Andy Lee

28 November 2020

MAIS 202 – Deliverable 3: Progress Report

I. Final Training Results

Many unexpected problems arose after training the GAN model for the first time. Since the project had not been run when Deliverable 2 was submitted, there were many pitfalls which did not reveal itself during beta-testing.

Primarily, further adjustments to the dataset had to be made through preprocessing. First, the images had to be reduced from 128x128 to 80x80 pixels. The main reason for this was because loading in and computing through 128x128 pixels per image made it extremely demanding on the computer's RAM and would often, if not always crash the computer.

In addition, the dataset size had to be reduced to 10000 images from over 180,000 images. This was solely due to the fact that the computations for 1 epoch took too much time to be computable on a regular desktop computer. Therefore, running multiple epochs would be impossible.

Lastly, training the GAN was an incredibly difficult process. The GAN would often collapse and converge due to unknown reasons generating meaningless images.

The discriminator was created using a relatively standard architecture, consisting of 5 x Conv2D and LeakyReLU layers. This architecture, upon research, was standard for elementary CNNs and discriminators. Adam was the optimizer of choice here as it is used with CNNs very often.
The generator was created in the "reverse" process by creating a linking of 4 x Conv2DTranspose and LeakyReLU layers connected to a Conv2D layer at the end.
Lastly the GAN was created by linking the generator and discriminator.

Here are the resulting "fake" images generated from the GAN (Epoch 10, 50, 100):

Visually, there is a clear improvement in the realistic-ness of the photos in increasing epochs. At epoch 10, only skin color, a slight outline of the face, and the positions of the eyes are noticeable. At epoch 50, a more pronounced outline of the face is visible along with other facial features such as the nose and mouth. However, at epoch 100, it appears that all of the facial features are recognized and understood by the model. Some photos in the Epoch-100 set are very realistic.

Statistically, for the models, 'binary-crossentropy' was used as the loss function. This was because the discriminator was built to discern between binary values ('real image' or 'fake image'). Near the beginning of training, the discriminator loss values ranged around **5-20** on average. However, as the training neared completion, the loss was greatly reduced to around **0.1-0.5** on average. The loss values for the GAN model remained higher around **2-4** on average, although it did not seem to have any significant impacts on the output.

Another interesting phenomenon was regarding the accuracy of the discriminator as the training neared completion. In the beginning, the discriminator was able to tell real images from fake images with near-100% accuracy. However, as the training kept going and the generator was able to create more convincing images, the accuracy slowly began to decrease to around **80%** after 100 epochs.

II. Final Demonstration Proposal

The proposal for the Final Demonstration is in the format of either a Poster / Slideshow or a simple Webapp.

In the case of the poster / slideshow, it will contain a detailed explanation of the project, how it works, discussion about the difficulties of creating / training GANs and most importantly, a progression of the generated images from the beginning of the training process to the end.

In the case of the Webapp, a simple landing page will be made in basic HTML and CSS using Flask. The landing page will have a 'Generate' button which will generate a fake image using the pretrained weights from the model. Then it will display the 80x80 image on to the screen to the user.