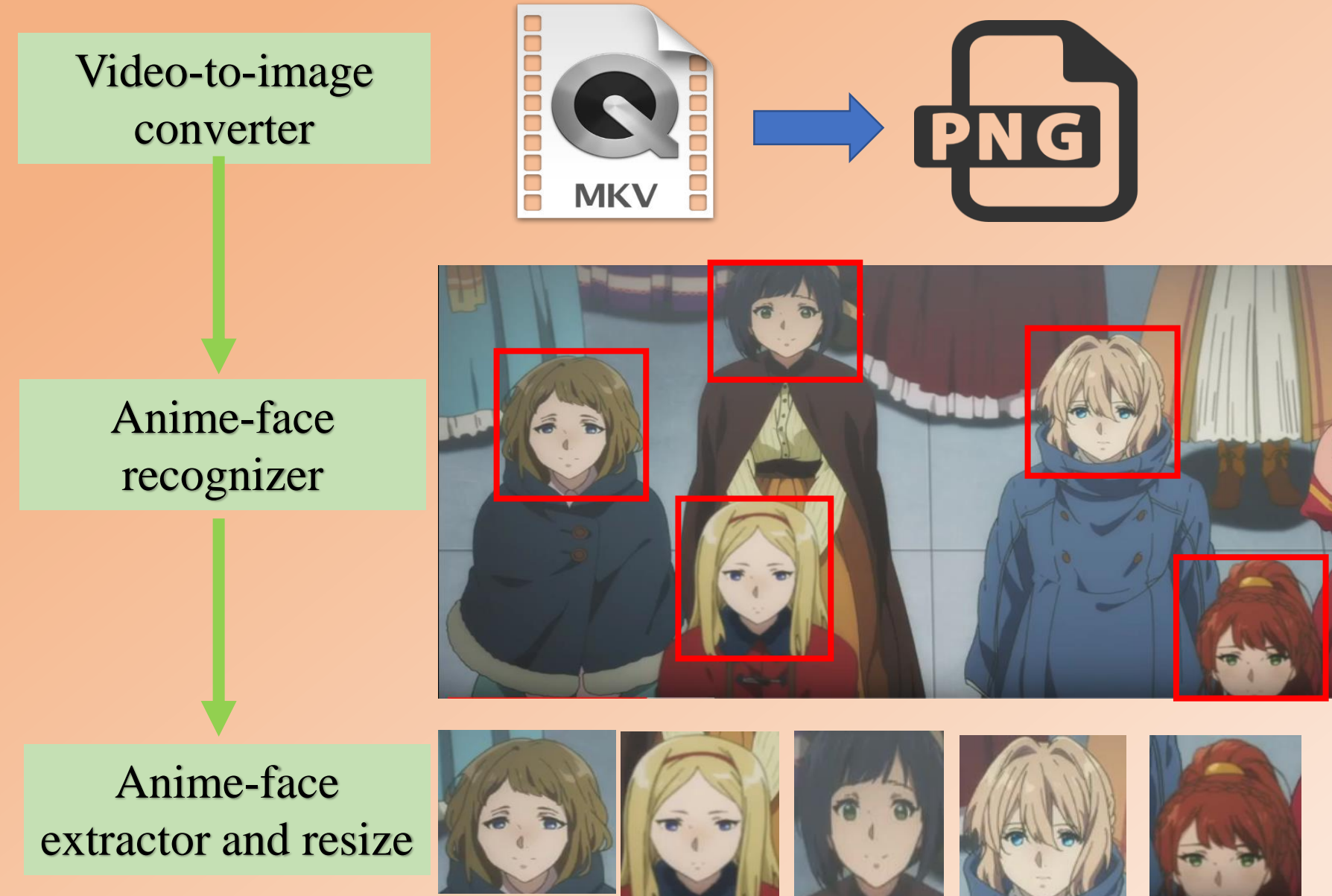
In CartoonGAN, We adopt the high-level feature maps in the VGG networks

Experiment environment and dataset

1. device: NVIDIA K80 GPUs X 2 for three days
2. Human face dataset about 20000 from CelebA
3. About 6000 cartoon image dataset from four animate produced by Kyoto Animation.



Pretreatment for cartoon data



1. Transfer video from mkv format to png format.
2. Fetch the face of all characters.
3. Filter out the image which is smaller 128x128
4. Resize all image to 256x256

Result

1. With less content lambda (5)



2. With normal content lambda (10)



3. Over transfer example



conclusion

We think that the keys to transfer the real image to cartoon style is training dataset and content lambda. The color of transferred image is seem to be light and white due to the pool of artistic style. In addition, larger content lambda may cause disfigured image, but sometime it results in good outcome.

reference

CartoonGAN: http://openaccess.thecvf.com/content_cvpr_2018/papers/Chen_CartoonGAN_Generative_Adversarial_CVPR_2018_paper.pdf
lbpascade_animeface: <https://freedomofkeima.com/pyconid2017.pdf>