CIS 680 Final Project Report

Kun Huang, Zhihao Ruan PennKey: $huangkun, ruanzh$

1 Visualization of Training

We trained our network for 20 epochs and got the loss curve as follows.

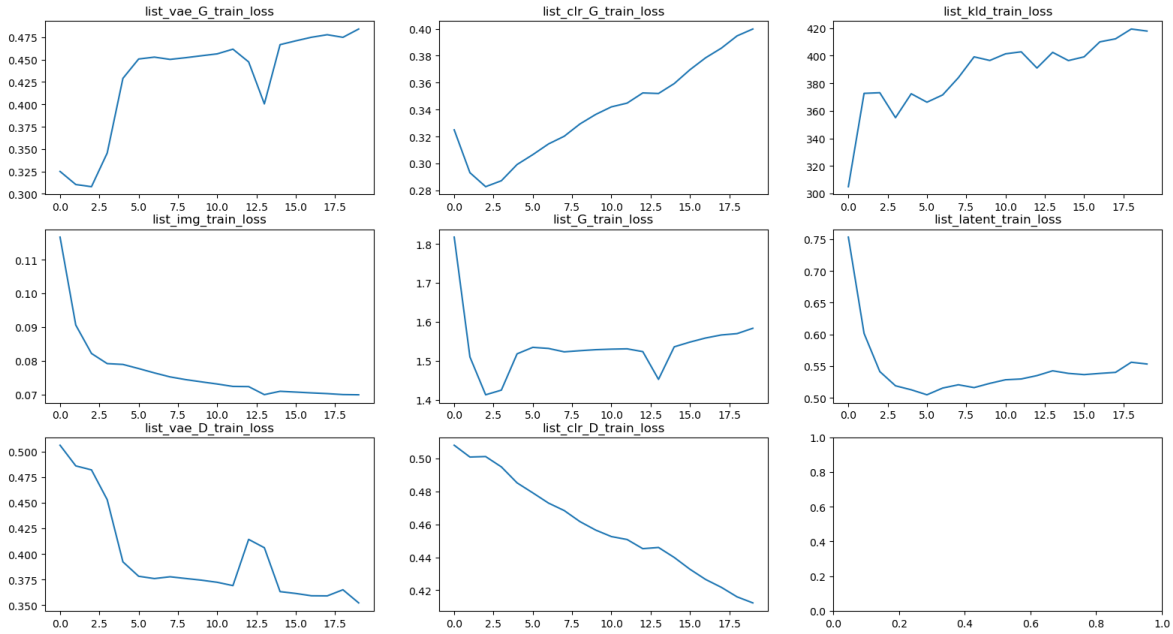
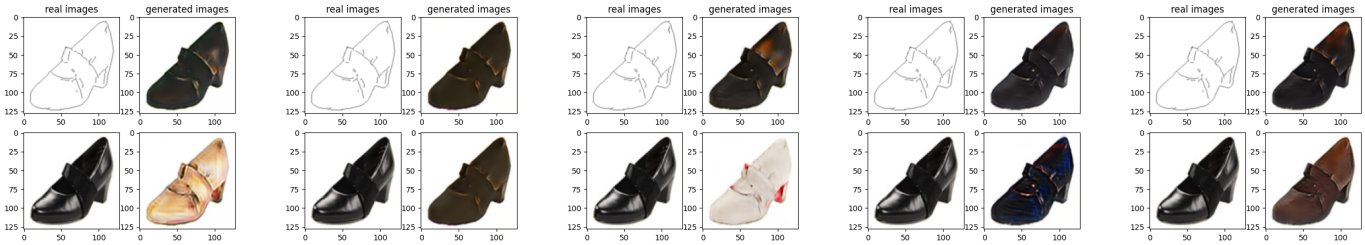
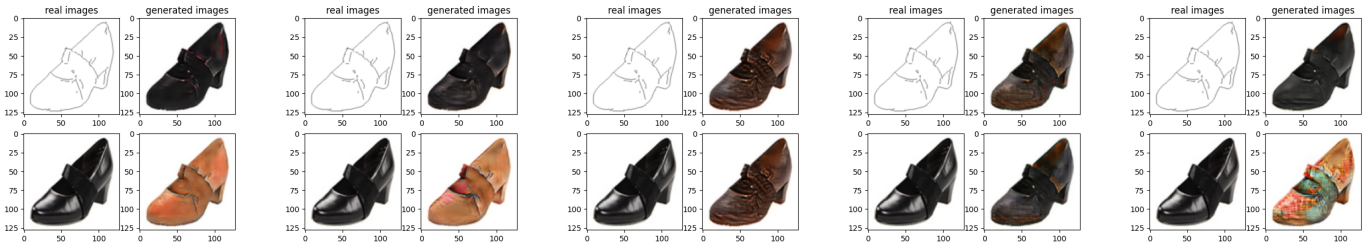


Figure 1: Plot of different training loss in one figure

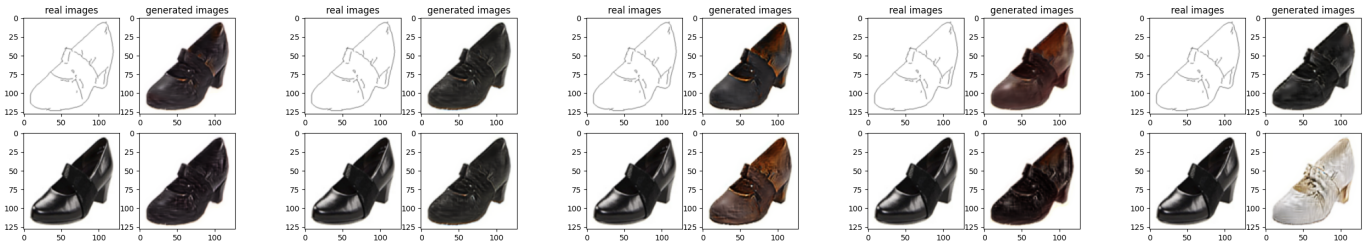
What's more, we also selected part of the data in training set and got the following visualization results:



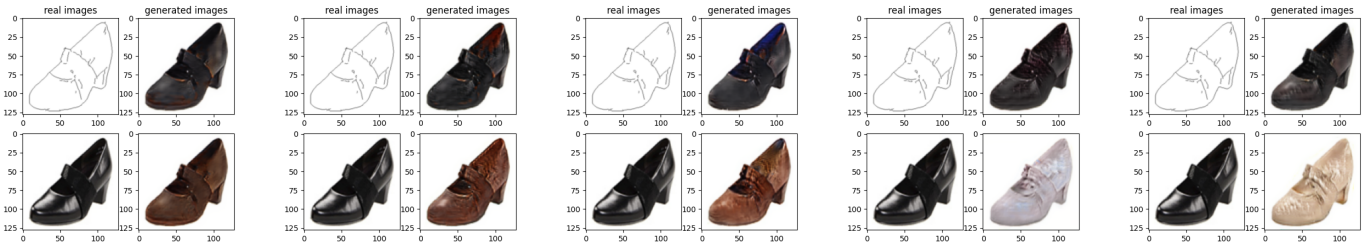
(a) Epoch 0 (b) Epoch 1 (c) Epoch 2 (d) Epoch 3 (e) Epoch 4



(f) Epoch 5 (g) Epoch 6 (h) Epoch 7 (i) Epoch 8 (j) Epoch 9



(k) Epoch 10 (l) Epoch 11 (m) Epoch 12 (n) Epoch 13 (o) Epoch 14



(p) Epoch 15 (q) Epoch 16 (r) Epoch 17 (s) Epoch 18 (t) Epoch 19

Figure 2: Visualization of a shoe across training epochs. Top left: image A; bottom left: image B; top right: encoded B; bottom right: random B.

From the visualization results we can see that as the network is trained, our generator produces more and more similar shoes to ground truth B in encoded B and learns to produce a variety of colors in random B.

2 Quantitative Evaluations

2.1 FID Score and LPIPS

We used FID score to evaluate the photo-realistic quality of the outputs of our generator. As a matter of fact, **we reached 76.8220502372865 in FID score**, showing that our generator is capable of producing pretty realistic photos.

To quantify the diversity of our generator results, we also made use of another metric called LPIPS. We ran LPIPS across all images within our test dataset, with each image A we used 10 different random latent vector z to generate 10 different image Bs. Finally **we compute the average of all LPIPS scores and it produced a result of 0.23297.**

2.2 t-SNE Graph Embedding Space Visualization

What's more, we also visualized the distribution of latent variables of our trained encoder using t-SNE technique. We collected all encoded latent variables from image A from training dataset and ran them through openTSNE. The figure below shows that our trained encoder is capable of producing rather compact uniformly-distributed latent variables for generator.

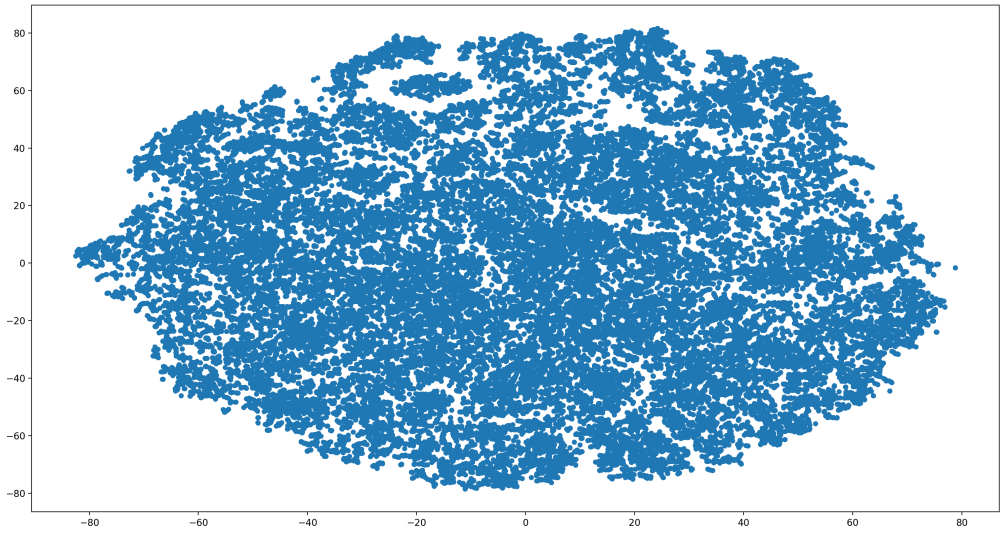


Figure 3: t-SNE visualization of latent variables in graph embedding space.

3 Qualitative Evaluations

In order to further show our trained generator, we also selected multiple image As from dataset and transformed it to different image Bs using 4 different random latent vectors. The results are shown as follows.

Figure 4: Selective inference results of multiple random latent random variables.

From the results we can see that as the random variable changes, our trained generator is able to produce different styles of shoes according to the sketch while keeping the images in a reasonable shape of shoes.