Optimization-based Neural Style Transfer

Tushar Verma

12 Oct 2024

1 Introduction

This project focuses on neural style transfer, where the style of one image is transferred onto the content of another image using optimization techniques. Instead of training a network end-to-end, we directly optimize the pixels of a target image by minimizing a style and content loss. The goal is to generate visually coherent images that preserve the content structure while adopting the artistic style.

2 Methodology

2.1 Dataset

The content and style images were sourced from the Art Images dataset, including categories such as drawings and iconography. Images were preprocessed by resizing to 256 pixels on the smaller side, followed by normalization using ImageNet statistics.

2.2 Model Architecture

A pretrained VGG-16 network (features only) was used for feature extraction. Key details:

- Content features were extracted from `conv4_1`.
- Style features were extracted from `conv1_1`, `conv2_1`, `conv3_1`, and `conv4_1`.
- Gram matrices were computed for style representation.

The VGG-16 model's parameters were frozen to prevent updates during optimization.

2.3 Training Strategy

The target image was initialized as random noise matching the content image size. Optimization was performed on the pixel values of the target image using the following losses:

- **Content Loss**: Mean squared error between the target and content features.
- **Style Loss**: Mean squared error between the Gram matrices of the target and style features.

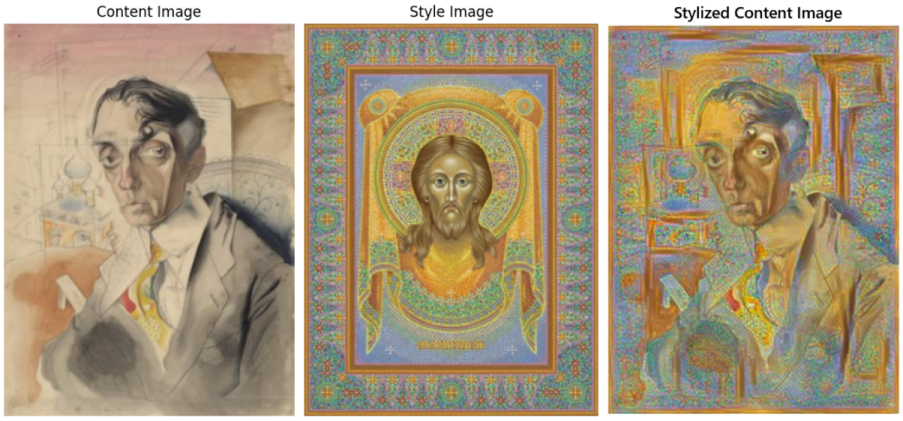
The total loss was a weighted sum:

$$\text{Total Loss} = 1.0 \times \text{Content Loss} + 3 \times 10^{-6} \times \text{Style Loss}$$

Optimization was performed using L-BFGS for 150 iterations. Frames were captured every few steps to visualize the gradual transformation into a motion picture animation.

3 Results

The optimization progressively stylized the content image, successfully blending structural features from the content with textures and colors from the style image. Visual inspection revealed smooth style application with reasonable preservation of the original content. Minor artifacts appeared in highly textured regions, but overall transfer quality remained visually coherent.



4 Conclusion

This project demonstrated effective neural style transfer through optimization without training a deep network. Using a frozen VGG-16 model for feature extraction and Gram matrix matching, the method achieved high-quality stylizations. Future improvements could involve multi-scale style representations or better regularization to enhance fine details further.