Vignesh Vasan

8/2/19

CS 4375.0U2

**Music GAN Report**

**Introduction**

           In my project, I am addressing the need for exploring new forms of music. Due to the rising popularity in general adversarial networks, I decided to use this form of network to produce new music, specifically as DCGAN. To create new music, I have used a .wav file and converted it into a .png to put in my network. I used two networks one is the generator and the other is the discriminator. I used deconvolution network as my generator and convolutional network as my discriminator.

           Regarding the overall results, I could not develop any good results. Some of the limitations included my laptop’s processing power ( the network required much more power than I excepted). The first time I tried to expand my network to handle larger parts of the .wav file, I lost my all epoch’s and code. After reconstructing for the last 24 hours, I was able to achieve similar results in creating 1-sec .wav files.

**Problem Definition and Algorithm**

I am trying to address creating a “new” music-based of current music. Due to the limited time and computing power, I was only able to use one song and within that only 30 small partitions of it.

           To begin with, I first convert the .wav to .png using an algorithm I found online. I will try to fix this because I believe that some of the information coming out the network is being lost/overwritten due to the algorithm. Then I use a method called crop by taking out 56x56x3 images (the third dimension is color). Using this method, I crop 30 smaller images and use that as my input. I would love to have used way, but the processing power gets to be too much. I had an alternate method of splitting the whole song, but I had some trouble using it in my network. The next step is to convert the image array into a numpy array and reshape it to fit my network. I start training the network with this data and set the epochs to 20 just in enough so the computer doesn’t burn out.

           In the training part of code, I use a loop for the epochs and the batches. Then I train the data to use the generator and discriminator. I also have multiple methods that save iterations through the epochs and save the images. Finally, I use the last epoch.png to create a .wav file at the end.

**Experimental Evaluation**

Regarding parameters and things, I had changed to see differences in my out, I changed the input number of music and mainly the size of the neural net. I couldn’t hear the difference in the songs output, but the epoch pictures did show change in how the neural net worked. At first, the epoch.png files were purple and gradually changed to resemble the look of the input png. I also decided to flatten a layer in the discriminator to see changes and the epochs start to diverge. I believe it may be due to the rgb ratio and it is impacting how the png is being converted into audio. One of the later changes that crashed my computer was taking in the whole song in with splitting it and somehow caused my .pynb to vanish. Overall, I did see some changes in how the pngs were being handled but not much on how the audio was working.

**Related Works**

Through some of the research I did, I came across a paper on Music Composition from Santa Clara. In this paper, they used midi files instead and worked with an RNNs instead. I also looked at the tutorial by TensorFlow for GAN with the new TensorFlow 2.0 beta. I also looked into Magenta and Wave net, but those algorithms were too complex for to both understand and run. I also found an article on Medium that uses an RNN-GAN to create piano audio files.

**Future Works**

           Currently, I have hardcoded paths to where I would receive, and store files and I would like to change that so that I would be able to run the code smoothly on my pc. I would eventually love to move this entire project to run on my pc. As of now, I having trouble running the new updated version of TensorFlow on my pc due I have to update and override a lot. I would also try to use GPU on pc to run this instead because my laptop had the hardest time just to produce one sec.

**Conclusion**

           Overall, I was able to work with creating a GAN and using it to create some music, rather a small snippet of one. This project was able to use a Bach song to create a second of using a DCGAN with deconvoluted and convoluted neural net.

**Link to Code**

<https://github.com/ViggyV/AI-Music-Generator>

**Bibliography**

“Music Generation with Neural Networks - GAN of the Week.” *Medium*, Cindicator, 27 Aug. 2018, medium.com/cindicator/music-generation-with-neural-networks-gan-of-the-week-b66d01e28200.

*MuseGAN*, salu133445.github.io/musegan/.

Wang, Shao-Fan, et al. *Music Composition with Deep Learning*. 2018.