McKay Bowcut

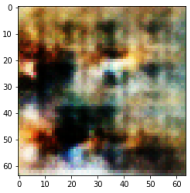
Deep Learning Final Project

So, I had very high, very naive hopes for this project. What started as trying to get a GAN to create interesting works of art, turned into trying to get a GAN to output anything relatively more interesting that random noise when using a small dataset with large variance in the distribution of images. Let's look at how it went.

My goal with this project was to develop a GAN that could create interesting pieces of art. The dataset I found to accomplish this is *The Paintings Dataset* created by Elliot J. Crowley, Ernesto Coto and Andrew Zisserman at Oxford University. The dataset includes 8629 images from the following classes: *aeroplane, bird, boat, chair, cow, diningtable, dog, horse, sheep,* and *train*. My initial GAN architecture is a Generator that Discriminator that operates on 64x64 images.

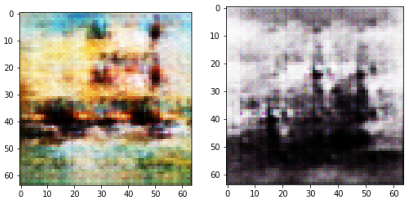
At first, I tried training the GAN on individual class sets, just to see what kind of output I would be getting. Specifically, because there is a large size discrepancy between different classes, I was worried that a GAN would have trouble training for smaller classes. After some experimentation, I decided to try some data augmentation, to increase the size for each set. I implemented a cropping augmentation, I cropped each image to a square at each corner, quadrupling the size of my data.

Below is an example output of training on the augmented data for the bird class:



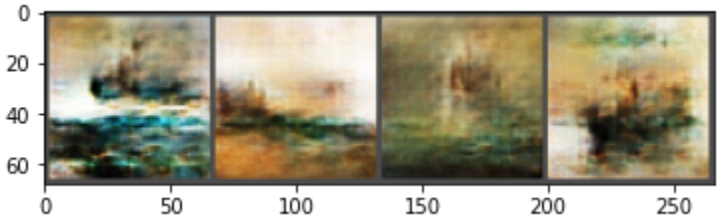
As you can see, there is not much structure going on in the generated output yet.

After more experimentation, I found that my GAN performed better when training on the largest class: boat. I also decided to try training on grayscale data to see if this would help with learning and induce more structure in the output. Here is a comparison of colored vs grayscale training for similar lengths of time:

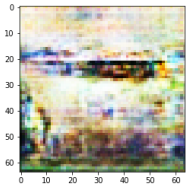


This led me to believe that the benefits of using grayscale data were negligible.

In the end, the greatest improvements were obtained from using the larger class datasets, and training for more epochs. The following are some of the best output for the boat class:



In comparison, here is output for the smallest class, aeroplanes:



It's also important to note that the variance of images within a class probably plays an important role in the quality of the GAN's output. For instance, when we look at the images produced for the boat class, you can see the network has learned structural elements such as water, sky, a horizon line, and a boat. It is likely that most of the boat images have these same elements, allowing for stronger training of the generator. (This is the same reason that GANs are able to achieve such remarkable results when trained on human faces.) It is likely not the case that other classes benefit from this same similarity across images.

Finally, I thought it would be interesting to implement a conditional GAN network on the Oxford painting dataset. I used the same architecture, but upped the image resolution to 128x128 and modified the generator and discriminator so they would take in labels. In this way, the model can learn to make predictions conditional on a given class label.

Early outputs from this model have not been impressive. I suspect that it will require a considerable amount of additional training to produce anything of interest.

LOG

Date	Hours	Activity	Total Hours
11/17	1	Reading	32
11/17	1	Dataset	
11/19	2	Dataset	
11/24	2	research - Numenta models	
12/2	3	dataset processing	
12/3	1	dataset	
12/4	4	Dataset & Initial GAN Model build, train	
12/9	2	Data Augmentation	
12/9	1	augmented data training	
12/10	1	splitting data into class sets.	
12/12	4	class set training, experimentation	
12/13	4	data augmentation experimentation -- grayscale training vs color -- boat class	
12/14	4	tried saving/loading model state, but something is broken --> fixed it! .eval() turns off training	
12/14	2	implemented conditional GAN, started training. Training will probably take too long.	