Image Styling and Transformation

Amarpreet Kaur, 501213603

March 13, 2024

Abstract

Visualizing image styling and transformations can be costly and time-consuming. Applying different styles before starting the remodeling work helps align with the vision. My method will be utilizing autoencoder architecture with block training, high-frequency residual skip connections, and bottleneck feature aggregation, achieving photorealistic style transfer for images.

Background

In 2015, Gatys et al. applied Deep Neural Networks to the realm of fine art [1]. They took natural images and stylized them with famous artworks by extracting the content representation and style representation of each image. In 2017, a method employing autoencoders was presented for the task [2]. I will expand these works to style transferring of images in general by using an autoencoder approach.

Research Questions

- How does using an autoencoder with specific techniques contribute to realistic style transfer in images?
- What role do hyperparameters play in optimizing style transfer for diverse applications such as virtual try-ons or automotive customization?
- How effective are the techniques adopted in refining the autoencoder model and enhancing the quality of style-transferred images?
- What limitations do we encounter, especially regarding scalability and adapting to different styles and images?

Methodology

I will be using an autoencoder architecture for photorealistic style transfer. The method will combine bottleneck feature aggregation (BFA) with block training and high-frequency residual skip connections to improve detail preservation and photorealism. The autoencoder will utilize a VGG-19 network as the feature extractor and performs style transfer by aligning the content and style image features through Whitening and Coloring Transformations (WCT). This process aims to maintain the global structure of the content while infusing the style from the style image, leading to more realistic and detailed stylized images compared to existing methods.

Dataset and Evaluation

I will use the MSCOCO [3] dataset and the ADE20K [4] dataset. I will assess my style transfer model using two key metrics.: Image Reconstruction Loss, which measures the accuracy of the reconstructed image compared to the original, and Feature Reconstruction Loss, which assesses the preservation of high-level features from the original content and style images in the stylized output.

References

[1] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. *A neural algorithm of artistic style.* 2015. https://arxiv.org/abs/1508.06576

[2] Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin lu, and Ming-Hsuan Yang. *Universal style transfer via feature transformations.* 2017. https://arxiv.org/abs/1705.08086

[3] MSCOCO Dataset. https://cocodataset.org/download

[4] ADE20K Dataset. https://groups.csail.mit.edu/vision/datasets/ADE20K/