~:Terminologies used and Some comments:~

Comments:

1. **Paper is successfully able to eliminate the water droplet like artifacts left behind such that Discriminator was not able to identify it. This was improved in this paper which is an amazing improvement over StyleGAN 1.**
2. **Model is designed so meticulously that, as the model grows and we learn more about the features in subsequent layers, skip connections are introduced in a way that does not drop high level features. This is where the model outperforms in quality and quantity as well. The images generated are high quality as well as faster convergence. The output starts with low resolution images and as the network learns more features it starts using high resolution feature layers. Thus skipping previous layers which had low level features. This way it has some kind of unsupervised mechanism inherently.**
3. **I would insist the reader to visit whichfaceisreal.com for hands on with StyleGAN2 and check how this model is able to beat time limited humans. (website)**
4. **GPU capacity practically required is 16GB ($2500) which is bare minimum for the training. Thanks to colab pro which made this available for me. (for $10)**
5. **At some point I found there is entanglement observed again in SG2 for 64x64 images training. However this was not a problem when trained with higher dimensions.**
6. **You may need to fluctuate the regularizer in the range 0 to 6 for facial features changes and 6 to 18 for color changes if used Face image datasets. Recall coarse styles, middle styles and fine styles from StyleGAN1.**
7. **I really enjoyed the training from scratch and then visualizing the intermediary images generated after 10 EPOCHS. And seeing this AI to converge near the training dataset.**
8. **For future scope, there can be a storehouse or marketplace for these pretrained pickle files for StyleGANs such that we can have pretrained pickles for multiple datasets. And one can use anyway they want to generate the images directly from the pretrained models. In addition to that, one can use transfer learning on it as well.**
9. **It is advised to have unlimited cloud storage of at least 100 gigs for the data and a fast GPU with larger running time. (Best combination will be colab pro and academic google drive with no bound on storage)**
10. **My Configuration:**

(snapshot in colab runtime)

Runtime	Colab pro
Python Version	3
TensorFlow	1.x
HIGH-RAM	25GB
GPU	16GB

(snapshot from google colab FAQs)

CPU	GPU
Intel Xeon Processor with two cores @ 2.30 GHz and 13 GB RAM	Up to Tesla K80 with 12 GB of GDDR5 VRAM, Intel Xeon Processor with two cores @ 2.20 GHz and 13 GB RAM

11. **Weight Demodulation:**
    a. Using experimentation, simplify how the constant is processed at the beginning.
    b. The mean is not needed in normalizing the features.
    c. Move the noise module outside the style module.
12. **Lazy Regularization: Performing regularization at every step but not calculating the metric cost at every step. Thus saving training time. It is calculated after 16 mini batches.**
13. **Path Length Regularization: Used for GANs performance. This is actually a kind of Perceptual Path Length analog where you calculate the distance between 2 images should be approximately the same in feature space or latent space. A penalty is added when the images distance differ when the distance between images and distance between latent features are different. In other words, we want the 2 feature maps to be as close to each other in feature space or latent space as they are in image space.**
14. **Progressive Growth: It is analog to Progressive GAN. Going to great depths with skip connections. As we grow deeper we get higher features which are more important in generation of high quality images than those of the early layers. Thus skipping connections is important. (also check point number 2)**
15. **Large Networks: This is where we concentrate on deeper layers to produce high quality images.**
16. **Removing Progressive growth is actually a kind of change here for getting better results with skip connections. This was done to remove water droplet like features when generated in resultant images. Hence this progrssive growth is replaced with a skip connection mechanism without compromising image quality.**