ELEC 475 - Lab 2

Steven Calles (20153117)

Sinan Sinan (20165092)

Hardware

For both the "1k" and "10k" datasets, a GPU was utilized to train the model. The specific GPU used was the GTX 3060 with 6GB of VRAM. The GPU performed well enough for this project to train different models and conduct tests.

Hyper-Parameters

Through numerous alterations. the best results for hyper parameters are as follows:

- Initial Learning Rate: 0.0001
- Learning Rate Decay: 0.00001
- Optimizer: Adam
- Scheduler: ExponentialLR
- Batch Size: 8

These hyperparameters were selected because they provided the best results with the least amount of time required to train a sufficient model.

AdaIN Inspiration

The implementations from AdaIN that helped us the most included their implementation of loss and train_transform function. The loss equation represented the addition of the content loss and style loss, and their implementation provided a framework to calculate the content loss using the network weights. The train transform function provided a simple method to resize the images and convert them into tensors.

Time to Train

For the three datasets, it took approximately 2.5 minutes to train the 100 image set, ~30 minutes for the 1k set, and ~5 hours for the 10k set.

Gamma Optimization

For training, we found it best to use a high value of gamma (0.9) to be precise. We chose this value after training with multiple gamma values on the 100 and some on the 1k dataset. After incrementing from 0.1 to 1, by 0.1 each time, we slowly tuned the parameter and compared the loss curves and the output images for each value of gamma, and narrowed down some higher options for the 1k dataset. We found that 0.9 gave us the best loss curves and output images, so we opted to use that one for further training.

Loss Curves Comparisons

The 10k loss curves are noticeably smoother than the 1k, and appear more like an exponential curve than the 1k loss curves. This is most likely influenced by the difference in the size of the

dataset. With more training data, the model has a much more diverse range of samples to train from, which helps the model better generalize, allowing for it to converge more elegantly and quickly to an optimal solution, thus the smoother loss curves.

Qualitative Differences

The difference between the 1K and 10K images were how well the 10k images were able to retain their content, while taking in more of the style of the style images used in testing. This again is due to the larger dataset, as the model was far more generalized with 9000 more images to train on, allowing it to better understand how to apply an image style without losing the contents of a target content image. The subjects of each photo were still clearly visible, with different styles in place of the old ones.