**LT2326 Assignment2 Report**

Tianyi Geng

-Bonus A

I changed conv2d to (3, 32, (3,3), *padding*=1), using smaller 3x3 window, more channels/pattern detectors (32).

I also convert all images to 256\*256 px

Model (wikiart.pth) accuracy after these modifications:0.08253968507051468

-Part 0 Documentation

E.g.

To run train.py with 10 epochs: train.py -e 10

To run autoencoder.py with 10 epochs: python autoencoder.py -e 10

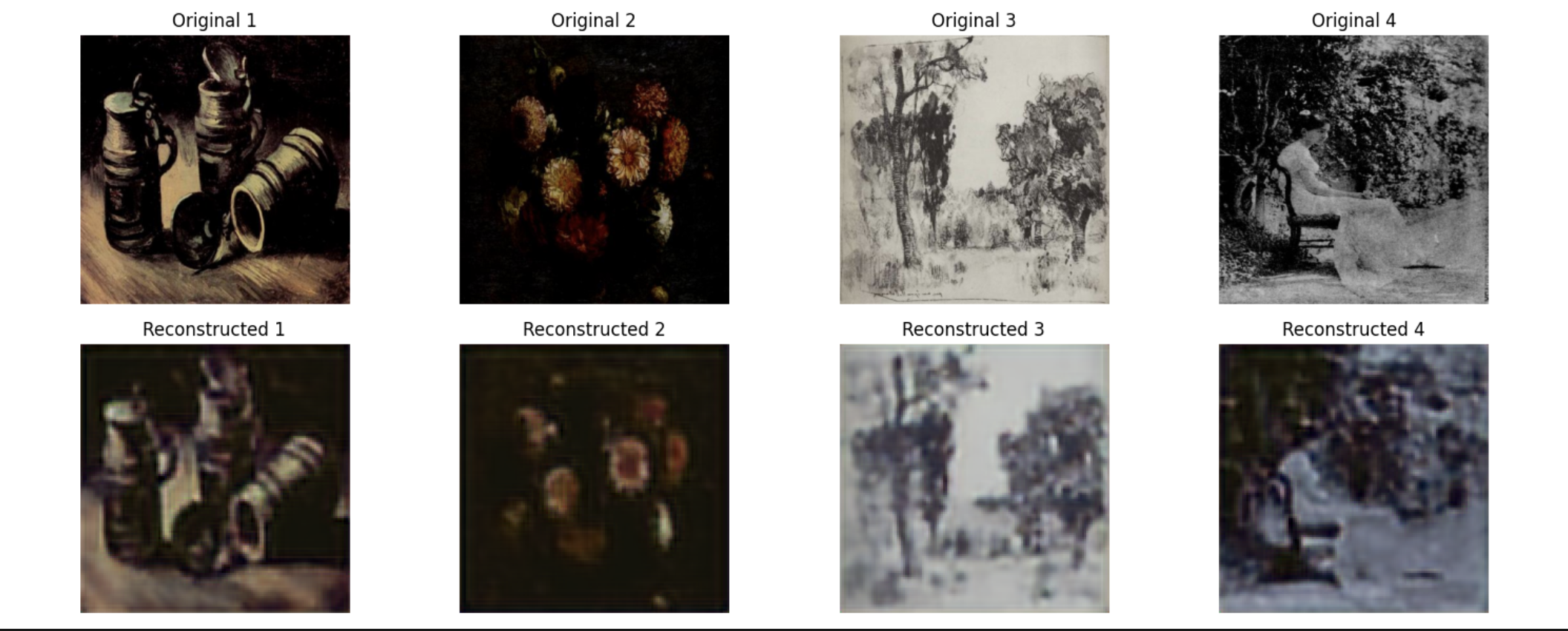
To run augmented\_autoencoder.py with 10 epochs: python augmented\_autoencoder.py -e 10

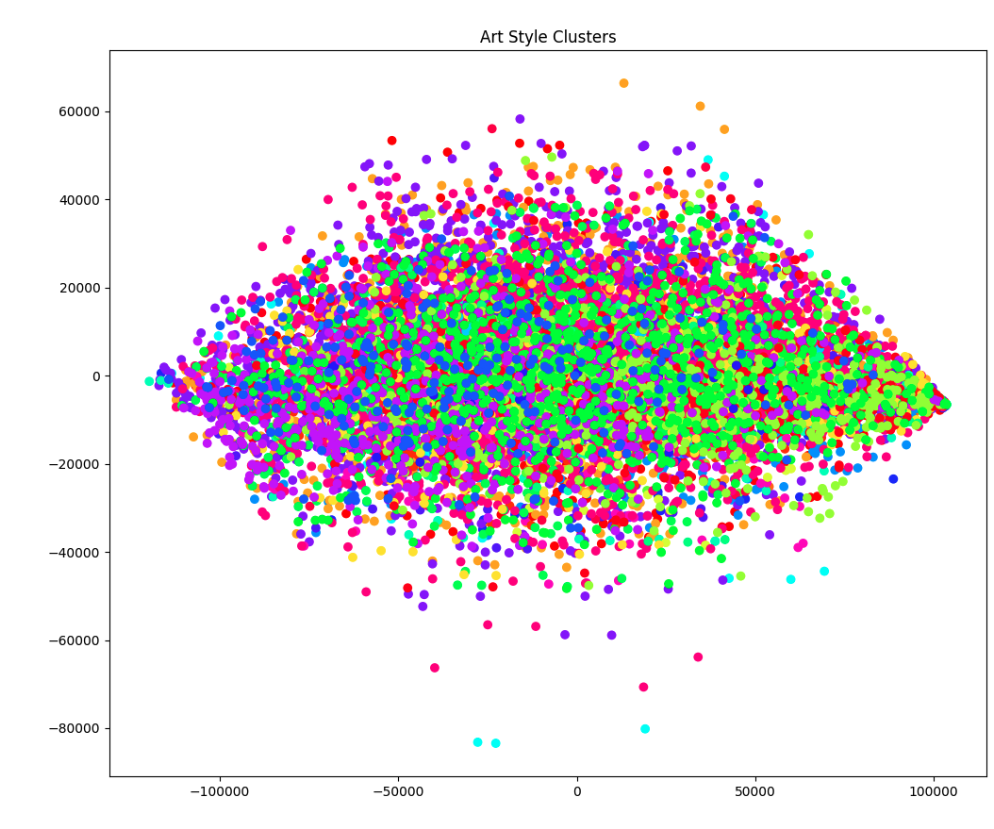
-Part 1 Class Imbalance

I used undersampling, reducing all classes sizes to the smallest class size.

Model (wikiart\_undersample.pth) accuracy after these undersampling: 0.0714285746216774

-Part 2  Autoencode and cluster representations

The clustering is taking form after 3 epochs as the image below shows. I also tried generating the reconstructed images to compare with original images for the final epoch before clustering to see how well the decoder is doing. It doesn’t work very well after 10 epochs since the reconstructed images are blurry and mildly changed in colors compared to original ones. I am not entirely sure how the reconstructed images should look ideally.



-Part 3  Generation/style transfer

For the generated results, I would say the difference is very trivial (but I only ran 3 epochs). I would attribute it to the architecture of the model and the lack of data, which lead to the model not being well trained to show distinctive variations. But it was really interesting understanding the logic of this part. Concatenating art style with original image to get versions of different art styles is definitely very inspiring for me.