Machine Learning Engineer Nanodegree

Capstone Proposal

David Pham

November 26th, 2017

Proposal

Domain Background

In recent years, deep learning showed its capacity to support art creation with neural style transfer. The idea is to copy the style from some image and to apply to the content of some other picture. Introduced by Gatys et al., the method has been improved by Johnson et al.. Instance normalization appears to provide a significant speedup to the process. Jing et al. provides an overview of the domain as well.

My motivation to study this and understand this topic is to combine my technical and artistic abilities together and also to provide some fun content easily to my friends.

Problem Statement

I would like to understand the original papers for neural style transfer and its improvements. The goal would be to create an stable and independent implementation of one of the concept and to easily apply the algorithm to different style images and to apply to any content image.

Datasets and Inputs

Creating a neural networks from scratch would take too many computational resources and the weights of Resnet-50 (from computer vision) are available and also used everywhere to compute a good representation for pictures.

Inputs would jpeg pictures. Style would be typically from famous paints or maybe some standard natural environment (beaches, forest, etc.). We would download pictures from either internet or use some standard dataset (cifar100 or imagenet).

Content picture will either be persons (when applied with an artistic style) or other environment (when applied to other environmental scenes).

Solution Statement

We will implement in tensorflow a part of fast-neural-style repository.

Neural style transfer is a technique that uses deep neural networks to learn representations of pictures and part of theses representation can be used to describe either the style or the content of a picture. The goal of this project is to provide a good summary of Neural Style Transfer, with the fast forward version as well as for instance normalization.

Moreover, we will provide a package that should apply style to a content picture in a reasonable amount of time (under 10 seconds for a 600x600 image.) I hope our model should be at least able to attain 10% of the state of the art implementations.

Benchmark Model

Models are provided in this repository fast-neural-style and in the keras example. An implementation in tensorflow is provided here by Lengstrom.

Evaluation Metrics

My metric will be the speed of my program to apply style to content. I will try to measure its distribution and benchmark my model and should run under 10s for most low resolution pictures.

Project Design

First part will be to understand the underlying concept behind Neural Style Transfer and its improvements. Then it will be to collect the data. Then, a part of the code will devoted to handle data (pictures): resizing them, moving them around and also caching the style to avoid recomputing their representation uselessly. As described above, fast neural style transfer is probably the main algorithm to be implemented with the trick of instance normalization.

A part of the project will be determine how to extract the hidden representation of the pictures and to implement efficiently the combining operations of these extracts. As I do not have GPU on my laptop, AWS instances will have to be set and run jupyter notebooks from there.

Obviously, the project will be implemented in Tensorflow, although a PyTorch implementation also seems really attractive.

Finally, a simple interface should be created so that the user could provide the paths of images which could be used for content and style and these could be used for creating new and fun pictures.

All in all, this is an engineering problem, where the solution uses deep learning to create new artistic content.