Learning to Cartoonize Using White-box Cartoon Representations Supplementary Material

Anonymous CVPR submission

Paper ID 6791

1. Overview

In this supplementary material, we show more experimental results, including the architecture of generator network and discriminator network, the influence of different loss functions in generative adversarial networks (GANs), illustration of our method in different scenes with different style, and examples used in the user study.

2. Network Architecture

We show the architecture of generator network and discriminator network in Figure 1. The generator network is a fully-convolutional network consisted of only convolution, activation and bilinear-resize layers, which enables it to be easily embedded in edge devices such as mobile phones. PatchGAN [1] is adapted in the discriminator network, where the last layer is a convolution layer. Each pixel in the output feature map correspond to a patch in the input image, with the size equals to the perceptive field, and is used to judge whether the patch belongs cartoon images or generated images. Spectral normalization [2] is placed after every convolution layer (except the last one) to enforce the Lipschitz constrain on the network and stabilize training.

3. Influence of Different GAN loss

In our proposed framework, least square GAN (LSGAN) loss [**?**] is used for adversarial training. We also tested the vanilla GAN loss [**?**], Wasserstein GAN loss with gradient-penalty [**?**], and LSGAN loss with spectral norm removed in discriminator. The results of different gan loss are shown in Figure **??**.

4. Influence of Different GAN loss

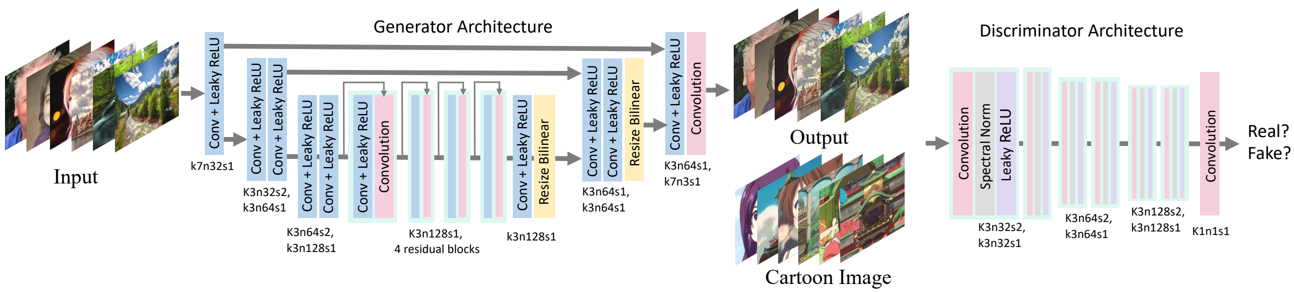


Figure 1. The architecture of generator network and discriminator network.

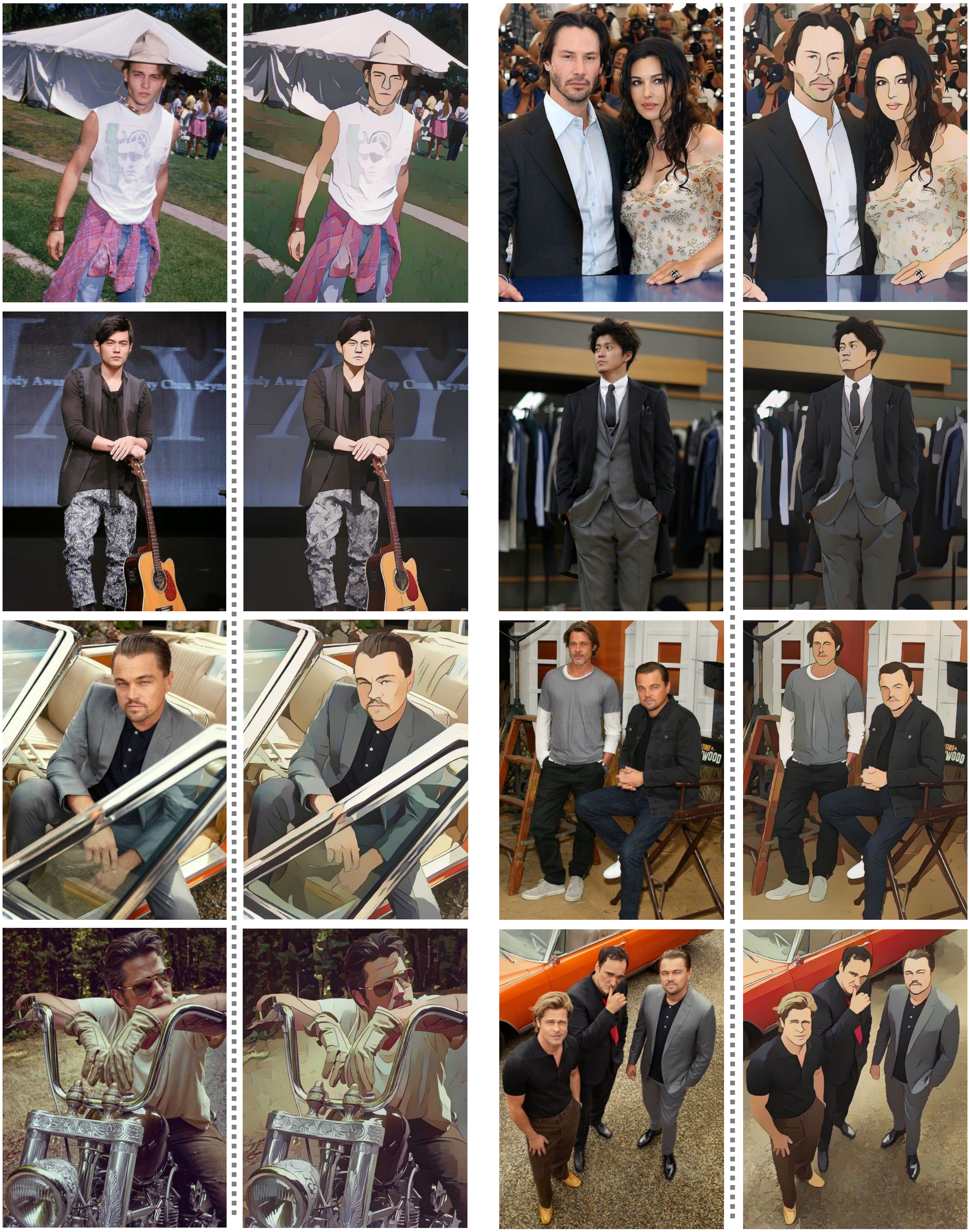


Figure 2. Cartoonized male Celebrities.

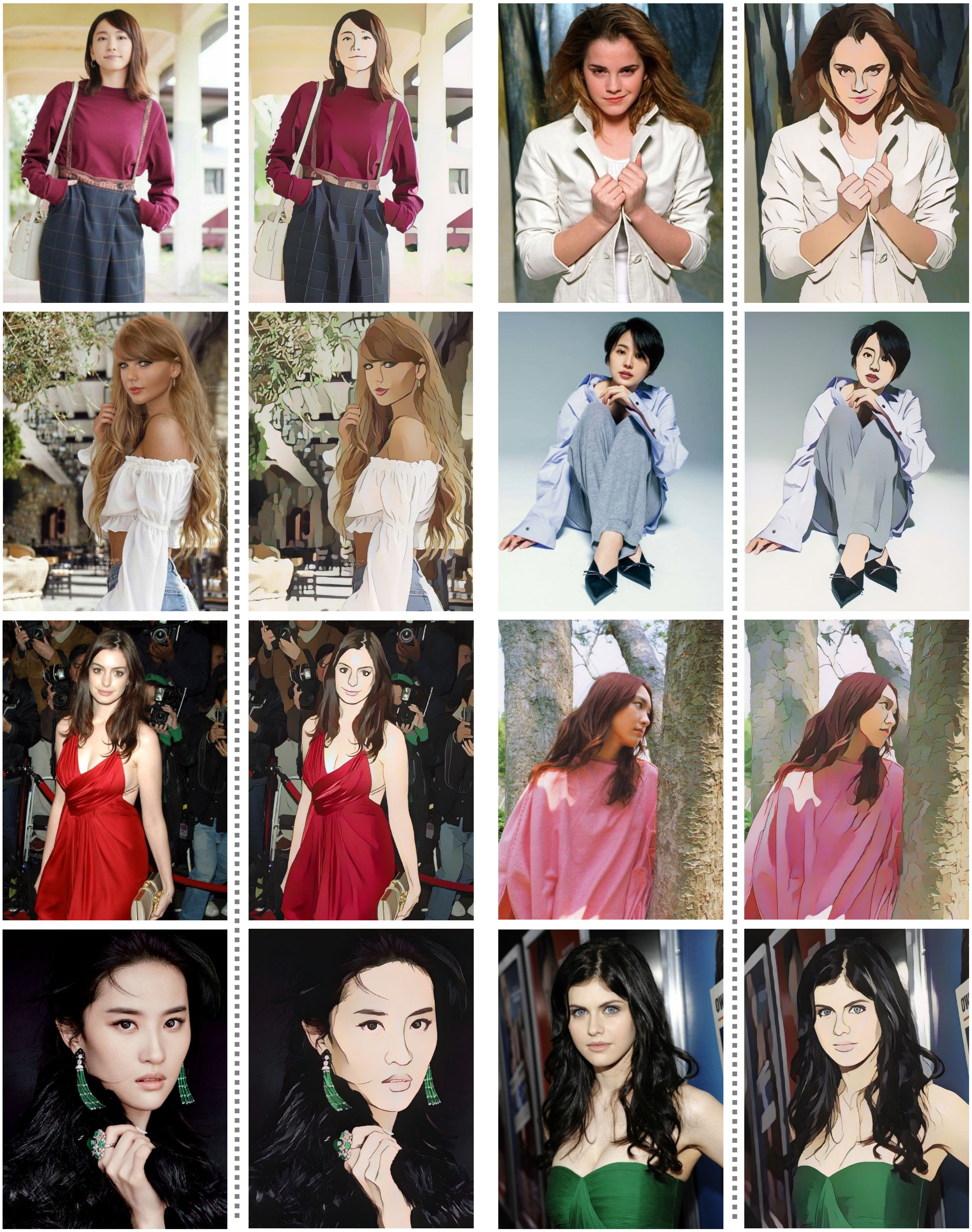
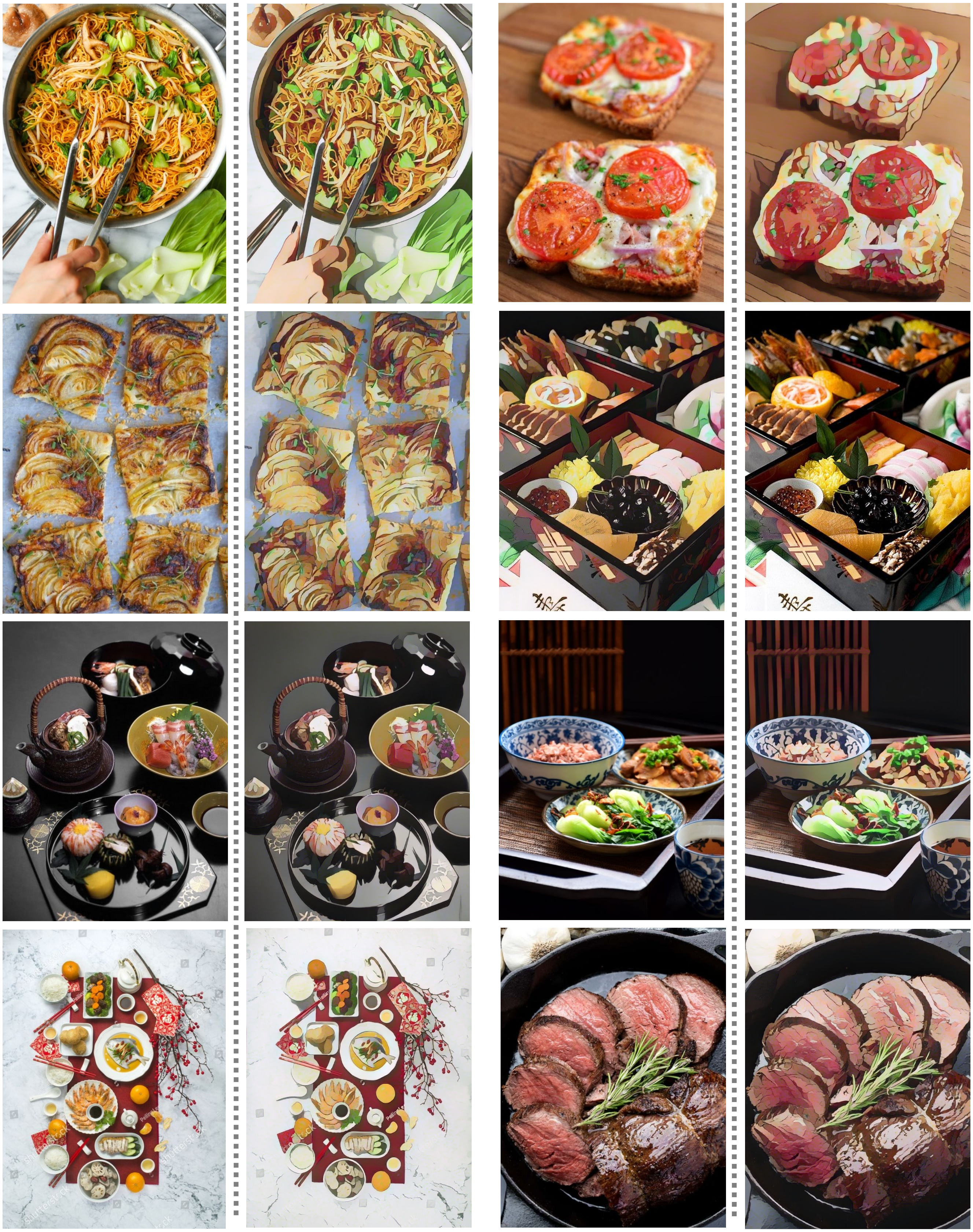


Figure 3. Cartoonized female Celebrities.



Real-world photo | Cartoonized result | Real-world photo | Cartoonized result

Figure 4. Cartoonized food.

Figure 5. Cartoonized scenery.

Real-world photo Cartoonized result Real-world photo Cartoonized result

Figure 6. Cartoonized scenery.

Real-world photo | Cartoonized result | Real-world photo | Cartoonized result

Figure 7. Cartoonized indoor scenes.

Real-world photo

Cartoonized result

Figure 8. Cartoonized city scenes.

CVPR
#6791

Real-world photo Cartoonized result Real-world photo Cartoonized result

Figure 9. Cartoonized city scenes.

References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017. 1

[2] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018. 1