

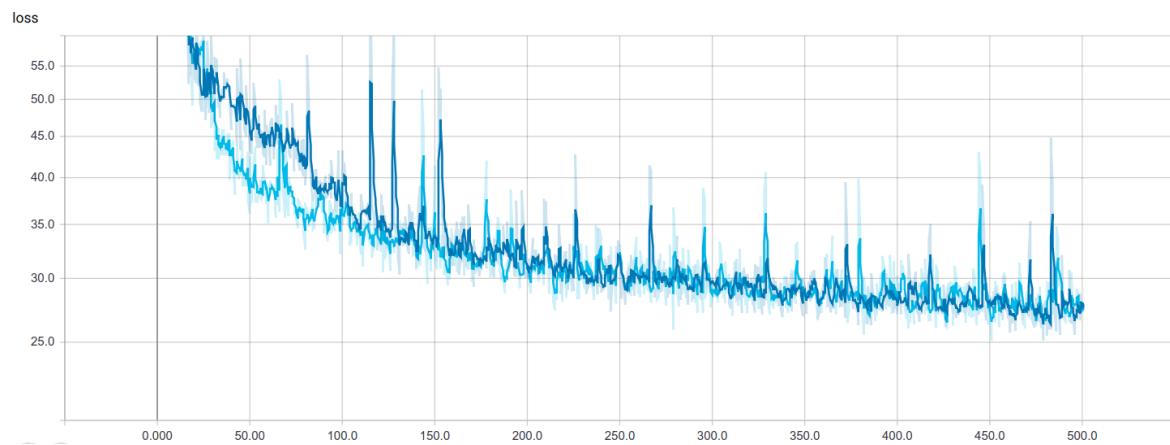
DL HW3

A072032 紀伯翰

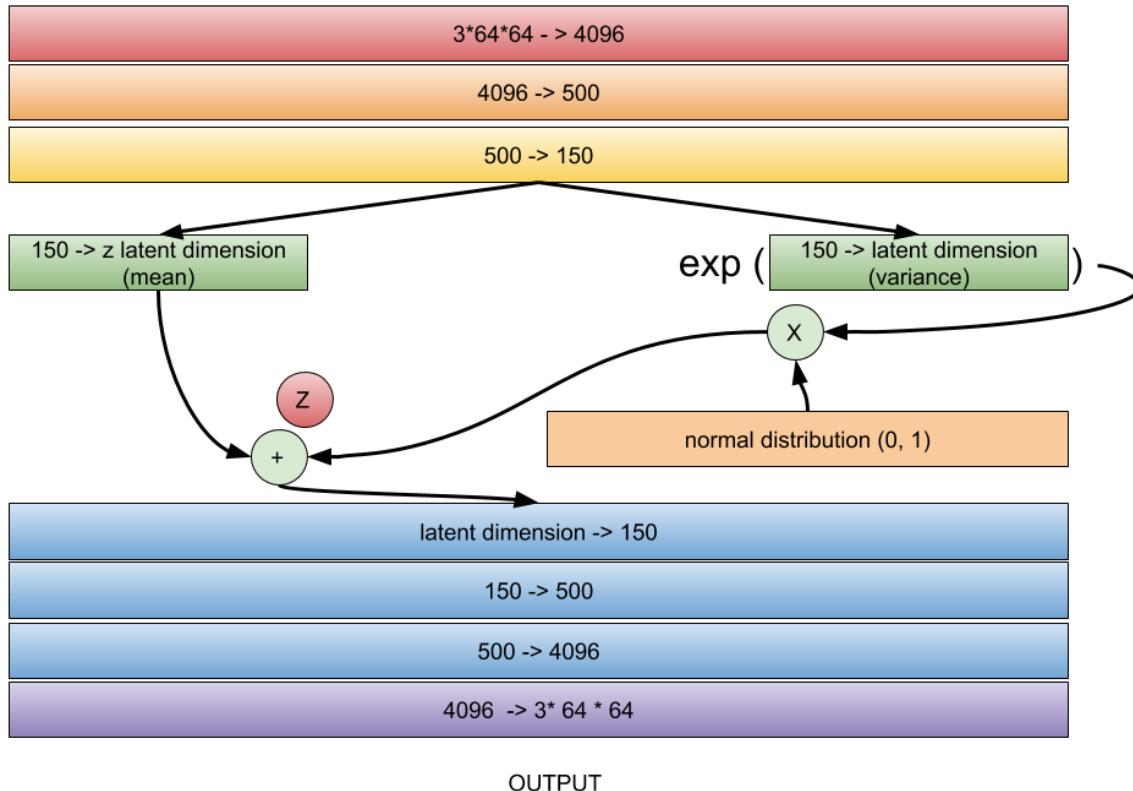
1. VAE

Describe in details how to preprocess images (such as resize). Implement a VAE for image reconstruction by using convolution layers or fully connection layers. You need to design the network architecture and show it in the report. Finally, plot the learning curve in terms of loss function or negative evidence lower bound.

1. learning curve:



There are two model which I sample when I tune the hyperparameters.



Above picture will showed my model architecture detail.

Dark blue one I set the number of latent dimension number is **5**.

Light blue one I set the number of latent dimension is **10**.

learning rate is **1e-3** on both.

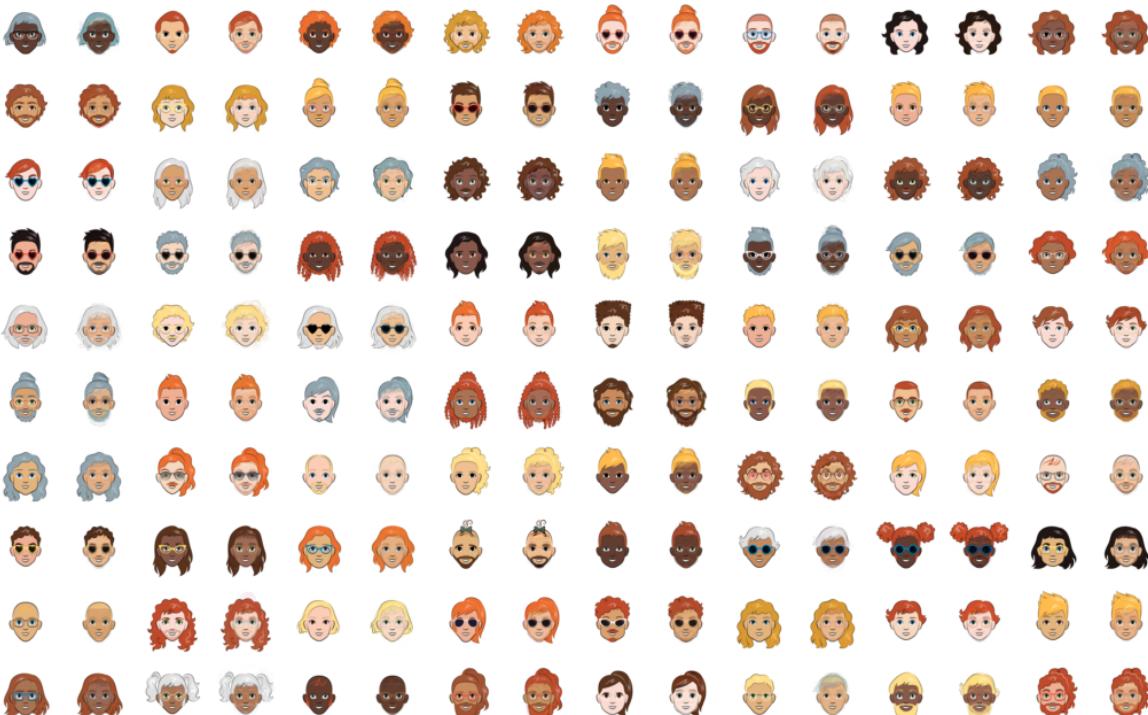
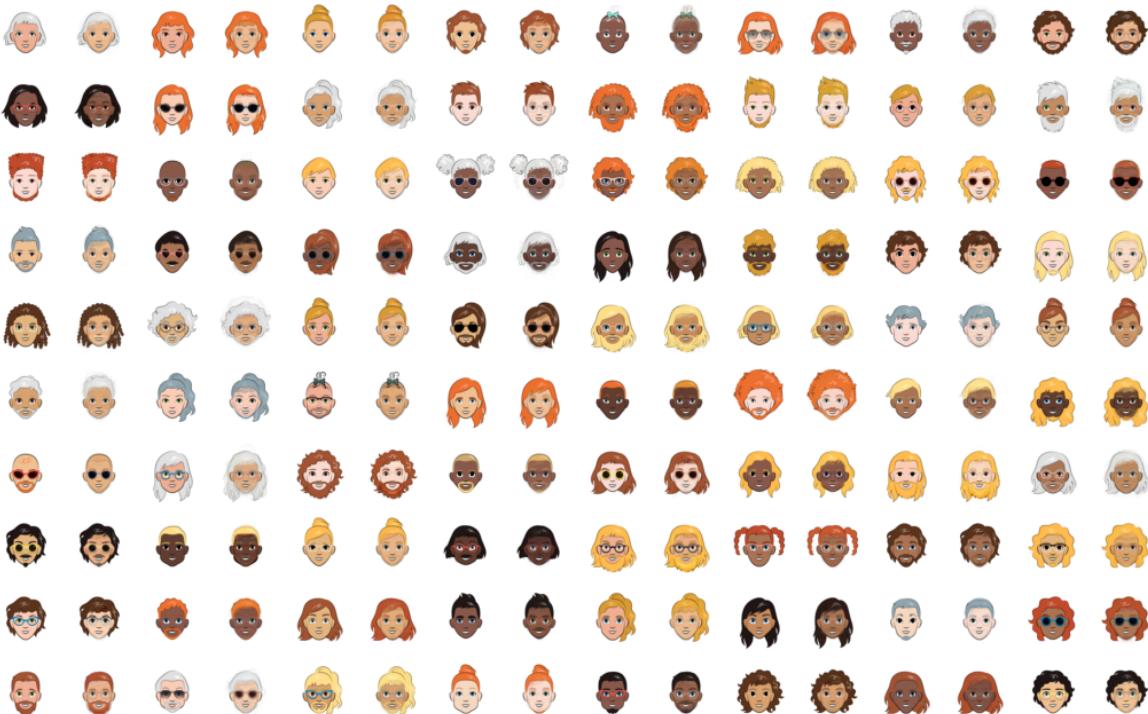
All I use are **fully connected** layer to build my model.

2. Reconstruct Images:

Below shows two results which sampled from two model above.

pohan

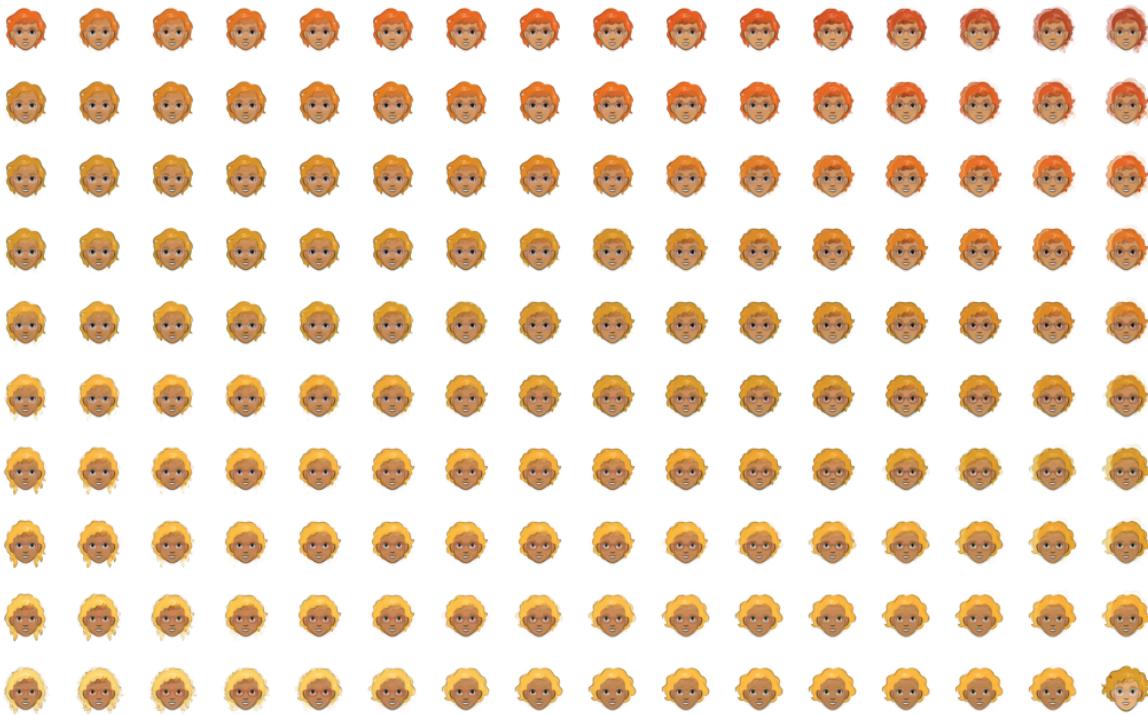
3/12



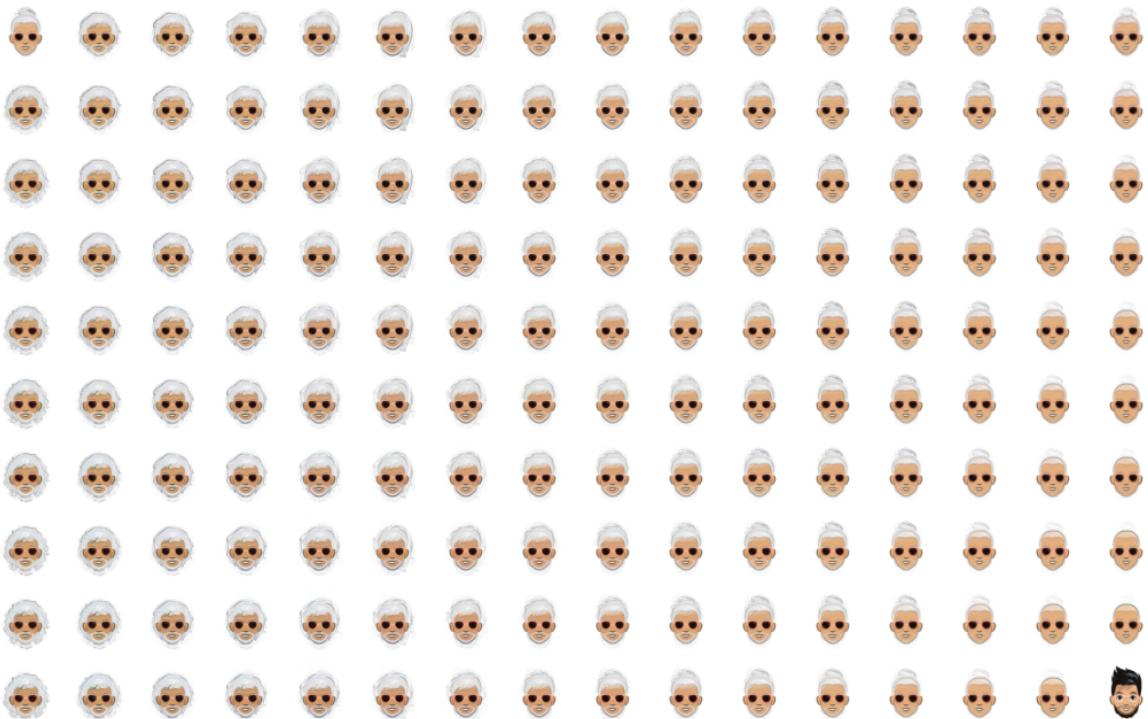
3. Sample Data from latent space:

pohan

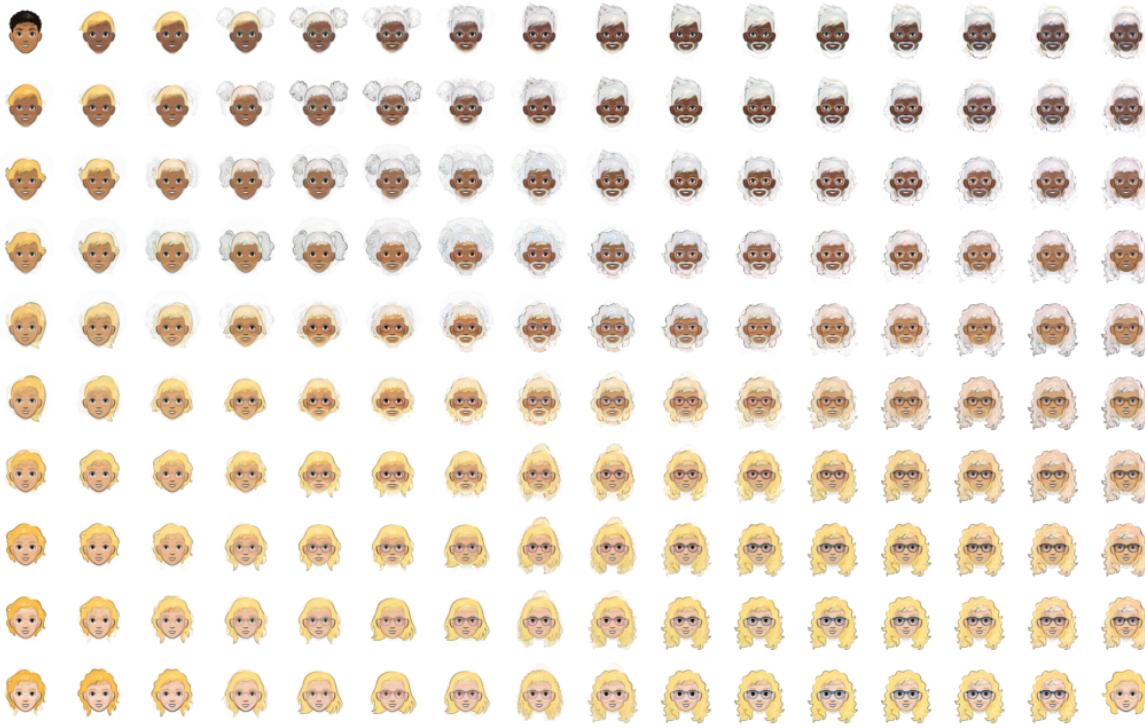
4/12



I take 5 dimension latent space of 2 image. fix index 2 to 4 dimension value and sample latent space between two latent vector.



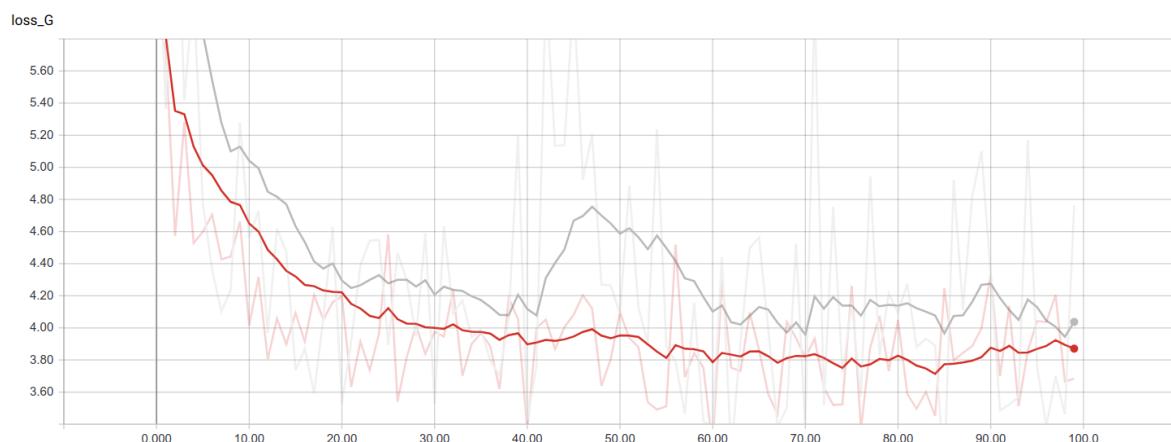
I take 5 dimension latent space of 2 image. fix index 0 ,2, 4 dimension value and sample latent space between two latent vector.

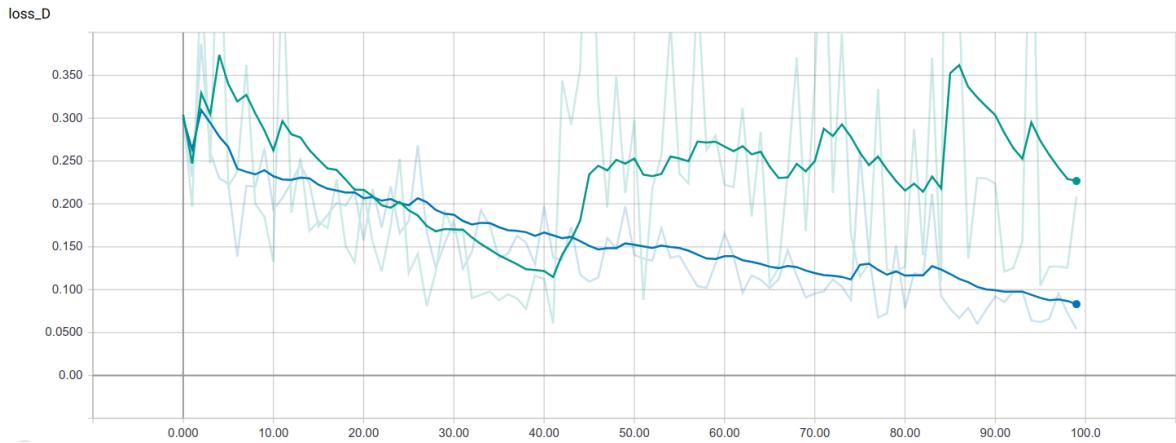


The final image come from just sample latent vector on distance between latent vector of two image.

2. Style Transfer

1. Construct a cycleGAN with the loss function below. Plot the learning curve of both generators and discriminators. You can sum up the loss of two generators and plot in one curve.





Above shows the results on two model with scalar lambda is 10 and 5.

Architecture I select 2 residual blocks in the G and D models.

batch size: 100

learning rate: 1e-4

amazing thing is that the scalar on loss consistency term. If I make scalar(lambda) higher, the picture quality will be higher on animation output.

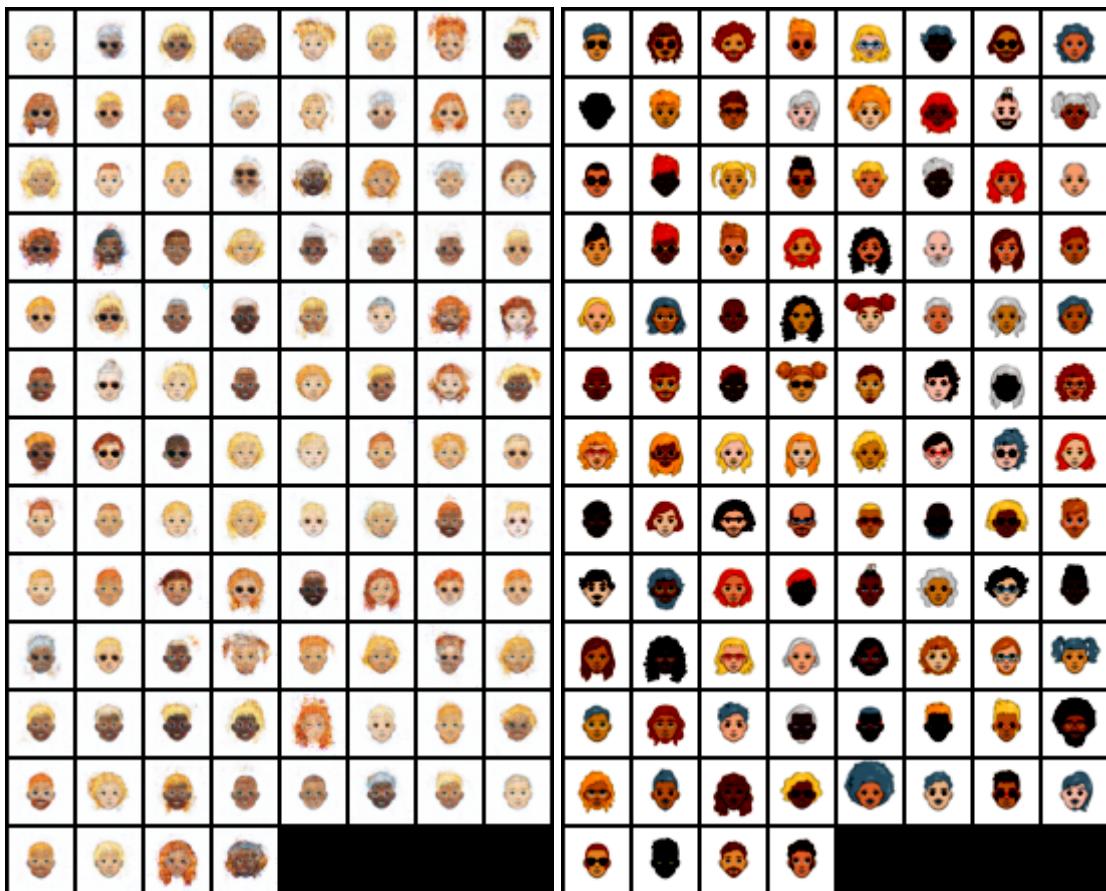
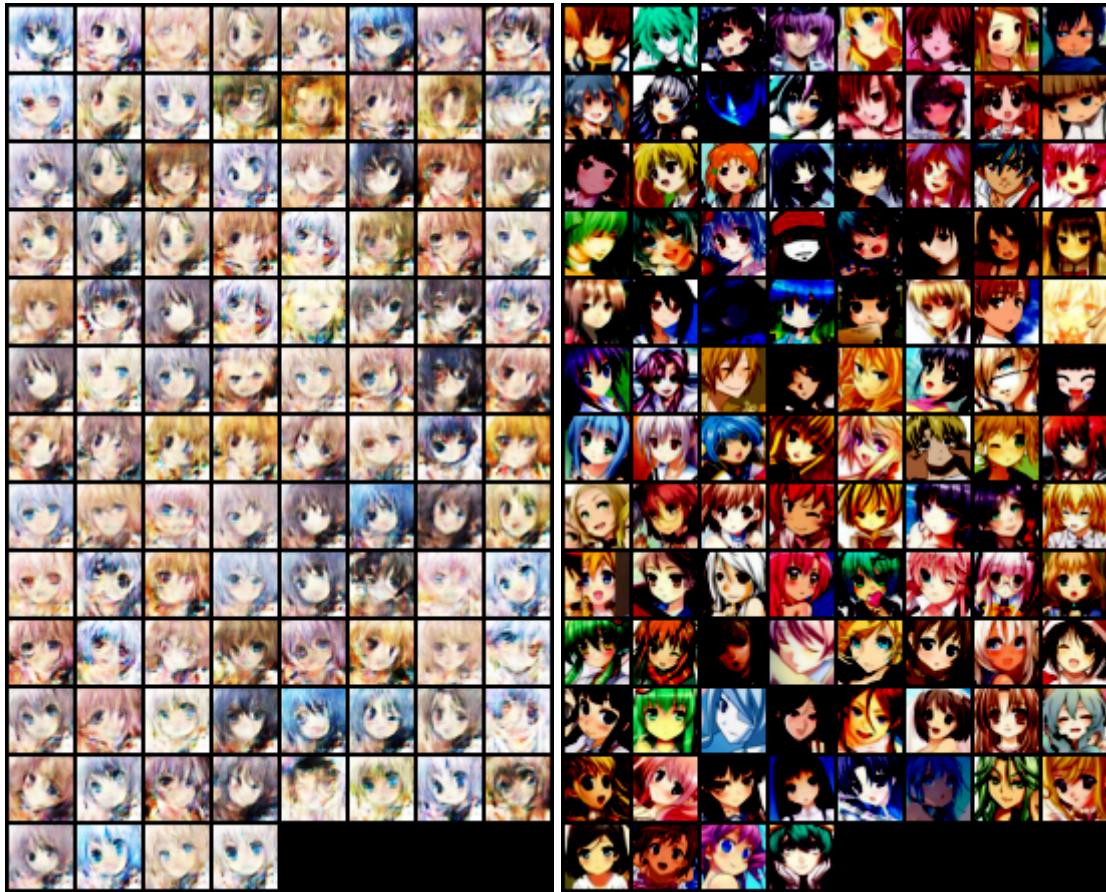
2. Please sample some cartoon images in animation style and animation images in cartoon style. Show your results and make some discussion in the report.

fake

real

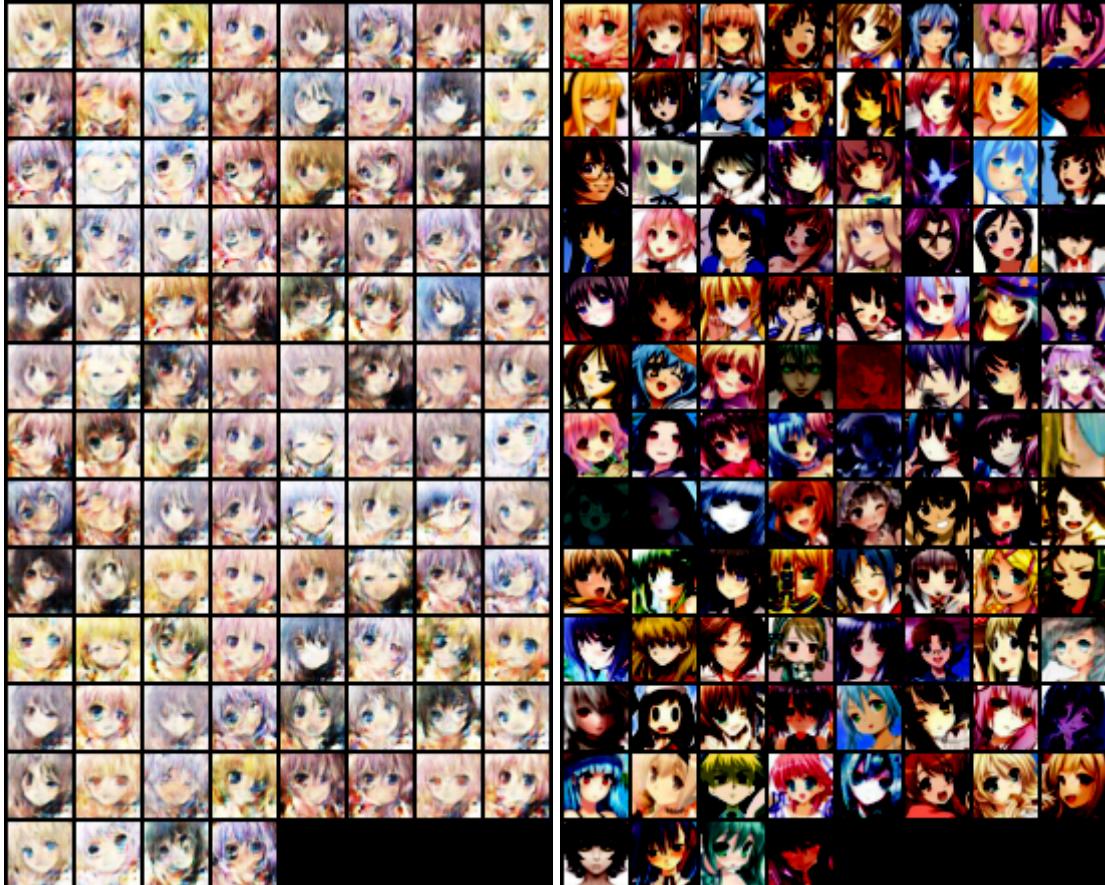
pohan

7/12



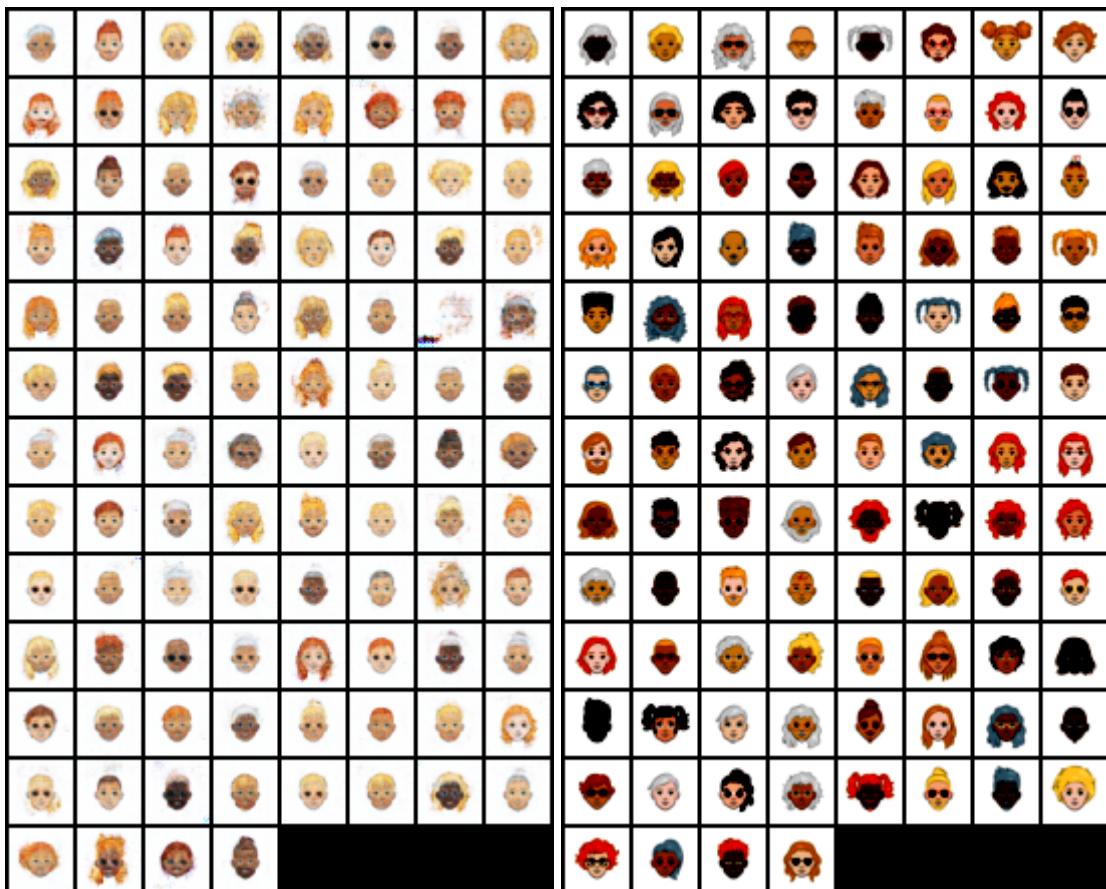
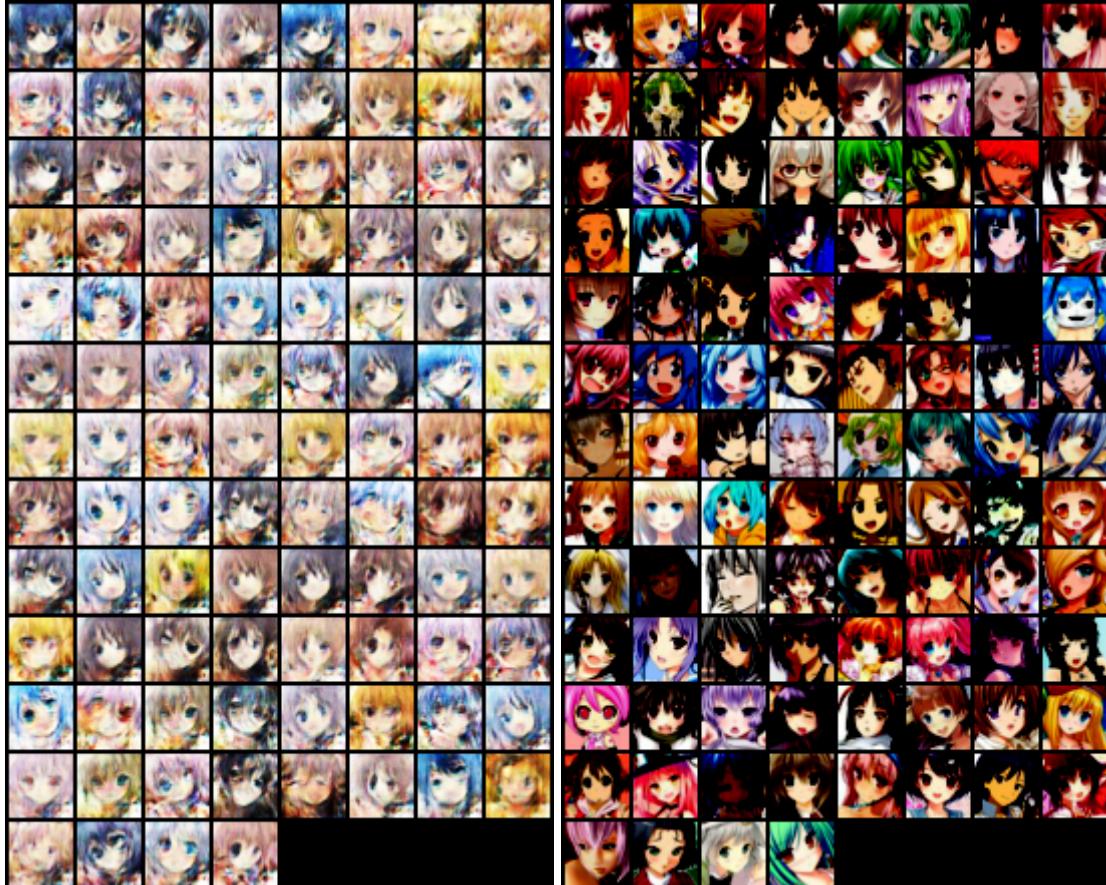
pohan

8/12



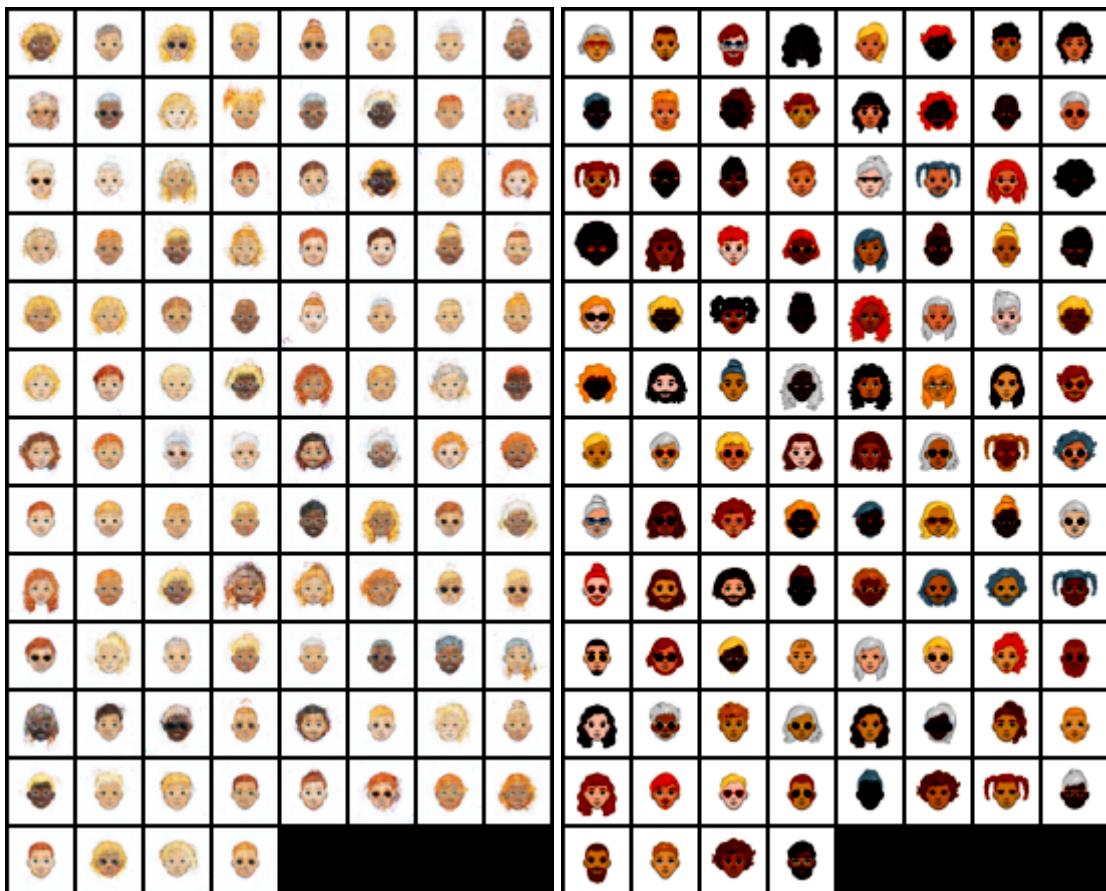
pohan

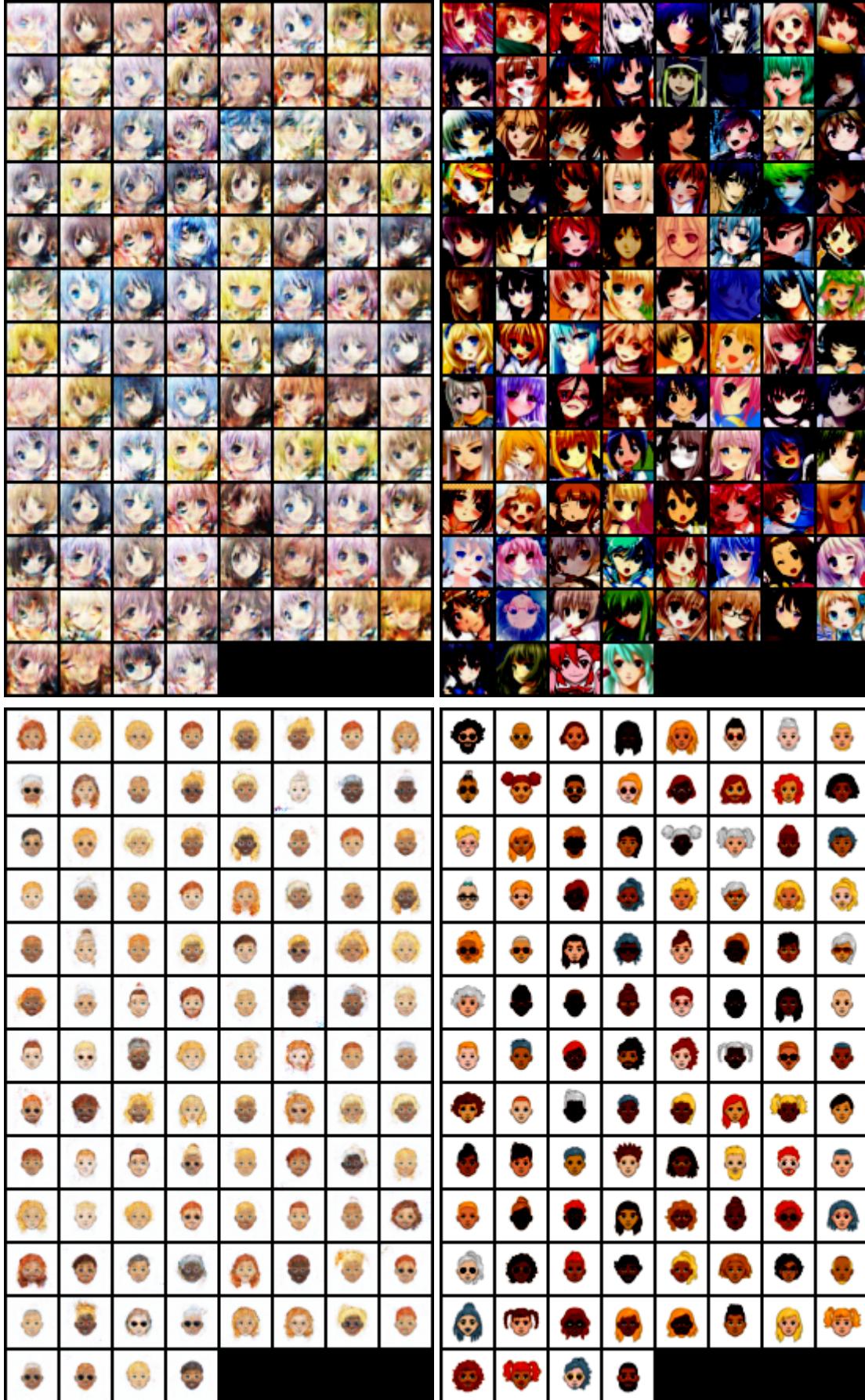
9/12



pohan

10/12





3. Briefly describe what is mode collapse. According to (ii), is mode collapse issue serious in this task? Why?

Mode collapse is that the output will all tend to generate same image or few kind image. In this situation, I think that it is serious on this task because it have lots of

different kind on animation dataset set. On the training, mode collapse will happen on the very beginning and after 100 epoch. In the appropriate training, we will get better performance on 40 to 90 epoch. During these epoch, the image will be more diverse and not convergence on only one to five image.