SUPPLEMENTARY MATERIAL FOR A CROSS-LAYER BASED NETWORK FOR FASTER IMAGE GENERATION

Zhaoyu Zhang, Yuechuan Sun and Jun Yu

Department of Automation, University of Science and Technology of China

ABSTRACT

In the main text we derived the cross-layer CNN architecture for the generator in generative adversarial networks (GANs) and demonstrated the superiority over two state-of-the-art GANs, namely deep convolutional GAN and Wasserstein GAN-GP. This document provides more details about the experimental results on the Fashion-MNIST and CelebA datasets.

1 Results on DCGAN

1.1 Fashion-MNIST Dataset

In the main text, we only presented part of the generated images. In this section, more intermediate results on Fashion-MNIST are given for the presentation of visual superiority of our method over time. The curves for the training loss are not given here for brevity and can be found in the main text.

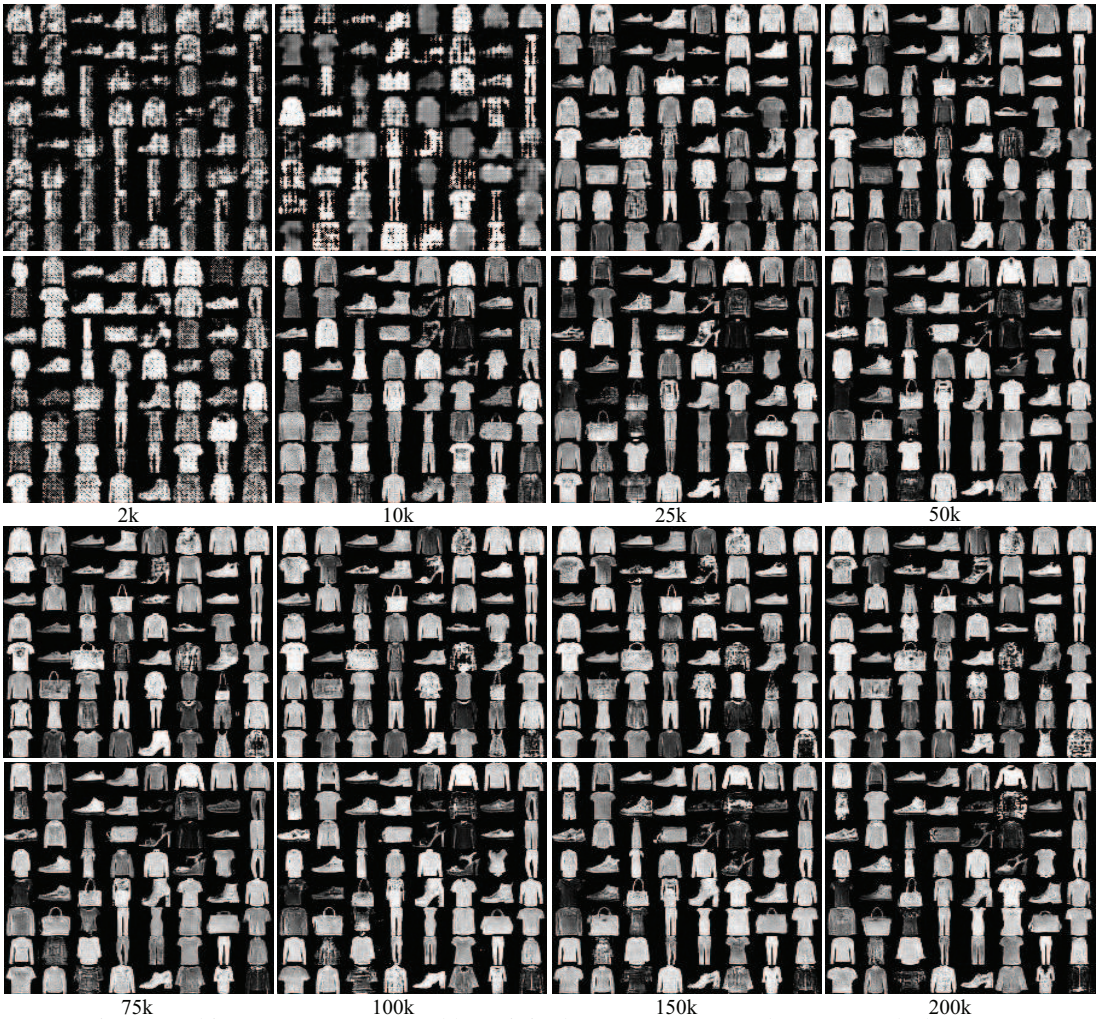
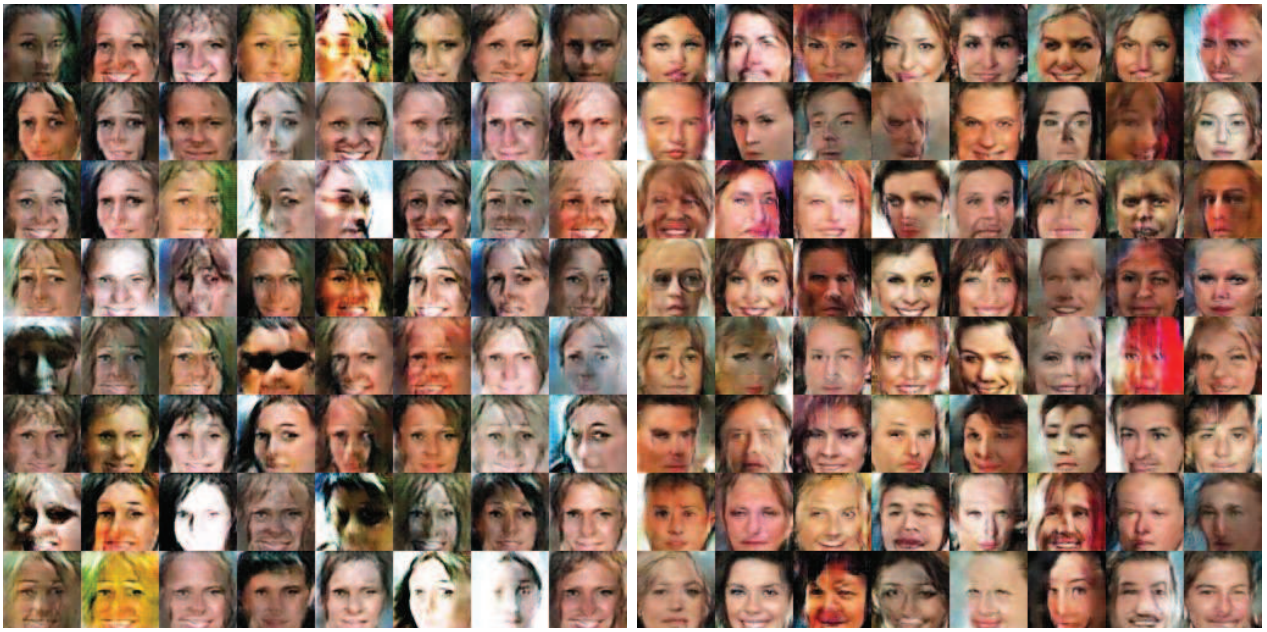


Fig. 1: Fashion-MNIST generated by original DCGAN (top) and our network (bottom).

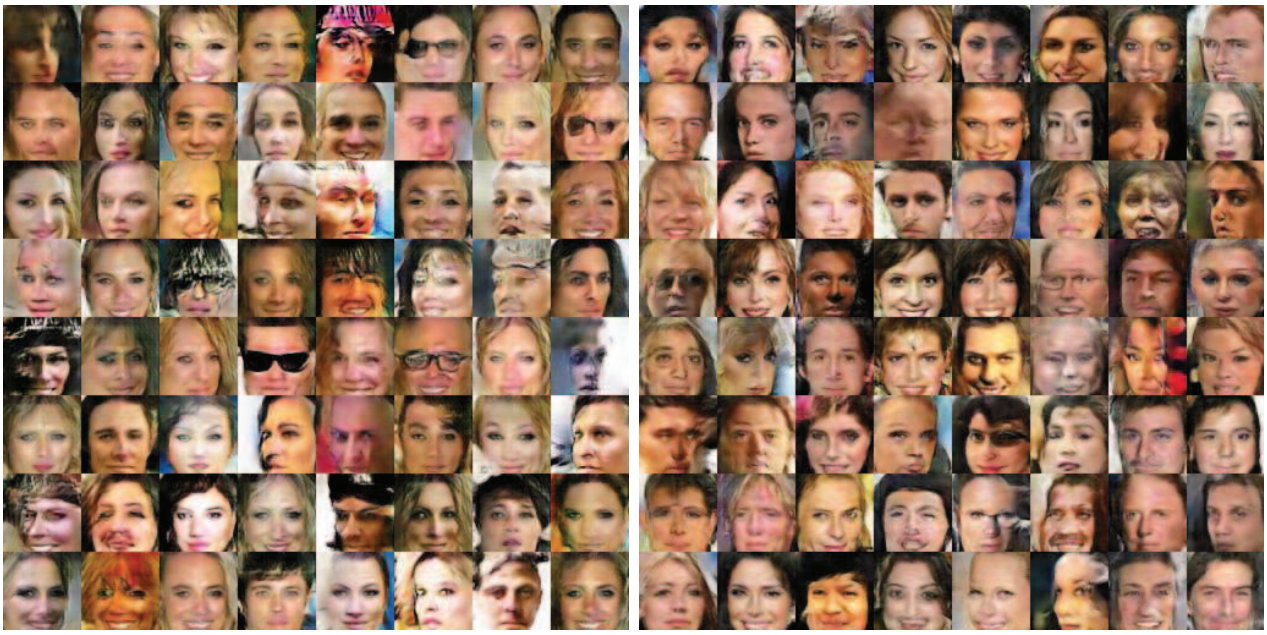
1.2 CelebA Dataset

Due to the space limitation, we only gave 25 faces at each step in the main text. Here, we show more generated samples to fully demonstrate the superiority of our network. The batchsize was set to 64 and thus 64 faces are generated at each step. It can be observed that the synthesized images by our network on the right carry more details and variations, resulting more genuine faces. The curves for the training loss are not given here for brevity and can be found in the main text.

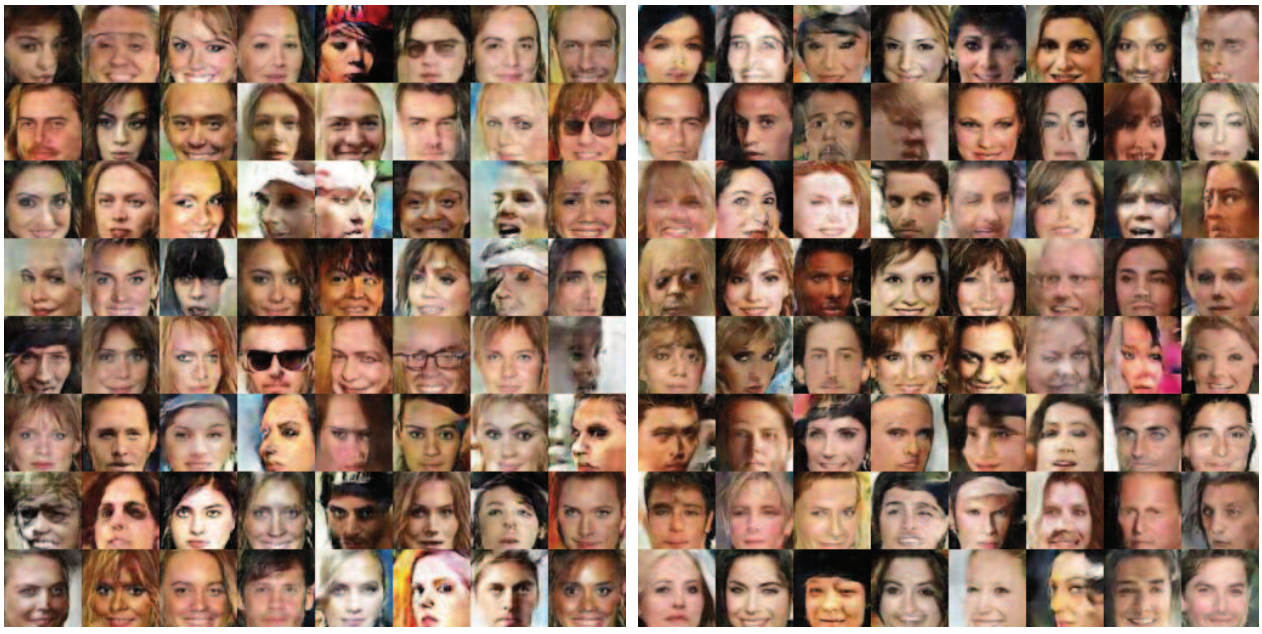
1k



3k



8k



16k

Fig. 2: Faces generated by DCGAN (left) and our network (right) at given steps.

2 Results on WGAN-GP

Fig.3 presents the face images generated by WGAN-GP and our network at different steps. The batchsize was set to 100 and we therefore obtain 100 faces at each step. Again, our network gives more confusing face images early in training, which indicates that the proposed architecture speeds up image generation. The curves for the training loss can be found in the main text.

1k

3k

8k

16k

Fig. 3: Faces generated by WGAN-GP (left) and our network (right) at given steps