ART AND MACHINE LEARNING
CMU 2019 SPRING
Final Project Report

Daily Rhythm

Group 8
Wooshik Kim
Yejin Lee

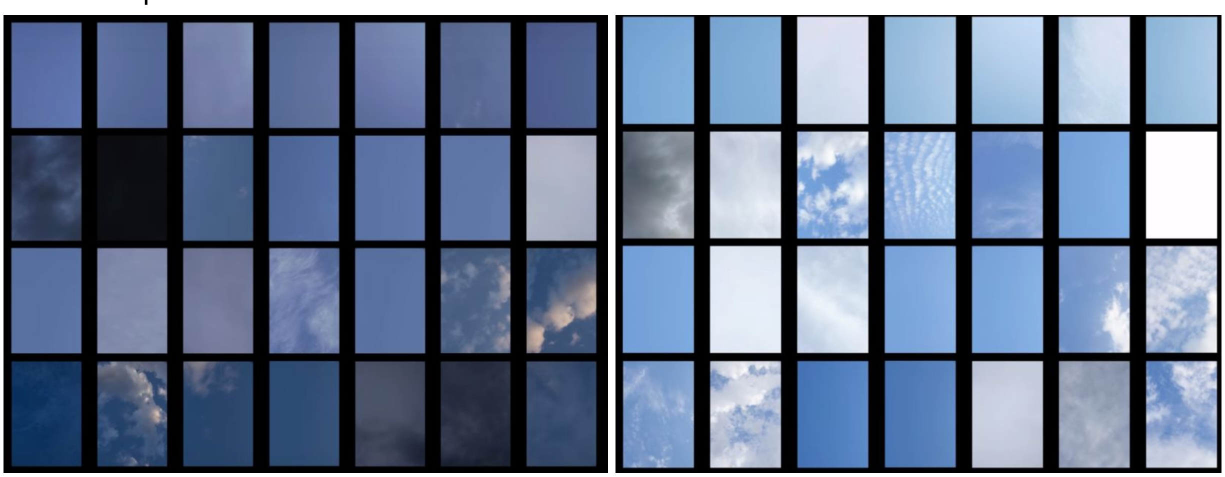
DESCRIPTION

Concept:
Daily Rhythm replicates our daily rhythm that includes natural rhythm and artificial rhythm. The natural rhythm is based on the uncontrollable time like the repetition of days and nights. The artificial rhythm is what we experience through social media, a phone, a computer, and a television disregarding the limitation of time and space.

Process:
We started our final project based on the idea that we wanted to explore from the third project. In this work, we wanted to create a piece of art/music from something that is not. Can everyday conversation be translated into a piece of art? What can we infer from these convesation data (maybe audio track from *Friends*) by feeding it into a machine. Would we hear a distinct flow?

It was a interesting idea, but was technically too challenging. Gathering useful data from voices is actually something researchers are still working on. The problem is our voice has too much dimension in it. It is hard to extract a single note, it is comprised of many notes, and it is hard to learn pattern from it (at least at our level).

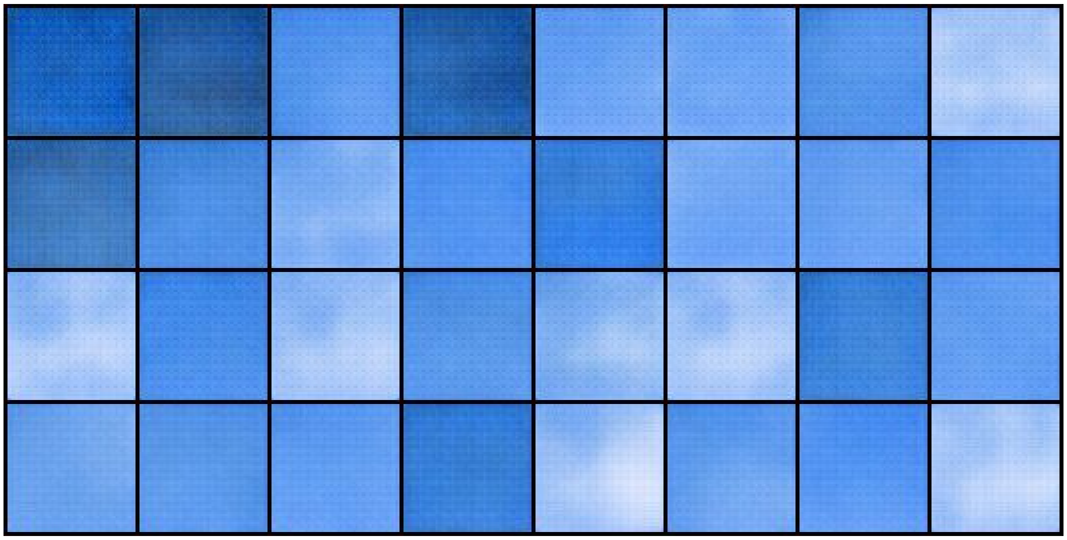
Next we decided to start our final project based on Yejin's previous work. The project was created using photography of the sky taken from the same location every five minutes for four weeks. The resulting arrangement of 28 videos over 28 days offers a visual rhythm analogous to music. [1] There is a clear flow that we can observe here from nature like music. We wanted to dive deeper into this idea.



Daily Rhythm. 2017 **[1]**

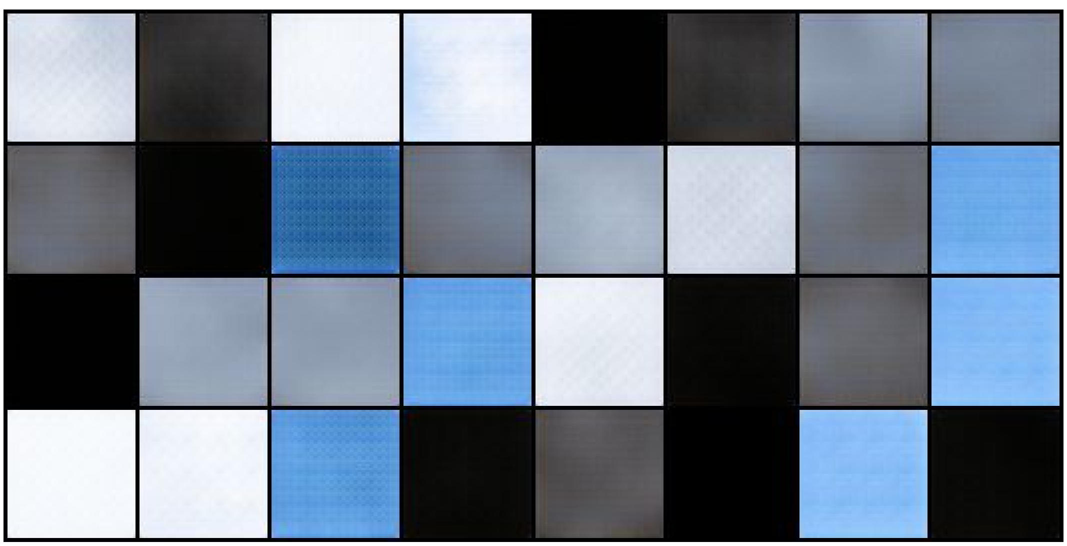
In nature, there is always a clear flow, a causal relationship, and the dependence of time. There is an underlying dynamics that all things in nature follow. What about artificial intelligence? We tried to recreate something similar by using Deep Convolution Generative Adversarial Network.

For this project, Yejin collected data of sky to elaborate the natural rhythm. In one spot, she takes photo of sky every five minutes every once a week for a year. To describe the artificial rhythm, we generated new images with the collected data of sky photos using artificial intelligence.



Generated sky from clear sky data

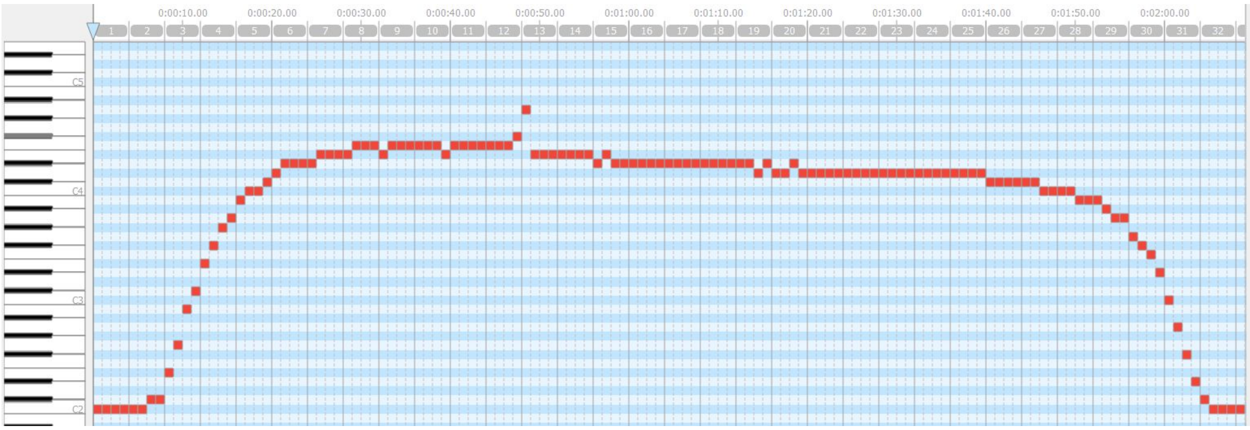
First, we divided the data into four different sets: morning, afternoon, night, and cloudy. However the categories were rather arbitrary and didn't have justification for not using the entire data set. Next we used the entire data set to create a sky spanning all time.



Generated sky from all sky data

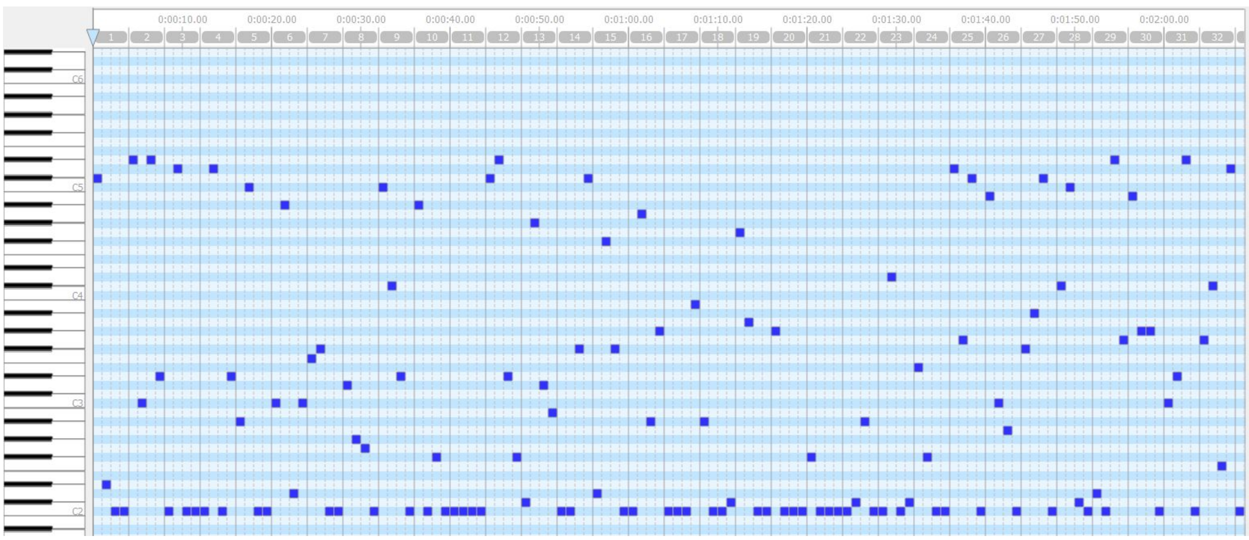
Next, we generated sky using all the data samples. We had ~1000 high resolution pictures. It was interesting to see how various noise (z value) created different types of sky.

The next step was to create another interpretation of the piece. We wanted to express the natural flow found in nature and the unnatural one found from the generated piece. We decided to experiment with another sense: auditory. We extracted average value of pixels (red, green, and blue), and then extracted the average value of those to express each picture with a single number. We translated this number into value between 36 and 81. In midi notation, the notes between 36 and 81 are the most common notes found in music.



Midi file generated from average pixel value of raw data

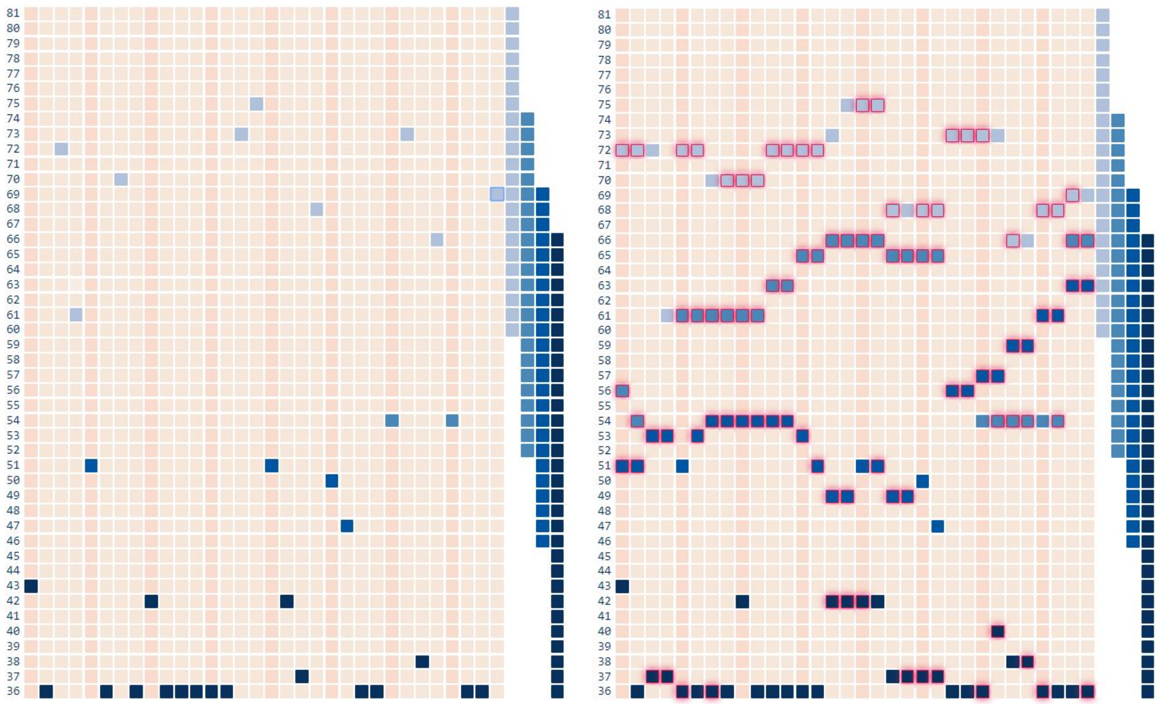
We also gathered the same data from generated sky images. Here it was harder to find the flow found in nature.



Midi file generated from average pixel value of generated images

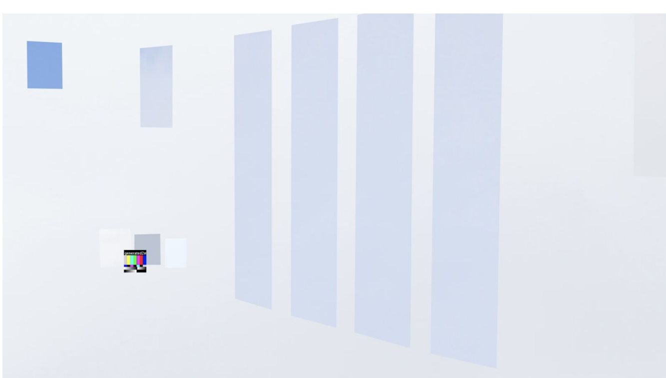
We decided to get help from artifical intelligence to create the flow that is not found in generated image. Although we could create sequence of generated image that has more flow, we wanted to get an “opinion” from a machine to make flow for us based on this input.

We used *Coco Net* from Magenta to create a flow and harmony at the same time. [4] The generator seemed to overcome the lack of flow by adding the missing part and making it more a musical piece.

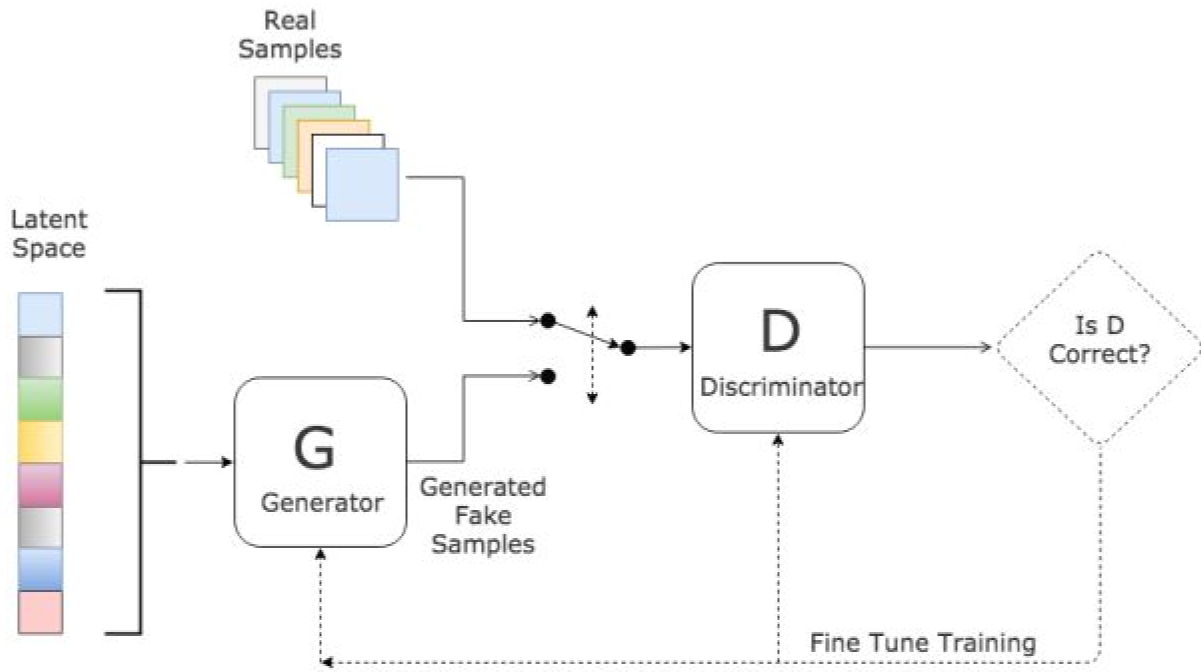


Harmony generated by the coucou net from raw monophonic midi

Using these data, we made four series of videos using Premiere and After Effect. To compose the frame, we referenced the spaces that have multiple frames.



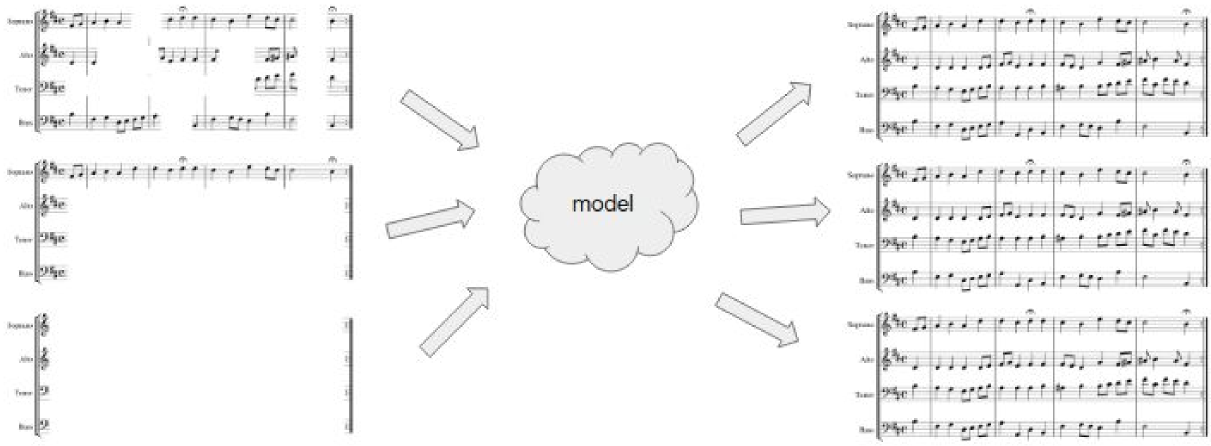
Technique:
First artificial intelligence we used was Deep Convolutional Generative Adversarial Network. Based on the idea of Ian Goodfellow's famous work, we were able to generate images that were similar to the data set we had. [2]



Generative Adversarial Network by Ian Goodfellow

This framework for estimating generative models via adverial process simultaneously trains two models: a generative model G that captures the data distribution (image) and a discriminative model D that estimates the probability that a sample came from the real data than G [2]. By making two models compete with each other, we are able to get images that are actually representative of the actual data.

Second technique we used was a state of the art network called Coconet by Google AI [3]. The network uses 306 chorale harmonization by Bach to generate polyphonic music in the style of Bach.



Coconet generates polyphonic music in the style of Bach

The model was trained to find the missing material from musical scores. During training, an example from the Bach chorales dataset of four-part counterpoint is randomly erased some notes, and the model is asked to reconstruct the erased notes. The music for each voice is represented in a two-dimensional array (one in time horizontally and one in pitch vertically). The stack is treated as convolutional featuremap just like images! The model itself is fairly straightforward convolutional neural network with batch normalization and residual connections [3].

Result:

To elaborate the natural rhythm, Yejin collected the data of sky. In one spot, she took photos of sky every five minutes every once a week for a year. To describe the artificial rhythm, Wooshik generated new images with the collected data of sky photos using artificial intelligence.

We made a series of videos composed with different frames. There are one frame that is the artificial rhythm and other frames are the natural rhythm that changes in same time but in different days. These videos contrast the relatively quiet, repetitive, and mundane daily rhythm with the noisy, random, and unexpected artificial rhythm. To emphasize the two different rhythms, we add the musical tones that correspond to the color of the images. As the image gets darker, it plays a lower pitch tone and as it gets brighter, it plays higher pitch tone.

Reflection:

We wish there is more variation on pitch tone that could match with the complex color variations.

Reference:

[1] Days and Nights by Yejin S. Lee http://yejinslee.com/drawing/2017/9/16/daily-rhythm2017

[2] Ian J. Goodfellow et al, Generative Adversarial Nets, 2014 https://arxiv.org/abs/1406.2661

[3] Cheng-Zhi Anna Huang et al, Counterpoint by Convolution, 2019 https://arxiv.org/abs/1903.07227

[4] Chengzhi Huang et al, Coconet, 2017, https://magenta.tensorflow.org/coconet

CODE: https://github.com/briankim13/art_ml_project_final

RESULT: