CSC 871 - Deep Learning

Team Project Proposal - Style Transfer

Mark Kim, Satvik Verma, Fabian Weiland

May 9, 2024

Instructor:

Prof. Robert Mateescu

1 Description

Advancements in machine learning have brought generative AI into the mainstream. Unfortunately, trying to reproduce these kinds of advanced models with limited compute resources is infeasible. Nevertheless, there are ways to utilize models that may not match the state-of-the-art and still produce amazing results while learning about the foundations such models are built upon. In our project, we attempt to implement neural style transfer using pre-trained models and multiple approaches to the style loss function. At the very least, we will examine the gram-matrix method, but aspire to experiment and compare this method with two other methods: optimal transport and Vincent loss. Finally, as a reach goal, we may attempt mixed style transfer using multiple style images to produce the generated image. With this reach goal, it will be interesting to see if we can produce more generalized results that mimic an artist's style rather than just a single painting. Likewise, it will be fascinating to see the results from mixing multiple varying styles into a generated combination image.

2 Dataset

The style dataset that we will be working with is a collection of paintings from famous artists ranging from Monet to Pollack which we source from Kaggle. This collection consists of a total of 8446 paintings from 50 different artists. Because we will be producing a generalized model, we will be experimenting with multiple artworks, which is why we need a variety of style sources. As for our base or content images, we will be using various personal, public domain, and fully licensed sources.

3 Approach & Task Assignment

We will be using the PyTorch library for our project and approach this problem incrementally and iteratively. Our primary goal is to implement a simple style transfer using the gram-matrix method

with a single style and content image. Our secondary goal is to expand on this with the optimal transport and Vincent loss methods and then analyzing the differences and similarities between the three. Our final and aspirational goal will be to implement mixed style transfer.

As our team has varied expertise, we will split the responsibilities accordingly. Satvik Verma has experience in computer vision, CNNs, and image processing, so he will take lead with researching and implementing the image pre-processing tasks. Since Mark Kim has the most experience and knowledge of Mathematics, he will be responsible for researching, advising, explaining, and writing about the mathematical underpinnings of the entire process. Fabian Weiland has broad experience in all stages of software design, so he will take lead for software design and implementation. The final report will be a group effort, with each member contributing according to their respective expertise and to the editing of the entire document.

4 Discussion

Since we are building our project incrementally with the intention of creating a model that may have the flexibility to use different style loss methods and mixed style transfer, we will need to employ good software design practices with extensibility and flexibility in mind. By doing this, we should be able to implement our primary goal and then build upon it with very little to no refactoring as our project progresses. Similarly, if time does not permit us to expand the project, we will have the option to abandon the more ambitious portions of the project without affecting the rest of the project.

This project is generative with no quantitative methods of evaluation, so our assessment of the project will rely mostly on qualitative analysis and subjective measures of performance. Currently, our ideas for quantitative evaluation will likely be based on how quickly each method converges on a suitable result with respect to the type of task provided to it. As we progress, we suspect that we will learn other ways to evaluate the project.