

Motivation

- Face-to-Manga translation can be useful in social application as image stylization.
- Lack of paired data makes this a novel and unique task.
- **Solution:** unsupervised translation with CycleGAN.

Problem & Challenges

Dataset:

- **CelebA:** 202,599 face images.
- **Manga109**^{[1][2]}: 10,619 pages with 26,602 characters.
- Cropped, scaled to 64x64px, converted into grayscale.
- **98% training, 2% test.**



"Rising Girl" © Hikochi Sakuya "Aisazu Niha Irarena" © Yoshi Masako

Problem: Translating face images into Manga characters style.

Challenges:

- Unstable learning in GAN.
- Unavailable paired dataset for supervised learning, making evaluation challenging.

Discussion - Next Steps

- Model learns decent face-to-manga translations.
- Spectral normalization makes the images stable and more realistic.
- Self-attention highlights more facial features.
- Pair manga images are not all centered and aligned.

Future work:

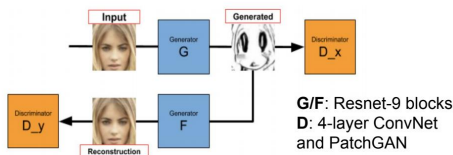
- Make the discriminator simpler to allow more freedom for the generator to style the images.
- Experiment with other generators that can highlight more facial features.
- Train best model longer

Approach

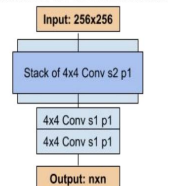
CycleGAN^[3]: GAN Loss + Reconstruction Loss

$$L_{GAN}(G, D_Y) = E_{y \sim P_{data}(y)} [\log D_Y(y)] + E_{x \sim P_{data}(x)} [\log(1 - D_Y(G(x)))]$$

$$L_{Cycle}(G, F) = E_{y \sim P_{data}(y)} [\|G(F(y)) - y\|_1] + E_{x \sim P_{data}(x)} [\|F(G(x)) - x\|_1]$$



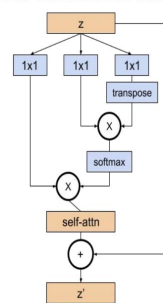
PatchGAN architecture



Spectral Normalization^[4]:

$$SN(W) = \frac{W}{\sigma(W)}$$

Self-attention Module^[5]:

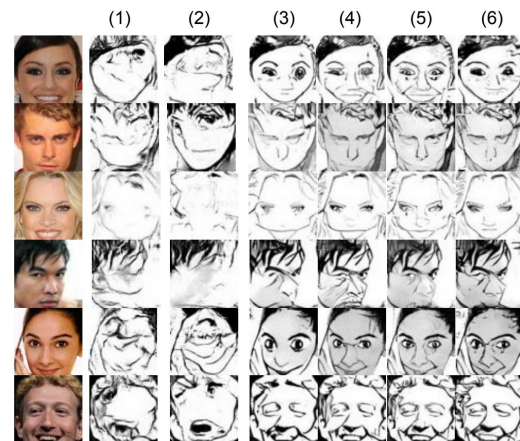


Results

Evaluation for quality and diversity: Inception Score & FID Score

d: depth of PatchGAN

Model	IS	FID
(1) ResNet + ConvNet	3.200	0.497
(2) ResNet + PatchGAN d=3	3.800	0.456
(3) ResNet + PatchGAN d=4	4.866	0.523
(4) ResNet + PatchGAN d=4 + SN	4.520	0.487
(5) ResNet + PatchGAN d=4 + Attn	3.960	0.478
(6) ResNet + PatchGAN d=4 + SN + Attn	4.074	0.523



References

[1] Yusuke Matsui, Kota Ito, Yuji Aramaki, Toshihiko Yamasaki, and Kiyoharu Aizawa. Sketch-based manga retrieval using manga109 dataset.CoRR, abs/1510.04369, 2015.
 [2] Toru Ogawa, Atsushi Otsubo, Rei Narita, Yusuke Matsui, Toshihiko Yamasaki, and Kiyoharu Aizawa. Object detection for comics using manga109 annotations.CoRR, abs/1803.08670,2018.

[3] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks.CoRR, abs/1703.10593, 2017.

[4] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks.arXiv preprint arXiv:1802.05957, 2018.

[5] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks.arXiv preprint arXiv:1805.08318, 2018.