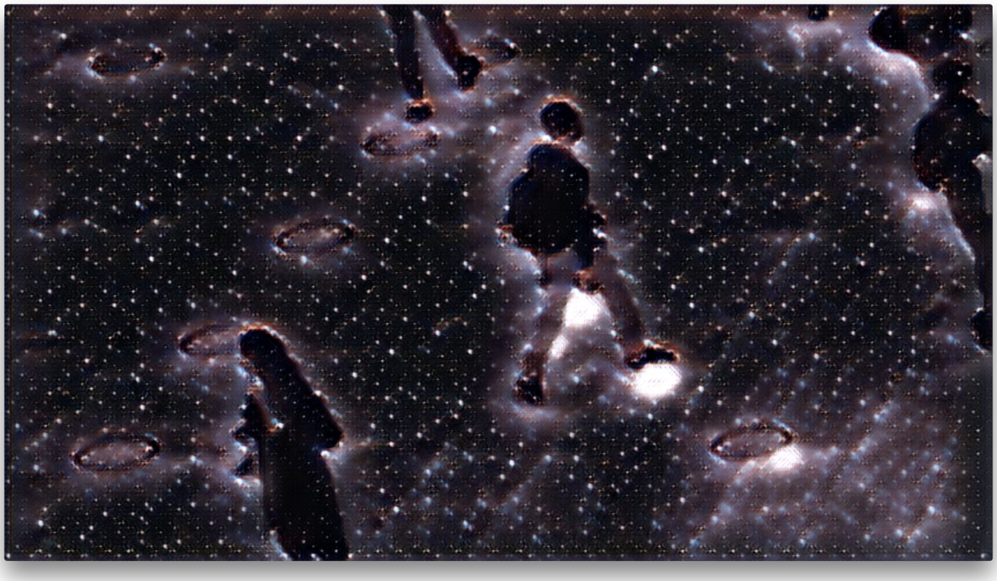
ART AND MACHINE LEARNING
CMU 2019 SPRING
PROJECT 1

THE EPHEMERAL EGO



Group 13
Wooshik Kim
Yejin Lee
Mimi Wang
Shengzhi Wu

ART AND MACHINE LEARNING
CMU 2019 SPRING
PROJECT 1

DESCRIPTION

Concept:

The idea of this projects are based on the Buddhist idea of Oneness removing ego.

Video 1: Combining ephemeral videos with stone images, we wanted to contrast the lightness of the ephemeral nature with the heaviness of human desires of permanence and possessiveness from stones.
Video 2: Many art style transfer efforts have been made to generate interesting and compelling images, but with our interests in 3D immersive spatial computing, we believe there are also opportunities to produce volumetric images using Machine Learning. Inspired by Refik Anadol studio's (http://refikanadol.com/) media artworks, we tend to make some experiment to generate dynamic 3D rendering using the art style transfer and Deep Dream. As such, the stylized videos are feeded into Unity Engine, with a real time generated particle height map using Unity Visual Effect Graph, we were able to render over million particles in real time with a GeForce 1080 GPU. The stylized pixels becomes sculptures showing the poetics of machine learning with an immersive experience. The approach, we believe, offers an alternative way to visualize data, not only make them materialized and tangible, but also poetic and beautiful. (Final video please find here)

Video 3 & 4: Followed by the Buddhist' idea of Oneness which talks about an interconnectedness between self and the others, we tried to connect different peoples and cars as a part of galaxy in a larger view and as a part of particles in a smaller view. The style content of video 3 is Kim Hwan Ki's painting in 1970s. His painting also combine the Buddhist idea of Oneness and Minimalist idea that everything is composed with geometry shapes.

Technique:

Main machine learning algorithms used for this project are style transfer for still image and style transfer for videos. The implementation detail of the two are slightly different. However both are similar in that they use optimization to minimize the style difference measured by Gram matrix of the features and content difference measured by the values of a convolutional layer.

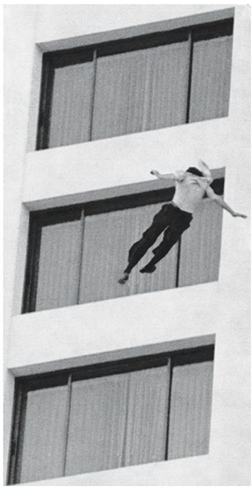
1. Style transfer implementation is that of "simulacre7" from Github, modified by Eunsu Kang ("kangeunsu") [2]. We experimented with different weights for style, content, and number of iterations.
2. Style transfer for video implementation is that of "lengstrom" from Github, modified by Eunsu Kang ("kangeunsu"). For video style transfer it takes longer time to learn new style. On a p2.xlarge instance of AWS, it took around 7 hours to train one epoch. Therefore, the experimentation was limited as we had little time to try different parameters. We found one epoch of training was enough so the style was learned for only one epoch. Another important parameter is weight of style and content. One of the

style we used was Hwan-ki Kim’s artwork. To achieve distinguishable content level, we reduced the style weight to 10 from 100. We kept all the other parameters constant.

Process:

1. First we tried a style transfer using different images. We tested with minimalist paintings for the style image to have sympler texture. We also used photos for the style image instead of just using paintings.

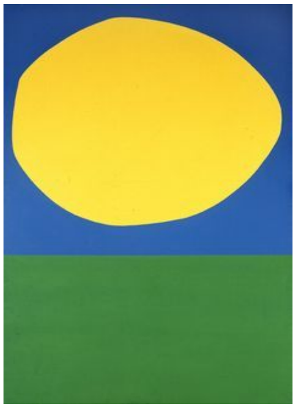
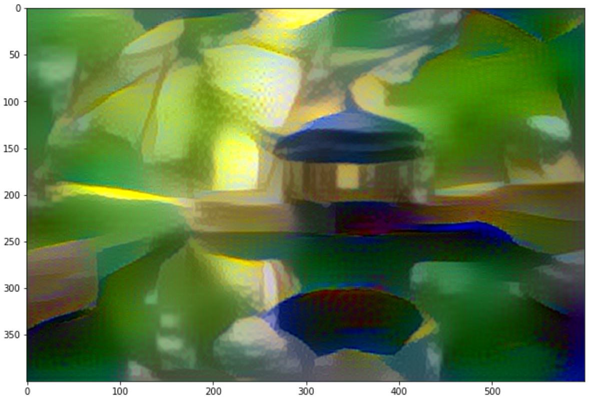
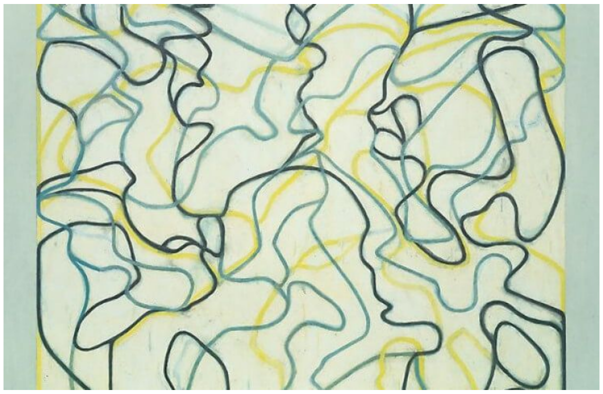
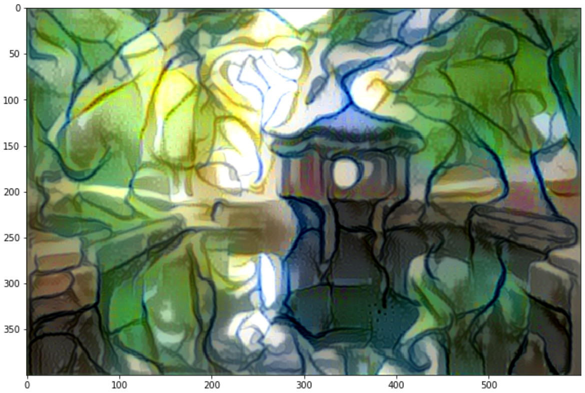
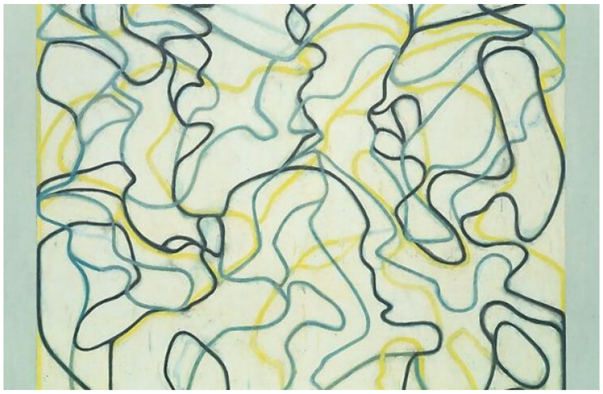
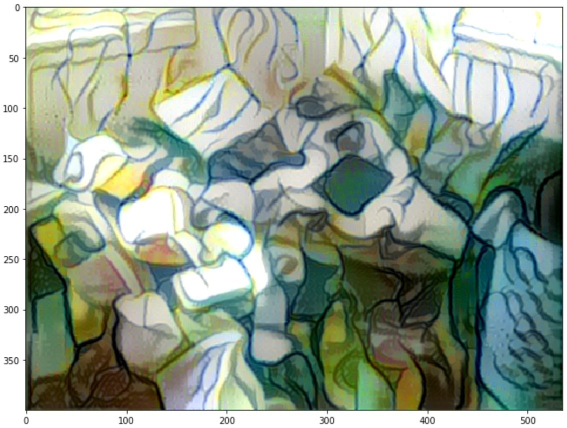
<Content images we used.>

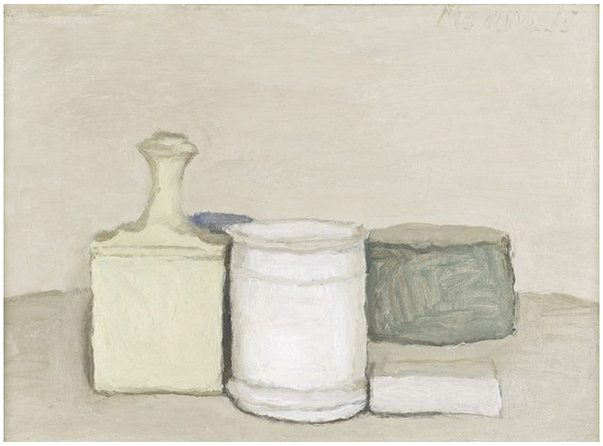
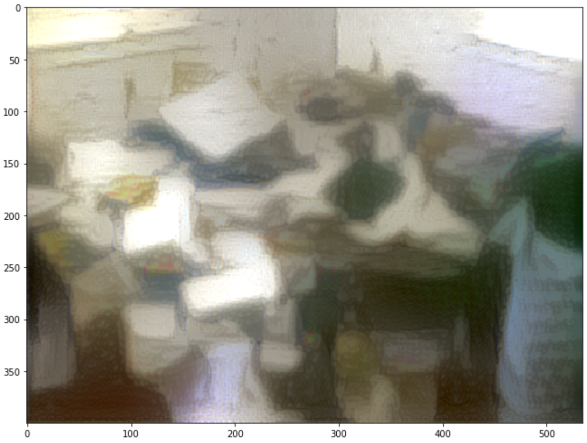
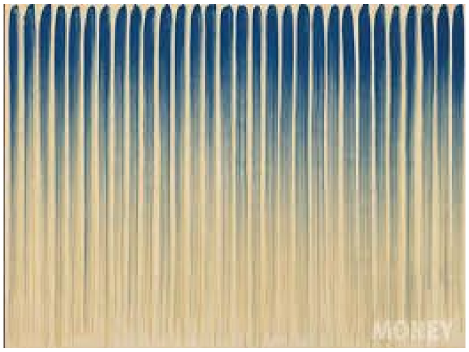
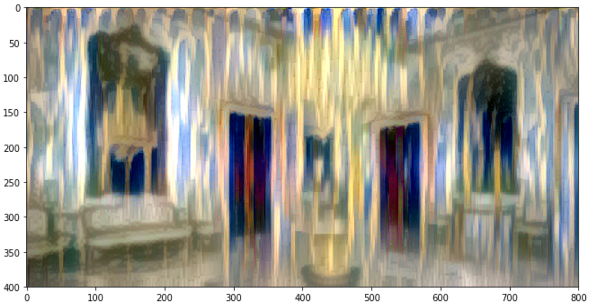
		

<Style images: Minimalist paintings>

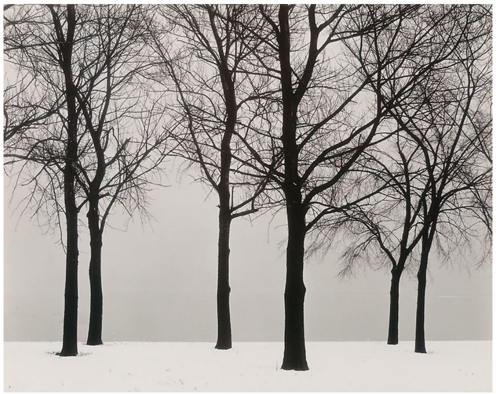
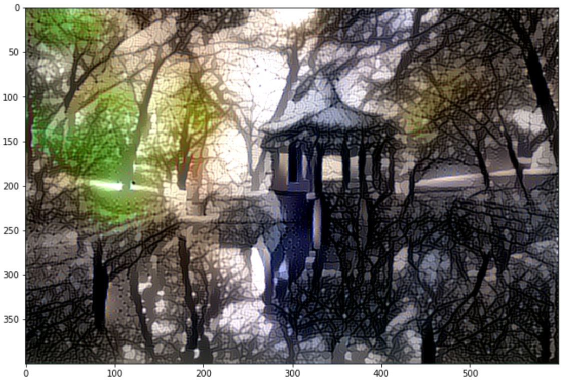
ART AND MACHINE LEARNING
CMU 2019 SPRING

PROJECT 1

<table>
<tr><th>Style images</th><th>outcome</th></tr>
<tr><td></td><td>
</td></tr>
<tr><td></td><td>
</td></tr>
<tr><td></td><td>
</td></tr>
</table>

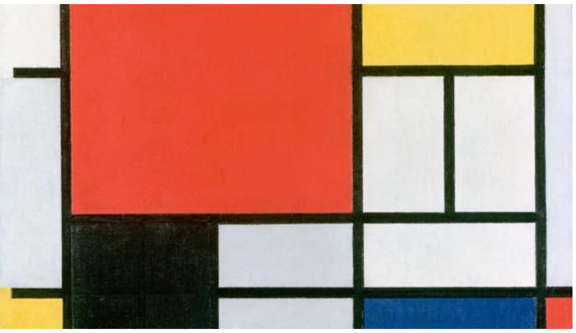
	
	

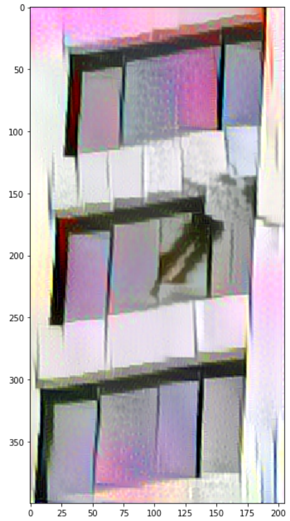
<Style images: photos>

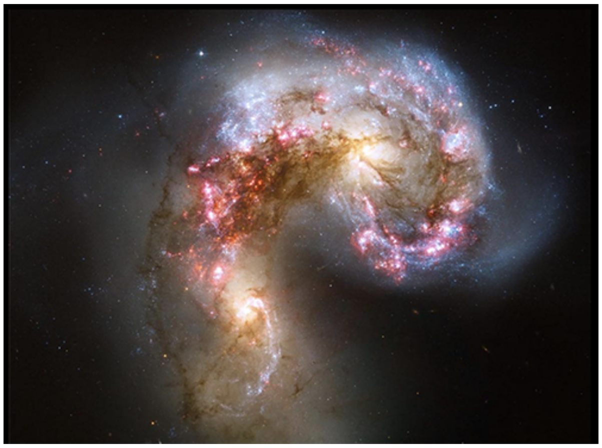
Style images	outcome
	

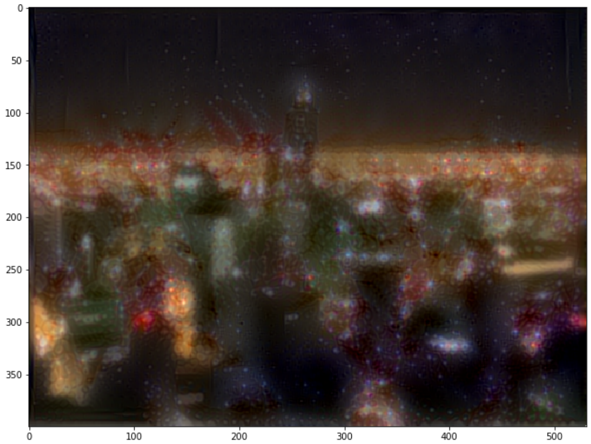
ART AND MACHINE LEARNING
CMU 2019 SPRING

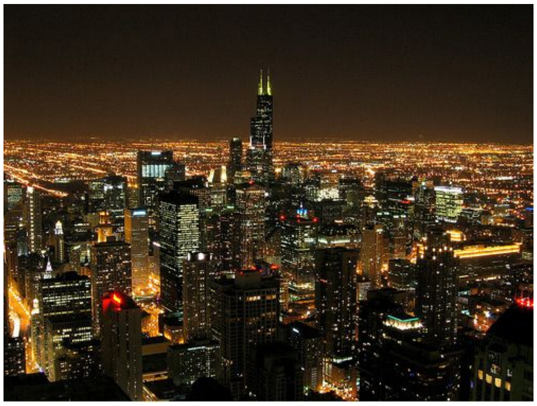
PROJECT 1

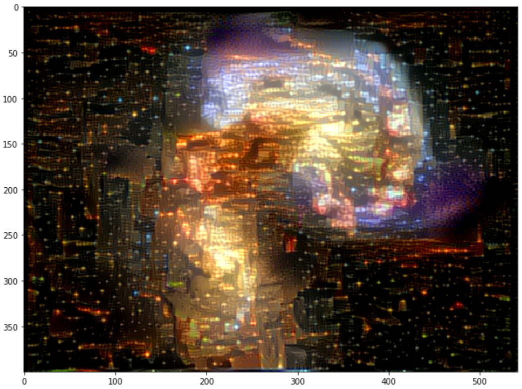


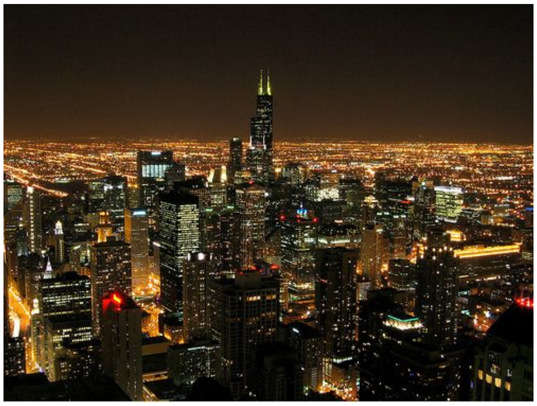
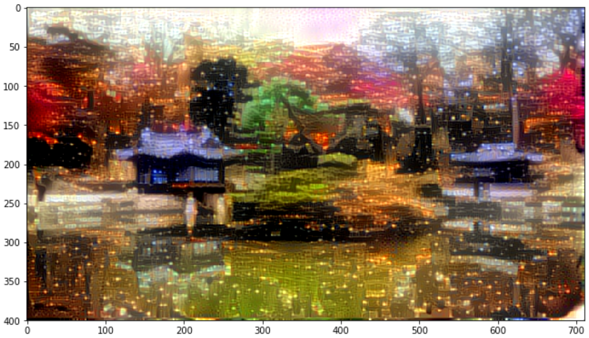
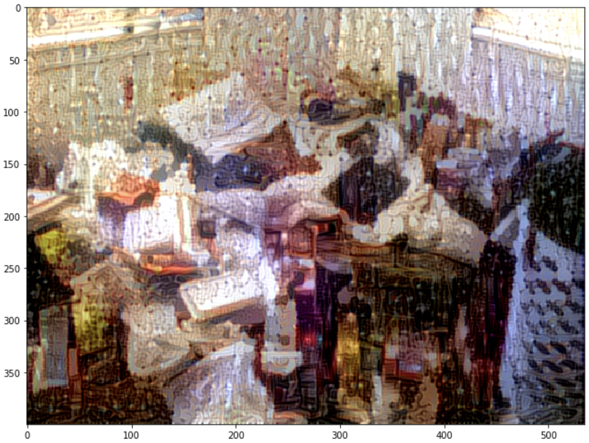




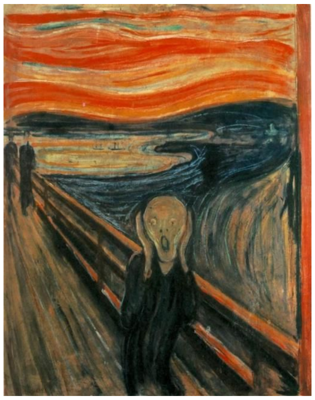
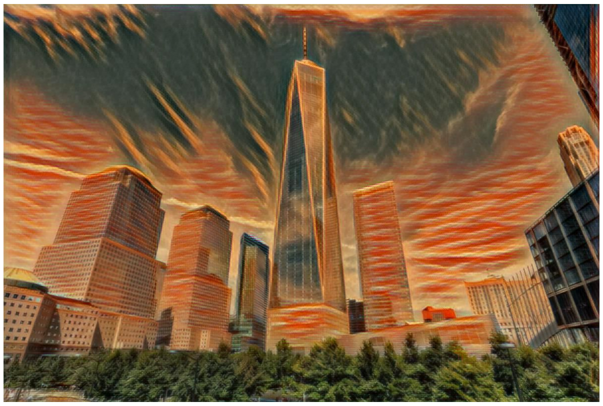


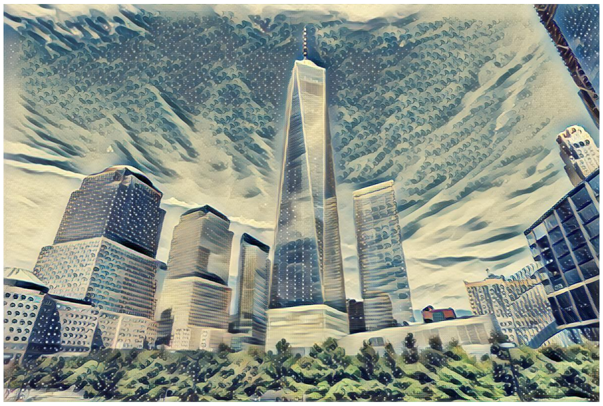
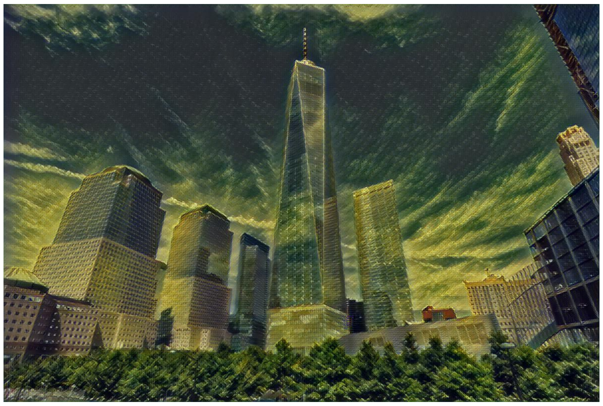
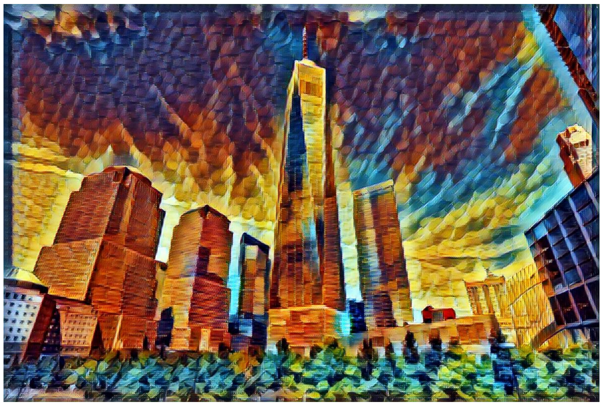




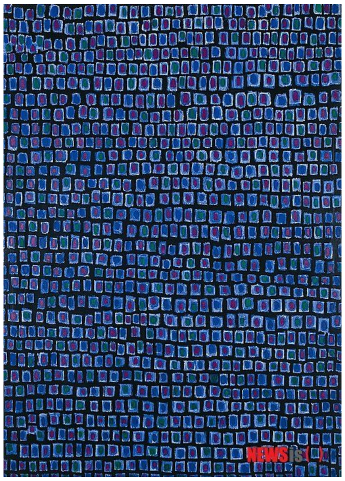
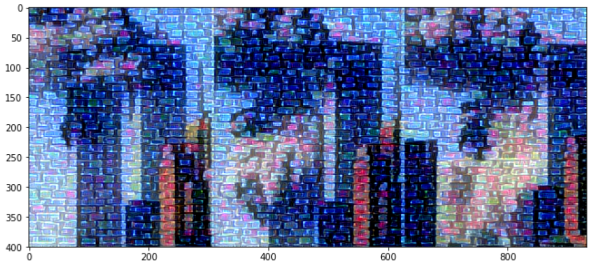
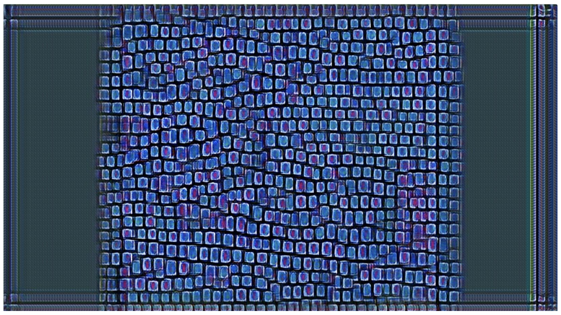
	 0 50 100 150 200 250 300 350 400 0 100 200 300 400 500 600 700
	

2. After we understood the style transfer, we tried the fast style transfer and find out these two techniques bring different outcomes. We tried evaluating image of world trade center with pre-trained styles to see what the outcome looked like.

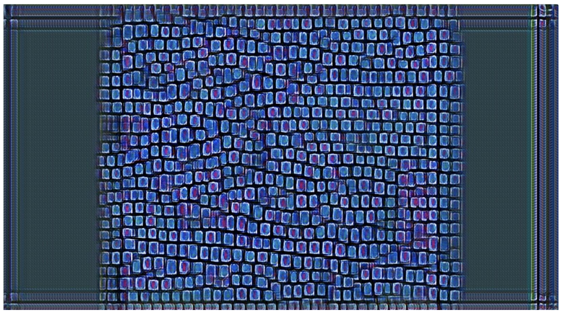
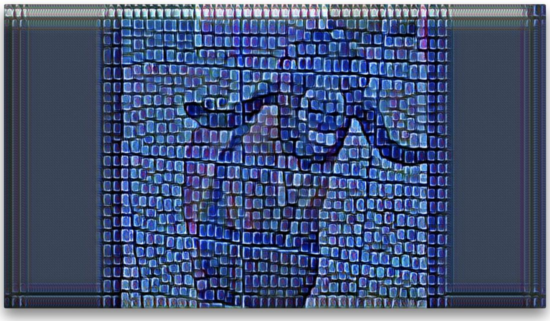
Style images	Outcome
	

3. At this point we decided to learn the style of Hwanki Kim's painting. With the same style image, the fast style transfer outcome was harder to read than the style transfer outcome. The implementation of fast style transfer picked up the style of image much stronger than the style transfer implementation. Below are results with original weights on style transfer and fast style transfer.

Hwanki Kim’s Painting	<Style Transfer>
	 0 50 100 150 200 250 300 350 400 0 200 400 600 800
	<Fast Style Transfer>
	

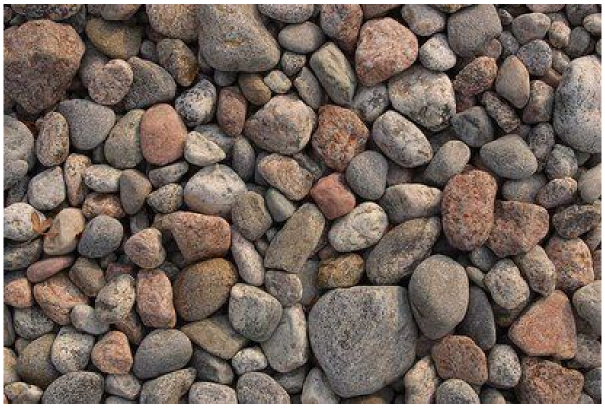
We needed to control the weight to achieve a readable output. We controlled the style weight to make a readable video.

<Fast Style Transfer with an original style weight>	<Fast Style Transfer with an lesser style weight> **
	

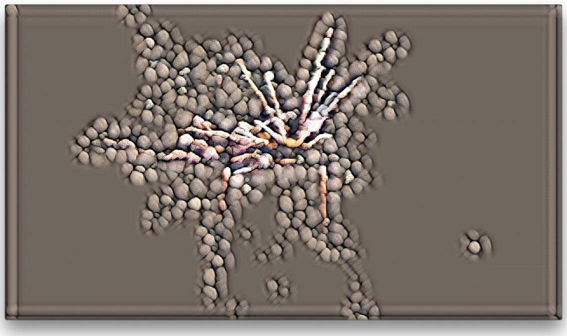
<video 1>

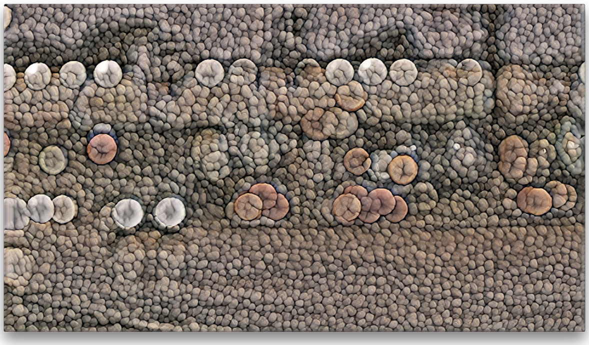
We collected videos for the content that talks about an ephemerality such as falling buildings, fireworks, and explosions.

We use stones for the style images to describe human desires of possession and permanence contrasting with the ephemeral nature.
<content image: stones>



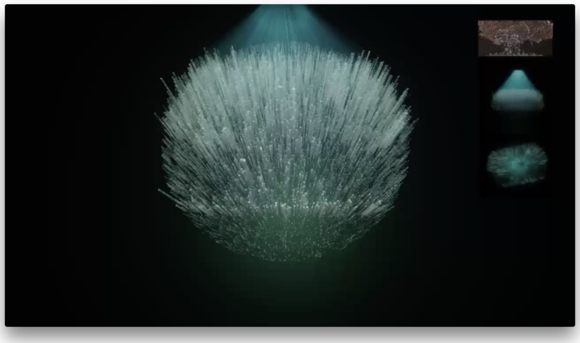
test.1	test.2

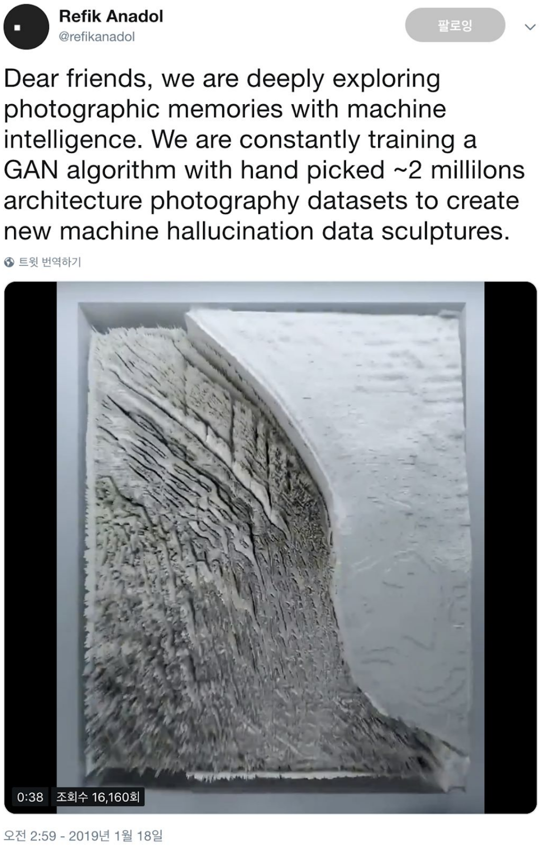




<video 2>

Adding Unity Height-Map Post Processing Effect





Refik Anadol
@refikanadol

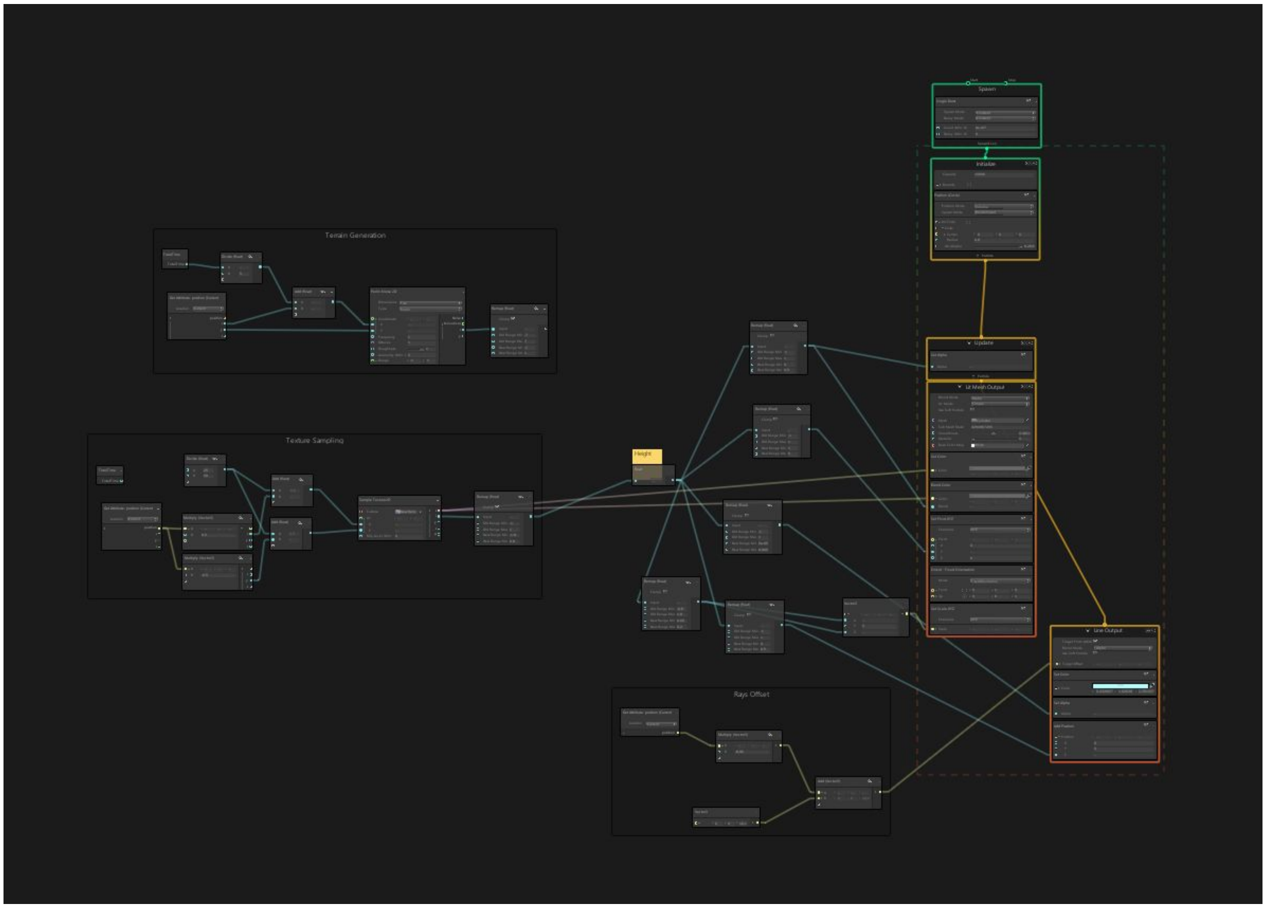
Dear friends, we are deeply exploring photographic memories with machine intelligence. We are constantly training a GAN algorithm with hand picked ~2 millilons architecture photography datasets to create new machine hallucination data sculptures.

트윗 번역하기

0:38 조회수 16,160회

오전 2:59 - 2019년 1월 18일

We used various of techniques to generate images, but we were also inspired by Refik Anadol's works where he transfers GAN results into a dynamic 3D sculpture. So we thought it might be also interesting to utilize the ML 2D image result to generate 3D contents. We want to see how the stylized texture looks like in a 3D environment.



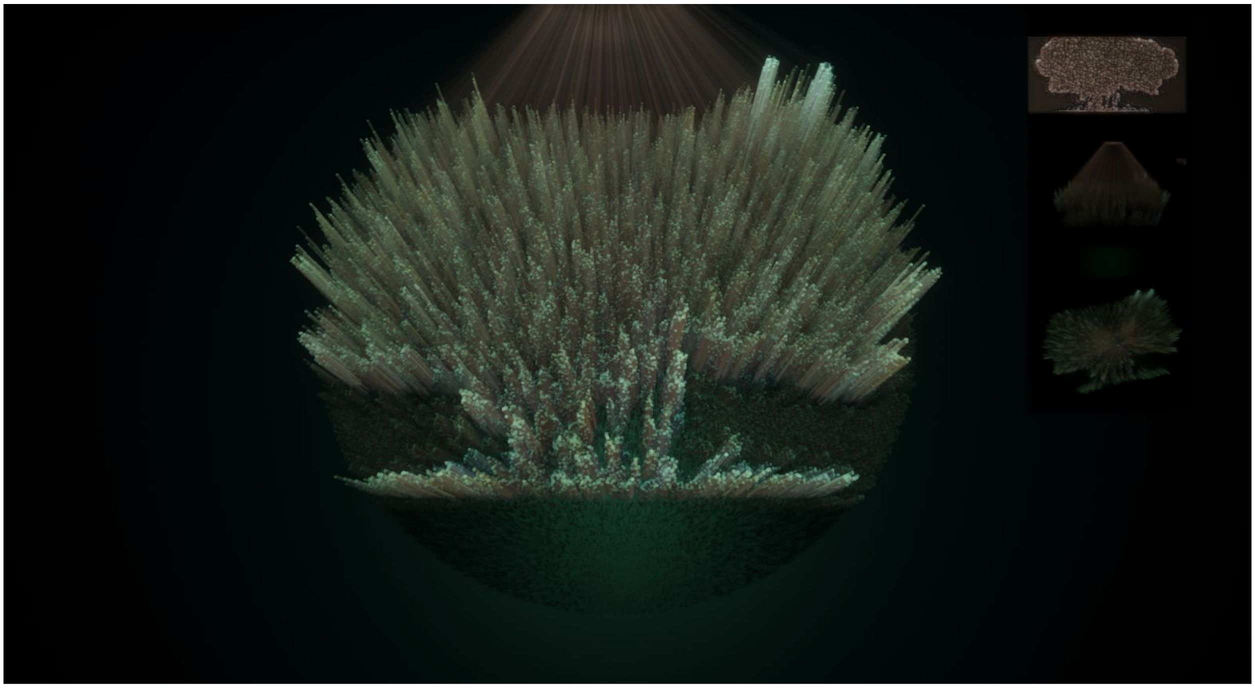
Visual Effect Graph Node Screenshot

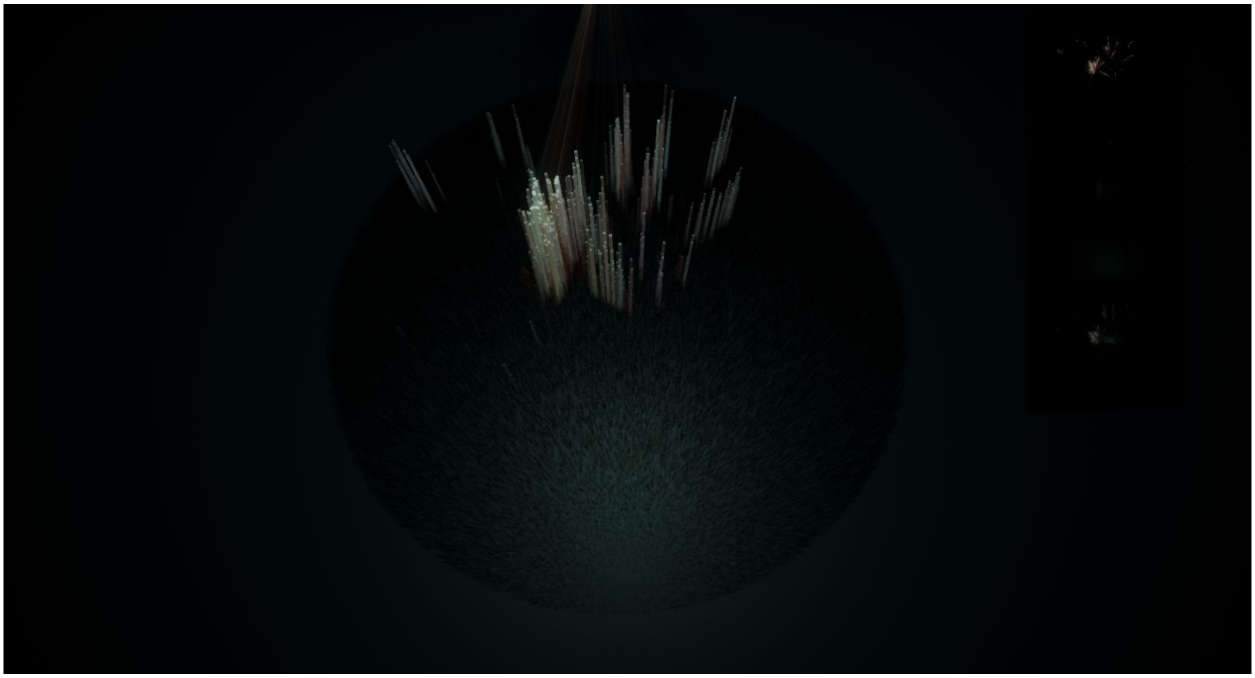
Therefore, we use Unity (a game engine) to procedurally generate a height map based on the video texture input. The way to realize it uses an Unity recent released plugin “Visual Effect Graph, which is a node based programming language for producing visual effects. The dynamic height map is generated in real time with over millions of individual particles with a GPU acceleration.

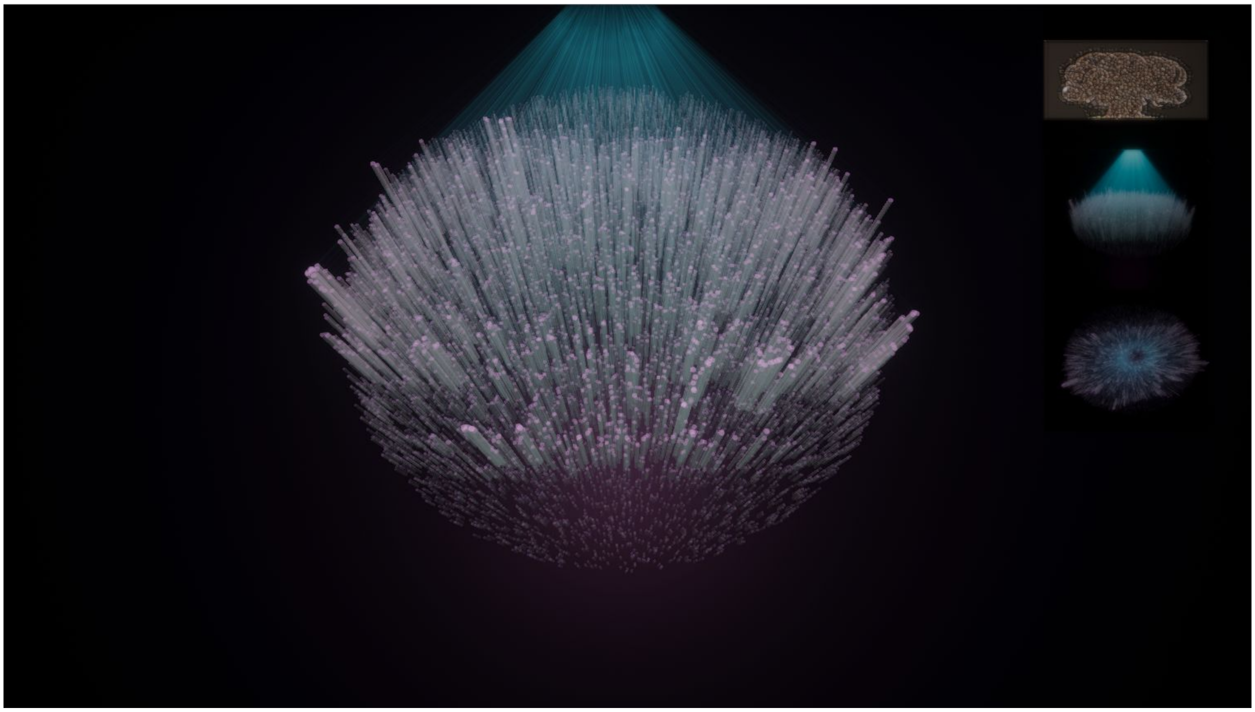
The concept is basically to map the pixel brightness from the original stylized video to the height map, and the brighter a pixel is, the higher the cylinder particle would be. We also played with various number of the total particles in the scene, and it’s no surprise that more particles usually helps us to get more visually compelling result, but due the the local real time computing power limitation, the rendering started to be lagging when we get to millions of particles, but we still found that 10,0000 particles can already provide satisfying result. Please check out the screenshots from our experiments. (videos can be found here https://drive.google.com/drive/folders/1QuxnqgUGds1SmG3dOh-j998j0aQ02n0s?usp=sharing)

ART AND MACHINE LEARNING
CMU 2019 SPRING

PROJECT 1

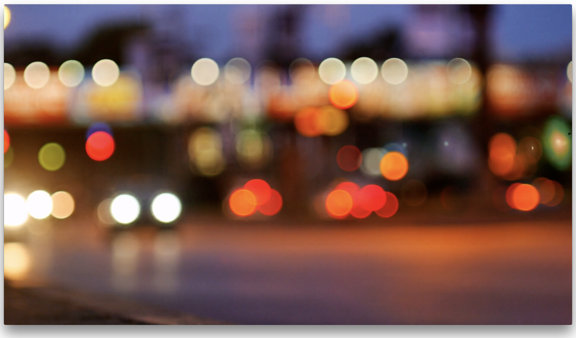
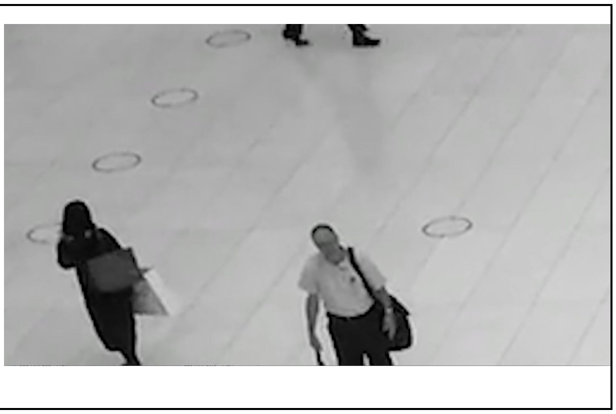




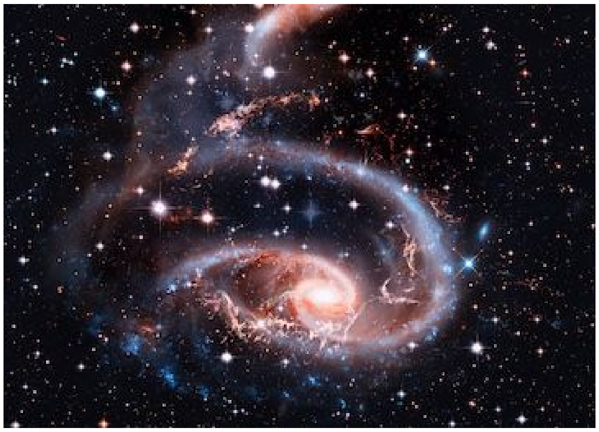
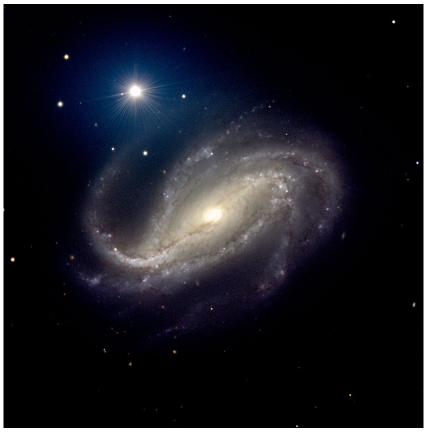


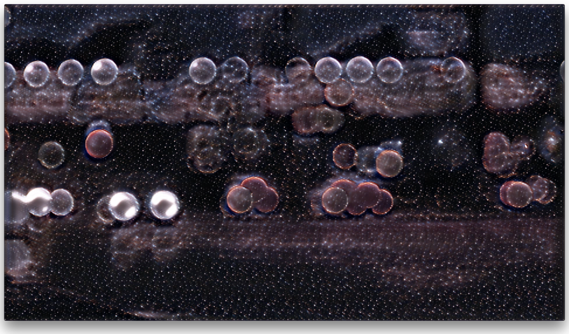
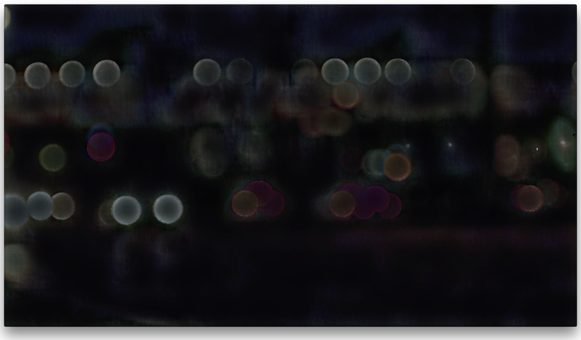
<video3 & 4>

we used daily scenes such as passing cars and people as a content.

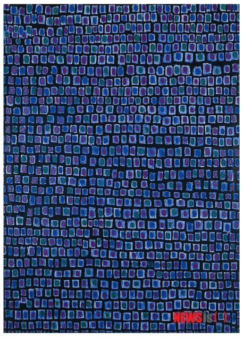
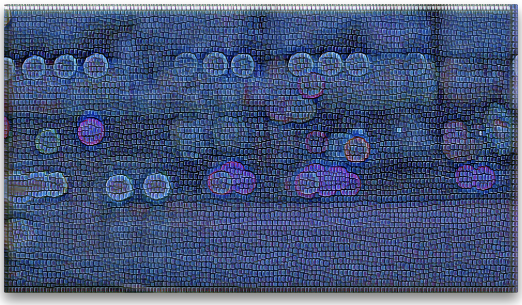
	

-we tested different kinds of galaxy images because most of galaxy images are hard to transfer the style of the images.

<Galaxy picture.1>	<Galaxy picture.2>
	

<Outcome from picture.1>	<Outcome from picture.2>
	

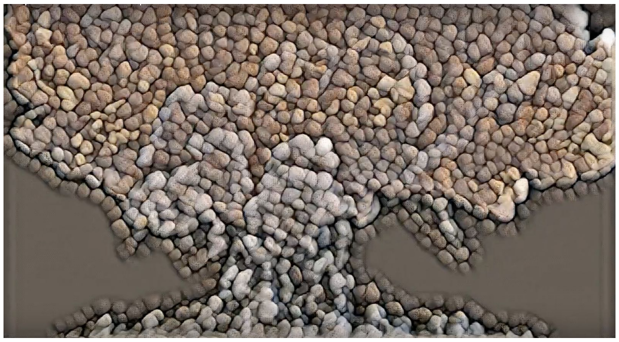
- We also used Kim Hwan Ki’s painting as a style transfer.

<content image>	<outcome>
	

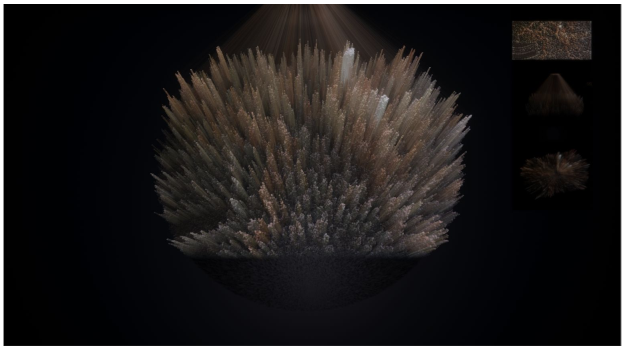
Result: Please describe your final result(s). Add thumbnails of your results as images if possible.

In the video 1 and 2, we used found videos of explosion and fireworks and combine them as a series. To contrast the lightness of the ephemeral nature with the heaviness of human desires of permanence and possessiveness, we used a stone image as a style image (video 1) and 3D Sculpting Style Transfer (video 2). We also slow down the video to give more heaviness.

<video 1>

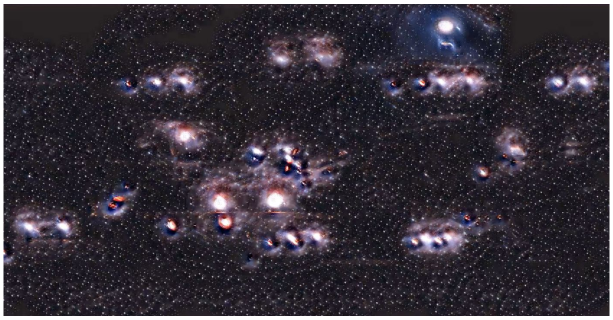
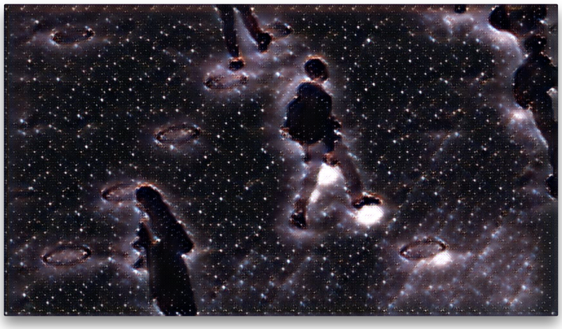


<video 2>

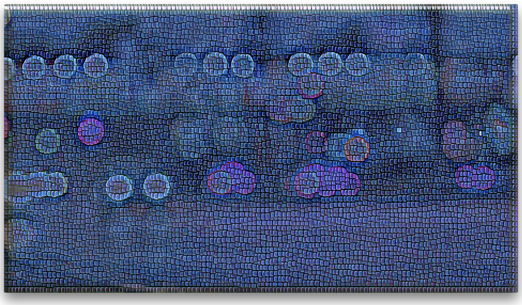
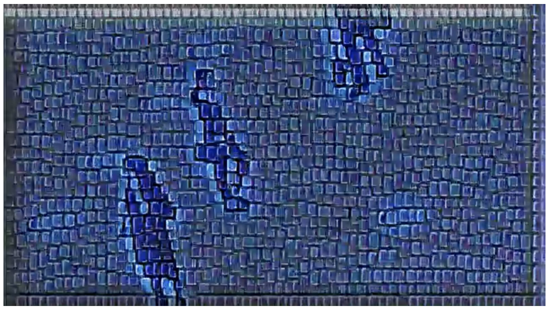


The passing people video is created from a found clip. For the passing cars video, we actually took videos on the street. To add a musical element, we combined different frames and show

them together. In video 3, we used a galaxy image as a style image and in video 4, we used Kim Hwan Ki's painting as a style image.
Add
<video 3>

<video 4>

Reflection: You may write why the final result was chosen among your results, how much satisfied with your ideas and techniques for this project, any new ideas that came into your mind, any reflections about trials and errors, and/or what you have learned etc.

After understanding the techniques, It was hard to match these techniques with the concept of an art work. If we could develop more, we want to develop the content video more to make it more sense with the techniques we used.

3D Sculpting Style Transfer:
We still find generating 3D content using machine learning is interesting, but the height map concept requires a continuous texture input to get an interesting results, otherwise the detailed texture would lose its fidelity through the transferring process and only turn out to be random noise. But with a video generated from Deep Dream and BigGAN, the result is quite satisfying.

Another question we still need to figure out is the logic of why the 3D content is important in this scenario. As we referred earlier, Refik Anadol's project transfers 10,000+ architecture images using ML, so it makes more sense to have it in 3D, while the question still remains unanswered in our project.

Another interesting concept came along the way is that what if we use 2D image generating ML method to produce 3D content directly. For example, in this project, we use a post-processing technique for the height map, but potentially the algorithm can generate 3D point cloud directly. One idea in my mind is that if we use 3D face point cloud data from Kinect, and transfer the z axis data into the alpha channel, make it into a 2D png image, but still keep its x,y position and RGB value, then use Pix2pix to generate new png images with the alpha channel. At last, transfer back the alpha channel data to the z axis for rendering, would we be able to get a 3D point cloud? Or instead of using 2D image as input, we apply the same method to DeepDream, it might generate very interesting 3D content.

Reference: NOT REQUIRED. If you have cited any reference, please add them here.

[1] Eric j. Heller, Transport 2, 2000
https://www.thedailybeast.com/larry-nassar-judge-aquilina-sentencing-victim-death-warrant?ref=scroll
[2] simulacre7, Github Repository,
https://github.com/simulacre7/tensorflow-IPythonNotebook/tree/master/neural-style

https://twitter.com/refikanadol/status/1086216491203264512?s=21&fbclid=IwAR2ogSufCLKtQOfh0ff8JZsLJE_rtZaM2ObLqb2ZEwsZRLA8IaCvzj5NNbI

[3] Visual Effect Project Reference
https://github.com/emeraldpowder/VfxTable

CODE:
Fast Style Transfer on Live Cam
https://github.com/briankim13/art_ml_project1.git

Height Map Generator with ML Video Executable File:
https://drive.google.com/file/d/10e1QUYFPm7T5qH0_6O7s0CulCU6gBW4h/view?usp=sharing
(PS: it only runs on Windows 10 with Nvidia GeForce 1070 GPU or plus)

RESULT:
https://drive.google.com/drive/folders/1pLtK7a15tTVIuqa4e4tP-pZntM93EY5g?usp=sharing