Supplementary Material

Hyojin Park[1]
wolfrun@snu.ac.kr
Youngjoon Yoo[2]
youngjoon.yoo@navercorp.com
Nojun Kwak[1]
nojunk@snu.ac.kr

[1] Department of Transdisciplinary Studies,
Seoul National University,
Republic of Korea.

[2] CLOVA AI Research, Naver Corp.,
Republic of Korea.

1 Additional Experiments on MC-GAN

In this supplementary material, we show the additional experimental results on the Caltech-200 bird [3] and Oxford flower-102 [2] datasets which are widely used in text to image Generative Adversarial Network (GAN). We show additional examples that we could not include in the Experiment section due to space limitation.

1.1 Another Example of Stack GAN

We give another synthesized examples of MC-StackGAN by Oxford-102 flower dataset. We also got satisfied images like Caltech-200 bird data.

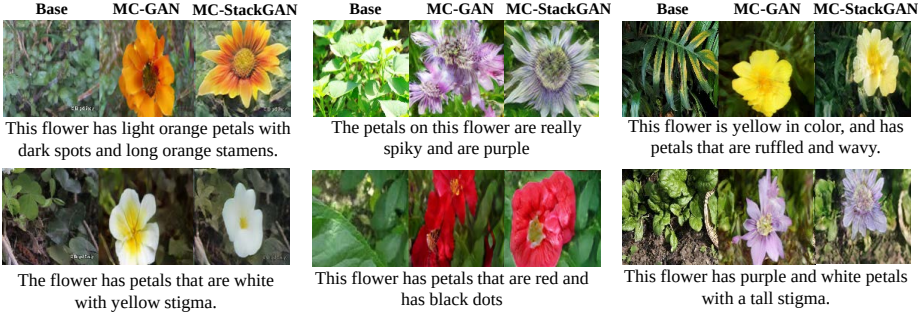

Figure 1: Examples of MC-GAN and MC-StackGAN. From left to right: Base image, the synthesized image from MC-GAN (128×128), the synthesized image from MC-StackGAN (128×128) on Oxford-102 flower test dataset

1.2 Comparison using the Synthesis Problem in [1]

Here, we show another examples of multi-modal conditional problem defined in [1]. We used Oxford-102 flower dataset and performed experiments with the same setup as for the

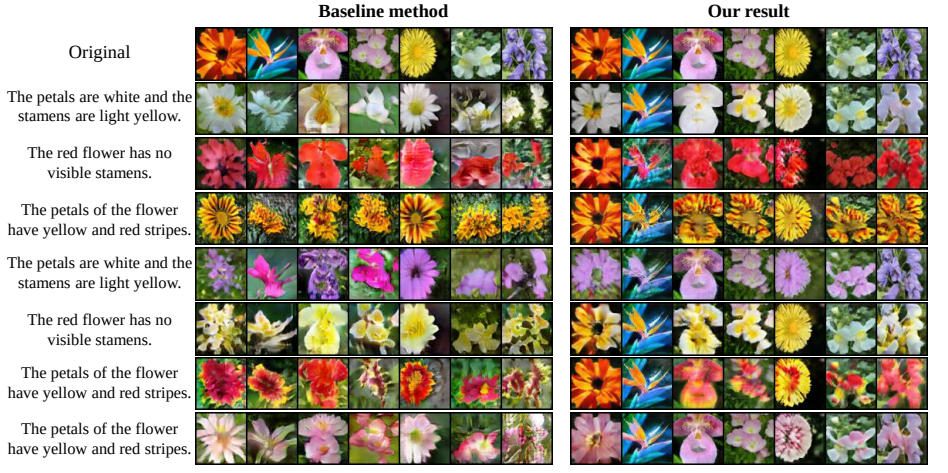

Figure 2: Example results of the image synthesis problem in [11]. The baseline method [11] and our method (MC-GAN without mask) on Oxford-102 flower dataset

Caltech-200 bird data. Ours also showed good performance on flower dataset. The style of the flower changed as the text description, the original shape was well maintained.

1.3 Another Example of Variety and Stability Experiment

Figure 3: Variety and stability of MC-GAN. The images are generated by linearly interpolating the noise vector z from all-zero to all-one vector on Caltech-200 bird test dataset

We generated more images using linearly interpolating the two noise vectors, which are all-zero (z_0) and all-one (z_1) vectors, under the same text and more various image conditions. For this case, we could also get around half of visually successful samples as can be seen in Fig 3.

References

[1] Hao Dong, Simiao Yu, Chao Wu, and Yike. Guo. Semantic image synthesis via adversarial learning. In *Int. Conf. on Computer Vision (ICCV)*, 2017.

[2] M-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing*, Dec 2008.

[3] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.