

# Statement of Work

Harvard AC215: Advanced Practical Data Science

Dongyun Kim, Vasco Meerman

## 1. Problem Definition

With the advent of Artificial Intelligence, it has changed the world in different ways such as an autonomous car in computer vision and language translation in natural language processing. However, AI is only limited for those who can know how to code, have statistical knowledge and mathematics background even though it has a lot of potential to be used in many fields. In particular, Generative Adversarial Network (GAN) has shown interesting features such as latent space and feature extraction, and skilled artists or affordable artists who have in-house engineering teams take advantage of GANs as a tool to help their artworks or create their own artworks, for example, Refik Anadol while normal artists have not had opportunities to take advantage of novel technologies so far.

## 2. Proposed Solution

In this project, we want to help artists and designers who are marginalized due to the technology gap and help those who are in the blind spot in technology. To open the AI gate for them, we propose a simple application that allows them to upload and/or scrap images and experience art-related AI techniques with an easy and intuitive interface. The application is a new kind of AI-powered creative tool for artists, designers, and creators. No need to have an AI background.

One of main advantages of using GANs is to extract features and create latent space based on input's characteristics. It means that it could be used to understand existing artworks in different ways through the lens of artificial intelligence and it can give us a new way to interpret its characteristics and unveil and discover hidden insights.

## 3. A rough timeline and components

- 1) We want to first create and/or use existing tools to scrape images from websites like Flickr and Pinterest to add to our datasets.
- 2) Then we want to compile all the images we have to our final dataset.
- 3) Clean and prepare our data into the right and efficient format to use for our models.
- 4) Research existing StyleGANs or similar models to use.
- 5) Apply transfer learning to create a new model for our needs.
- 6) Create a hybrid application that can use the model to change images and/or videos. We want to use a hybrid app/web framework like React.js or Flutter to create our application.

## 4. Project Scope

For this project we will try to create an application that can apply different art styles to existing images and videos. To achieve this, we will use images from multiple internet sources, like existing datasets from kaggle and we will also look at scraping images from platforms like Flickr and Kaggle.

We will use Docker to containerize the different parts of our project; compiling our dataset, the data pipelines and training of our models and our application.

Since training a full model would take too much resources and time, we will apply transfer learning to create a custom model.

## 5. Datasets and models being considered

### - Datasets

<https://artsandculture.google.com/category/artist> and <https://artsandculture.google.com/category/art-movement> (Google Arts & Culture database that categorizes art works based on artist and art movement. The category could be our topic in which GANs will generate images)

<https://www.kaggle.com/ipythonx/wikiart-gangogh-creating-art-gan?select=portrait> (Wiki-Art : Visual Art Encyclopedia)

<https://www.kaggle.com/ikarus777/best-artworks-of-all-time> (Best Artworks of All Time)

### - Models being considered

#### 1) Neural Style Transfer ([https://www.tensorflow.org/tutorials/generative/style\\_transfer](https://www.tensorflow.org/tutorials/generative/style_transfer) )

Neural Style Transfer manipulates digital images, or videos, in order to adopt the appearance or visual style of another image.

#### 2) Pix2Pix (<https://phillipi.github.io/pix2pix/>)

Pix2Pix uses a conditional generative adversarial network to learn a mapping from an input image to an output image. An example of a dataset could be input image which is a building sketch and output image which is a the color version of the building. Pix2Pix model tries to learn how to colorize the image from the sketch.

#### 3) CycleGAN (<https://junyanz.github.io/CycleGAN/>)

Although Pix2Pix shows awesome results, there is a limitation in the training dataset which requires a labelled dataset. To tackle this problem, CycleGAN allows us to use an unpair dataset.

#### 4) StyleGAN2-ADA (<https://github.com/NVlabs/stylegan2-ada>)

StyleGAN is a model that learns features from dataset images and creates latent space based on its features. It could be used to generalize features in the dataset and explore the features in the latent space. By traversing latent space, it might give us hidden

#### 5) pixel2style2pixel (<https://github.com/eladrich/pixel2style2pixel>)

pixel2style2pixel (pSp) framework provides a fast and accurate solution for encoding real images into the latent space of a pre-trained StyleGAN generator. The pSp framework can additionally be used to solve a wide variety of image-to-image translation tasks including multi-modal conditional image synthesis, facial frontalization, inpainting and super-resolution.