CS 445 Final Project Proposal

Theo Xiong TJ Caron Erie M. Adames Sagnik Chakraborty

I. Motivation

We aim to extend traditional CNN-based style transfer (Gatys et al., CVPR 2016)[1] by introducing depth awareness, so that artistic stylization varies naturally across spatial layers (foreground, midground, background). Conventional style transfer produces uniformly stylized outputs that flatten depth perception. By integrating depth estimation (MiDaS v3) and adaptive blending, we will create 3D-consistent stylized photographs that preserve perceptual depth cues. Through this project, we hope to deepen our understanding of the intersection between neural based image generation, geometry, and computational photography aesthetics.

II. Milestones

	Milestone
1	Prepare dataset and build baseline neural style transfer model using pre-trained VGG19 following the Gatys et al. (CVPR 2016) method.
2	Compute per-pixel depth maps to separate image regions into foreground, midground, and background layers.
3	Build pipeline to apply style transfer independently to each depth layer, allowing stylization strength to vary by distance.
4	Merge the stylized layers into one seamless image using blending techniques to avoid visible boundaries.
5	Evaluation and Report

[1] https://openaccess.thecvf.com/content_cvpr_2016/html/Gatys_Image_Style_Transfer_CVPR_2016_paper.html

III. Evaluation

We will assess results both quantitatively and qualitatively:

- Metrics: Structural Similarity Index (SSIM), Learned Perceptual Image Patch Similarity (LPIPS)[2], and style-content loss ratios across depth layers.
- Visual evaluation: Human perceptual study comparing flat vs. depth-aware stylization on realism and aesthetic preference.

A successful outcome will produce images where stylization strength correlates smoothly with perceived depth and shows fewer halo artifacts at layer boundaries.

IV. Resources

- torch, torchvision, numpy, opencv, matplotlib, etc.
- MiDaS (https://github.com/isl-org/MiDaS) for depth estimation
- LPIPS (https://github.com/richzhang/PerceptualSimilarity) for evaluation
- Pretrained VGG19 style model (from Gatys et al.)
- Dataset: 100–200 high-resolution images

V. Group

Member	Responsibility
Erie	Depth estimation pipeline
Theo	Neural style transfer implementation
Sagnik	Evaluation & perceptual metrics
TJ	UI/interactive demo, documentation

[2] https://github.com/richzhang/PerceptualSimilarity