

Neural Multi-style Transfer with Attention Masking

Richard Li, Yizhou Liu, Zhaoyi Zhang

Introduction

Style transfer is currently an intensive research area in computer vision, in which a digital image is manipulated to adopt visual style from another image. In the classic style transfer setting, such transfer is performed uniformly, namely, every part of the source content image is equally modified to adopt a specified style. To the best of our knowledge, the current state of the art algorithm proposed by Dumoulin, V., Shlens, J., & Kudlur, M. (2017) managed to generate images with n different styles with one forward pass, given a content image c and style images s_1, s_2, \dots, s_n . In the work of Schekalev, A., & Kitov, V. (2019), they addressed that the classical style transfer algorithm imposed a uniform transfer on the image, but people may desire to keep the original quality of the main object. They propose a central object detection algorithm to preserve the original feature and quality for the main object while performing style transfer. However, it is common the case that there could be multiple main objects in an image, and instead of preserving their original qualities, people may want to impose a specified style for each object, namely, objects of multiple styles coexist in one output image. Here we take a first step to explore the possibility of achieving this goal.

Importance

It is a trend that more and more people are getting familiar with photo editing tools to modify their pictures. Many softwares on the market enable users to add filters to the pictures they took, and beautify the figures in the picture. Moreover, people also want to be their own content creators and impose styles on their figures to make them look more stylistic and artistic. However, when there are multiple figures in the same image, instead of picking a uniformed style for all figures, people may want to choose a unique style they want to impose on their own figure. Our project takes this desirable trait into account.

Possible Approach

Conditional instance normalization is proposed by Dumoulin et al (2017). It is a method in which the activations in the convolution neural network are normalized, scaled, and shifted to apply a certain learned style in a feedforward convolutional neural network. Essentially, the network is learning the embedding for different styles. One possible approach is to borrow the insight from conditional instance normalization and try to

design a new construction of embedding space with multi-masking technique to achieve our goal hopefully with one forward pass.

Baseline

We plan to present our results and compare them against the work of Dumoulin, V., Shlens, J., & Kudlur, M. (2017) and Schekalev, A., & Kitov, V. (2019) to show the innovative development we made.

Time Table

Dates	Goals
10/15	Re-implement the feed-forward style transfer in <i>Texture Networks: Feed-forward Synthesis of Textures and Stylized Images</i>
10/30	Re-implement conditional instance normalization in <i>A Learned Representation For Artistic Style</i>
11/15	Explore ways to implement Multi-style Transfer with Attention Masking
11/31	Wrap up and prepare presentation