Ishmael Obeso
November 7, 2019

Udacity Machine Learning Engineer Capstone Proposal

Proposal

Build a Neural Painter that can create convincingly human-like paintings by using transfer learning followed by adversarial training.

Domain Background

Ever since I watched Ian Goodfellow's interview with Lex Fridman, the idea of GAN's has caught my interest. It's fascinating that a neural network can imitate what is commonly thought of as a uniquely human trait, art. However, I found that most GANs tend to generate images pixel-by-pixel while humans create art through other means such as brushstrokes, stippling, and other techniques that use the constraints of the artists tools to create a certain style. I wanted to work on creating a GAN that closely imitated the way humans create art. To this end, I stumbled across a blog-post called "The Joy of Neural Painting"[1] that introduced the idea of brushstroke GAN's which more closely imitate the way a human generates paintings.

Instead of generating paintings pixel-by-pixel like traditional neural networks, brushstroke GAN's such as the ones used in the paper 'Neural Painters' [2] generate their images by successively painting brushstrokes on a canvas to create a painting-like image. As evidenced in this paper, this technique is able to create more convincing painting-like textures than other models. I believe by leveraging these types of GAN's to more closely imitate human art, the outputs of these models will be more convincing and able to be used in a multitude of ways. For example, leveraging these models to simply create more human-like art will be helpful when it is not feasible to pay an artist for their original work, such as in small video-game studios that want to include some type of artwork in their games. Instead they can simply make use of these GAN's in their games. More generally, being able to generate images that closely resemble human-made paintings will be useful whenever art or fashion is concerned, such as when work-shopping new clothing textures, creating designs for websites, or simply for creating original non-copyrighted work that is aesthetically pleasing without having to pay exorbitant amounts of money for original art.

Problem Statement

With the recent advances in deep learning we are seeing more and more computer-generated images being used in things like advertising, web design, and fashion. However, it is still relatively easy to spot when a computer-generated image is being used, especially when creating painting-like images. These images tend to be created using style transfer and GAN's that build the image pixel-by-pixel, but if we want to create convincing computer-generated paintings we need to start using human-like techniques. In this way, brushstroke GAN's imitate human artists by using brushstrokes on a canvas to create convincing art.

By creating more convincing art, money and time could be saved by using computers instead of human artists to create art in places such as video-games, TV, and movies. This way, creators don't have to worry about paying a human to create art, when a brushstroke GAN can simply do it for them quickly and easily. This will bring us closer to art is not a uniquely human trait, but one that can be taught to and learned by computers.

Datasets and Inputs

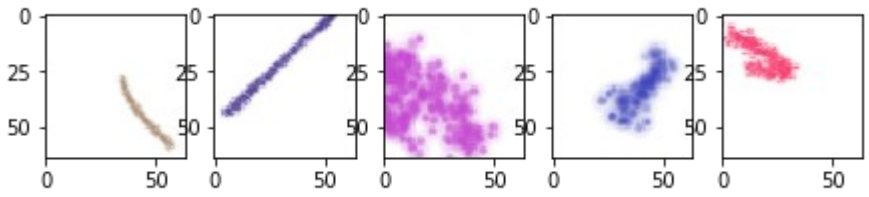
I plan to use the MyPaintBrushstroke dataset provided by researcher Reiichiro Nakano to train the brushstroke GAN to create brushstrokes. This dataset consists of 100k examples of:

actions - these consist of 12-dimensional vectors which characterize the brushstroke. The vectors define start and end pressure, brush size, RGB color values, and brush coordinates.

ex.

array([[0.330522, 0.249296, 0.638134, 0.656268, 0.539904, 0.435634, 0.623549, 0.494268, 0.38073 , 0.896127, 0.924754,
0.466711]])

brushstrokes – images of the brushstrokes that correspond to the action vectors.
ex.



I also plan to use the MNIST and CelebA datasets to generate painting- like images using the trained brushstroke GAN. The MNIST database consists of images of integers from 0-9 and the CelebA dataset consists of images of celebrity faces. The MNIST dataset will be the ‘easy’ test for my GAN and the CelebA dataset will be the ‘hard’ test.

Sources: https://www.kaggle.com/reiinakano/mypaint_brushstrokes
https://www.kaggle.com/oddrationale/mnist-in-csv
https://www.kaggle.com/jessicali9530/celeba-dataset

Solution Statement

The goal of this project is to create convincing painting-like images using brushstrokes. I will use the MNIST and CelebA datasets to create these images, and use VGG loss to measure the difference between the images generated by the benchmark model and my model.

Benchmark Model

I will use the neural painter model used in the ‘Joy of Neural Painting’ Blogpost as a benchmark to test my results against.

Evaluation Metrics

I will use VGG loss as the evaluation metric. VGG loss is a content loss function which uses feature extraction to calculate the Euclidian distance between the feature maps of the two images.

Reference: https://www.oreilly.com/library/view/generative-adversarial-networks/9781789136678/c399f455-2670-4321-92cf-de6cf75cd543.xhtml

Project Design

- Language: Python 3.7
- Libraries: Pytorch, fastai, numpy, PIL, glob
- Goal: Create a brushstroke GAN that can create convincing painting-like images from input images.
- Steps:
  - Non-Adversarial Training
    - Train Generator and Critic non-adverbially to speed up training.
  - Adverserial Training
    - Connect Generator and Critic to create adverserial GAN training.
  - Brushstroke Painting
    - Use trained brushstroke GAN to create painting-like images from input images of digits from MNIST and faces from CelebA datasets.

References:

1. https://medium.com/libreai/the-joy-of-neural-painting-e4319282d51f
2. *Neural Painters: A Learned Differentiable Constraint for Generating Brushstroke Paintings*. Reiichiro Nakano
arXiv preprint arXiv:1904.08410, 2019.
3. *Teaching Agents to Paint Inside Their Own Dreams*. Reiichiro Nakano.
https://reiinakano.com/2019/01/27/world-painters.html , 2019
4. *Fast.ai MOOC Lesson 7: Resnets from scratch; U-net; Generative (adversarial) networks*.
https://course.fast.ai/videos/?lesson=7 ; Notebook: https://nbviewer.jupyter.org/github/fastai/course-v3/blob/master/nbs/dl1/lesson7-superres.ipynb [Accessed on: 2019–08]