CS 231n Project
{Tianchang He}@ee.stanford
{Fabian Chan, Xueyuan Mei}@cs.stanford

Fast Mixed Style Transfer

Demo of mixed style transfer at http://[ip_address]:3000/ ip_address is _____

Background

Among all the applications of Convolutional Neural Network (CNN) and visual recognition, style transfer has been a very heated topic. Style transfer is the technique of separating and recombining the content and the style of an arbitrary image.

Problem and Dataset

The style and content of an image can be disentangled and applied independently, and the fast style transfer algorithm by Johnson [2] was able to speed up style transfer by training a feed-forward image transformation network to replace the optimization based method of Gatys [1], and ended up many times faster, allowing the transformation of video input in real-time. We would also like to experiment on training a network that can generate multiple styles and arbitrary styles.
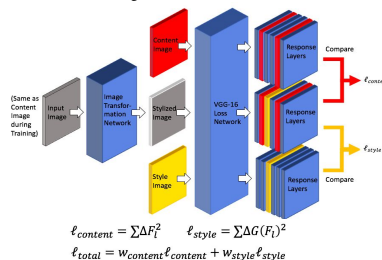
Our dataset includes 13 GB's worth of images from Microsoft's COCO 2014 dataset and artworks of various artistic styles collected from the web, many of which come from WikiArt.org.

Related Work

Gatys [1] proposed the neural style algorithm in 2015, Johnson et al [2] improved its implementation by adding a feed-forward network. Keiji [3] and Ghiasi et al [4] proposed methods for taking style choice or style image as inputs to the feed-forward network, thus generalizing the algorithm to unseen styles.
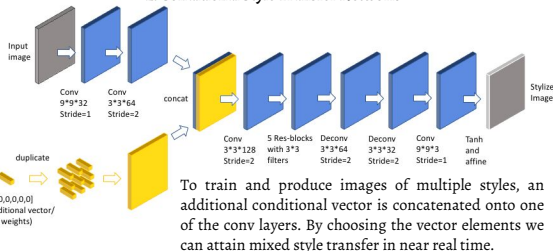
[1] Gatys et al, A neural algorithm of artistic style, arXiv:1508.06576 (2015)
[2] Johnson et al, Perceptual losses for real-time style transfer and super-resolution, arXiv:1603.08155 (2016)
[3] Yanai, Keiji, Unseen style transfer based on a conditional fast style transfer network (2017)
[4] Ghiasi et al, Exploring the structure of a real-time, arbitrary neural artistic stylization network, arXiv:1705.06830 (2017)

Models/Algorithms

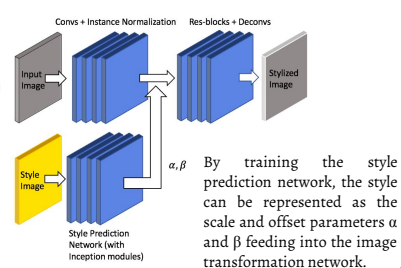**1. Image Transformation Network**

$$\ell_{content} = \sum \Delta F_i^2 \qquad \ell_{style} = \sum \Delta G(F_i)^2$$
$$\ell_{total} = w_{content}\ell_{content} + w_{style}\ell_{style}$$

2. Conditional Style Transfer Network

To train and produce images of multiple styles, an additional conditional vector is concatenated onto one of the conv layers. By choosing the vector elements we can attain mixed style transfer in near real time.

3. Arbitrary Style Transfer Network

By training the style prediction network, the style can be represented as the scale and offset parameters α and β feeding into the image transformation network.

Results & Analysis

Figure 1. Single Style Transfer (runtime < 1s)

Shrimp by Baishi Qi

Figure 2. Mixed Style Transfer

Figure 3. Arbitrary Style Transfer. Runtime ~ 1.5s. The network overfits to only the 5 styles seen at training and does not generalize well. Our implementation was limited by GPU memory and number of styles.
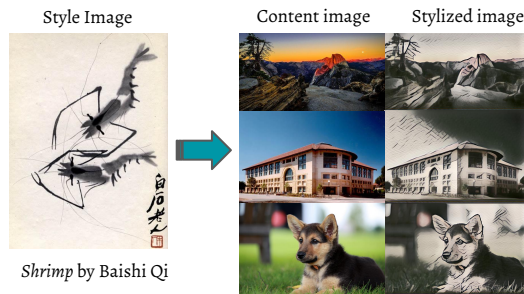
Due to the nature of the problem we evaluate our results based on visual appeal and run time. Figure 1 shows the style of a Chinese painting applied onto 3 different images. One can easily observe the broad brushstrokes and the modest palette that is characteristic of the artist's style. Figure 2 shows at every corner a distinct artistic style applied to the same image; and those shown between corners are a mixture of the styles from the nearest corners. The middle image is stylized from an even combination of all four styles.