

# Image to Image translation using GAN

## Project Proposal | DS5500 | Fall 2019

Anish Narkar, Nikhar Gaurav, Prasanna Challa, Prathwish Shetty  
Northeastern University, Boston, MA, USA

### Summary:

Image-to-Image translation is to learn a mapping between images from a source domain and images from a target domain. In this project, we investigate generative adversarial networks (GAN) for the purpose of converting satellite images to maps. The performance of traditional GAN system in detecting edges could be improved by introducing a superior edge detection architecture like U-Net to detect edges in the input, before feeding it to GAN. This could potentially improve the process of translation as the edges are well defined in the modified input, generated by feeding the actual input to U-Net.

The dataset has been obtained from pix2pix which is comprised of 1096 satellite images of New York and their corresponding google maps pages. The image translation problem involves converting satellite photos to Google maps format. We intend to leverage the Google Maps Static API (part of Google Cloud Platform) to obtain additional aerial satellite view along with its processed view to train our models.

### Proposed Plan:

The project aims to develop on top of an already existing research of Image-to-Image translation with conditional adversarial nets [1]. The satellite images will be segmented to detect the exact edges/ blocks of each entity in the image. For the purpose of segmentation, we will be using U-Net [2] which is a widely used technique in the field of medical science in applications like finding anomalies in medical images. The obtained segmented image will be superimposed on top of the satellite image to obtain a modified input with well-defined objects and sharp edges. The output of U-net will be passed through GAN to obtain the image map.

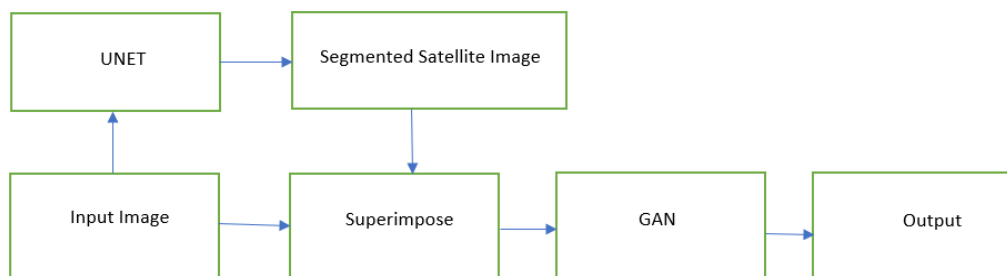


Figure 1 Current Plan of Execution

The need for superimposition is illustrated by Figure 2. The colors in satellite image(left) are basically different shades of green, this makes it difficult to contrast streets from the landscape, thus the predicted image (center) fails to identify these obscure paths leading to inaccurate output. This will be the major problem which we will try to address in our project.



Figure 2 Limitation of reference

### Preliminary Result

Applying U-Net will give us output which identifies distinct segments in the given image. We applied this technique to the image on left in Figure 3. The contrast of the street along the river bank will make it difficult for GAN to identify it, however the result which we obtained (Right) clearly differentiates streets from the landscape thus increasing the likelihood of successfully identifying the mentioned street.

Also, the dataset which we will be using has ~1100 images, apart from these we have also used Google Maps static API to obtain images from additional locations.

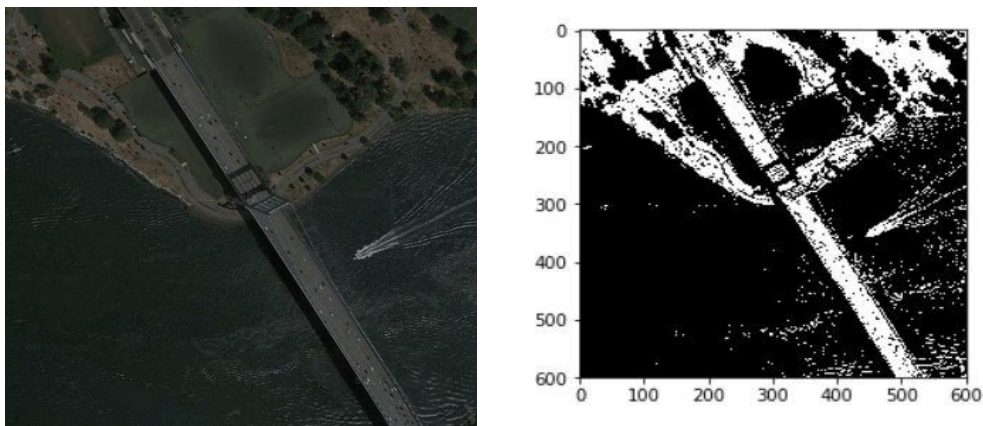


Figure 3 Preprocessing of Images

**Evaluation:**

To evaluate the accuracy of the predicted output, we will leverage the actual map of the input. There will be two segments (classes) in these images, streets and landscape. There are three metrics which are commonly used to evaluate image segmentation:

1. Intersection over Union (IoU) It measures the degree of overlap in the predicted and ground truth. In our case, there are two classes, so IoU for prediction will be the average IoU for Streets and IoU for Landscape.
2. Dice score (F Score): Dice Score is similar to IoU in extreme cases, but tend is not as severe as IoU in penalizing a single instance of bad classification (IoU has a squaring effect relative to F1). The predicted score for these metrics should be closer to 1.
3. Pixel Accuracy: This is the percentage of Pixels which are correctly classified. This metric is inaccurate when dealing with class imbalance which we predict we will encounter in this case.

Benchmark for our project will be value of these scores for the output of our reference paper.

**References:**

1. [Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks](#)
2. [Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation](#)

**Github:** <https://github.com/anishnarkar/DS5500>