Photorealistic Neural Style Transfer

Generating Realistic Images **without** GANs

Richard R. Yang

Department of Computer Science, Stanford University

Introduction

- **Neural Style Transfer** allows users to generate new artwork from a *style* and *content* image. However, this algorithm fails to preserve details when photographs are used as input images.

- State-of-the-art neural Photo Style Transfer (**PST**) use various forms of the **Laplacian** matrix to preserve details, which is expensive to compute.
- Generative Adversarial Networks (**GANs**) are state-of-the-art in realistic image generation, but are difficult to train.
- We introduce a **novel** approach to photo style transfer **without using Laplacians or GANs,** and achieve realistic results.

Methodology

- PST as an optimization problem:
  - Randomly initialize an initial image
  - Update pixels to minimize a loss function and repeat until convergence
- Our algorithm contains two stages: **global** style transfer and **regional** style transfer, and applies guided image filter after each stage
- The input is a style image, content image, style mask, content mask.
- For evaluation, we use Adobe's dataset of 60 pairs of diverse real world photographs with semantic masks. This dataset was used to evaluate the current state-of-the-art PST.

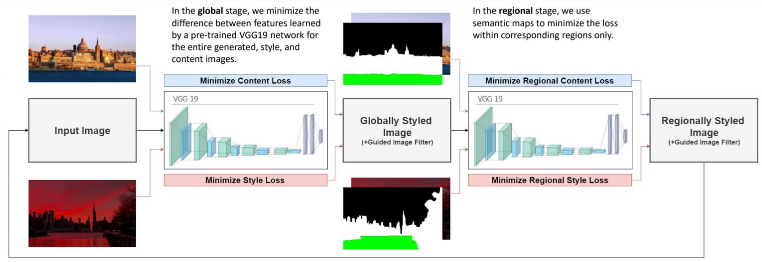
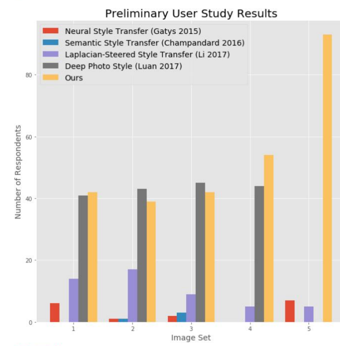


Figure 1. Our algorithm pipeline

Qualitative Results

Figure 2. A qualitative comparison between the content image, other state-of-the-art photo style transfer algorithms, and ours. Insets shown to magnify details of the images.

Quantitative Results



- Preliminary user study to evaluate how realistic our generated images are compared to previous style transfer work and current state-of-the-art
- **105** users were shown **5** sets of images, each set contains **5** randomly-ordered images generated by techniques from related work.
- Users were asked to *"Select one image that looks the most realistic, as if it was captured in the real world by a camera."* Blank answers were allowed.

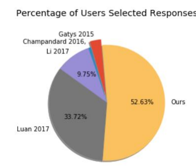


Figure 3. **User study responses. The images generated by our algorithm is comparable to the current state-of-the-art (Luan 2017), and performs significantly better on certain images. Image Set 5 is the mountain image in Figure 2 (second row).**

- Compute the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) score for 60 **content** images and **generated** images from all algorithms
- Compute the mean distance between all corresponding content scores and generated scores

Distance Metric	Gatys 2015	Champandard 2016	Li 2017	Luan 2017	Ours
MAE (L1)	19.762	16.426	14.078	6.429	**3.011**
MSE (L2)	498.062	512.008	249.795	64.080	**23.616**

Table 1. **Mean distance between BRISQUE scores of the content image and generated images. Our images are the closest to content images, which are real world photographs**