

Style-Guided Inference of Transformer for High-resolution Image Synthesis

Supplementary Material

Appendix A. Implementation Details

Basically, we followed hyperparameter settings from the baseline paper with a few modifications. In Table 1, we present major hyperparameters with only two hyperparameters varying. We trained models on 4 GPUs with a batch size of 1 per GPU.

Appendix B. Random Sample Images of Baseline

In our paper, we used random outputs of the baseline to run style transfer methods. For reference, we include those random outputs in Figure 1.

| Experiment | input size | n_{layer} | n_z | $ \mathcal{Z} $ | length(s) |
|----------------|------------|-------------|-------|-----------------|-----------|
| Flickr | 256 | 24 | 256 | 2048 | 512 |
| WikiArt | 256 | 24 | 256 | 1664 | 256 |

Table 1. Hyperparameters for training. All the other settings not in this table is same as the baseline paper.



Figure 1. Random outputs from the baseline method. We used these images to apply style transfer.