Machine Learning Project Proposal - Neural Style Transfer

Amith Gopal
Akriti Kapur
Amruta Rajopadhye
Sayali Sonawane
Sharan Srivatsa

Abstract

Style Transfer is a technique in which we extract the style of one image and overlay it on the content of another image. This task has been made possible by evolution of convolutional neural networks that create a semantic representation of the image at hand. Artistic style transfer by Gateys et al. [1] is the seminal paper which gave an algorithm to perform style transfer on any pair of images. In 2017, this technique was commercialized in an amazing app called Prisma which generated a lot of buzz. The app transforms any photo that is uploaded into works of art by famous artists. The objective of this project is to study and implement some of the style transfer techniques.

Motivation

Creativity is a core human trait. With time, sophisticated techniques to tackle complicated problems are being developed and instead of just learning from available data, we want our algorithms to be creative. Style transfer is one such application which tries to merge machine intelligence with art.

Why Machine Learning?

This seems daunting at first, which makes us wonder how we would even begin such a task. Traditional algorithms would do a pixel-level image analysis on the style image to extract the colors or some aspects of the texture. Through Machine learning, can we teach our system that the texture and color we are interested in is at the scale of the brush strokes and not the overall shapes that are present in the painting.

Possible Approaches

Currently there are two main approaches to neural style transfer viz.

1. Descriptive Neural Networks based on Image iteration

This approach is used in Gatys et al.[1], In this method, features such as color or texture

are extracted from the first image (called the style image) and structural or spatial features are extracted from the second image (called the content image). Finally, features of the style image are super imposed on the content image and a third style-transferred image is created.

To carry out this transformation, a loss function γ_c which represents the difference between the generated image and the content image is computed along with another loss function γ_s which represents the loss between the style image and the generated image. The objective of this task is to constrain the generated image in a way that minimizes a linear combination of the overall loss. The degree to which we want to apply the transferrence can be tuned with the parameters of this overall loss function.

2. Generative Neural Networks based on Model iteration

This approach is used in Johnson et al. [2], In this method, instead of performing a pair-wise transformation of style and content images, a feed-forward network is trained over several images for each specific style image, then, style is transferred to the content image in a single pass, this is faster than the previous method, but has the additional overhead of training one network per style. Further, style images can be grouped in terms of their overall theme (ex.Van Gogh's artistic style) and this approach can be made less redundant by training a neural network per theme instead of style. Dumoulin et al.[3] have detailed this enhancement.

Dataset

The initial approach for extracting high level image representation for our images is by using Oxford's VGG network[4].
As an enhancement, if the results of the VGG network cn be improved we intend to use the following datasets via trial and error.

- DAQUAR dataset (which is based on the NYUDepth V2 dataset).
- MS-COCO dataset which was developed and orchestrated by Microsoft.

References

[1] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge *A Neural Algorithm of Artistic Style*

[2] J. Johnson, A. Alahi, and L. Fei-Fei. *Perceptual losses for real-time style transfer and super-resolution.* In European Conference on Computer Vision, pages 694711, 2016.

[3] V. Dumoulin, J. Shlens, and M. Kudlur. A learned representation for artistic style. ArXiv e-prints, Oct. 2016.

[4] `http://www.robots.ox.ac.uk/~vgg/research/very_deep/`