הפקולטה להנדסת
חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי

SIPL
Signal and Image Processing Lab

TECHNION
Israel Institute of Technology

Creating Image Segmentation Maps Using GANs

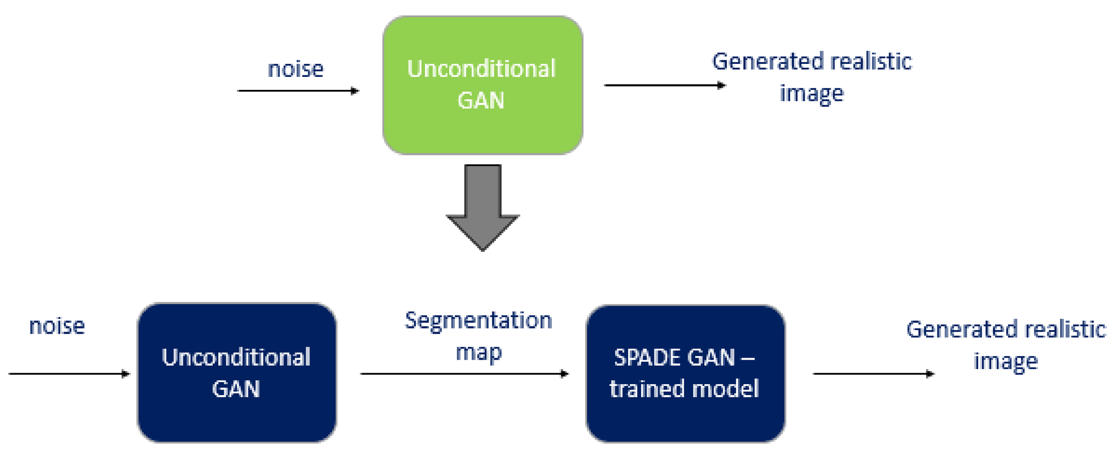
Inbal Aharoni and Shani Israelov Supervised by Idan Kligvasser

Introduction

- The use of GAN has drastically affected low-level vision in graphics, particularly in tasks related to image creation and image-to-image translation.
- With the success of GANs we will produce segmentation maps. With these maps and with the help of the generative model we can get a semantic understanding of the data set and even create completely new scenes.

Goals

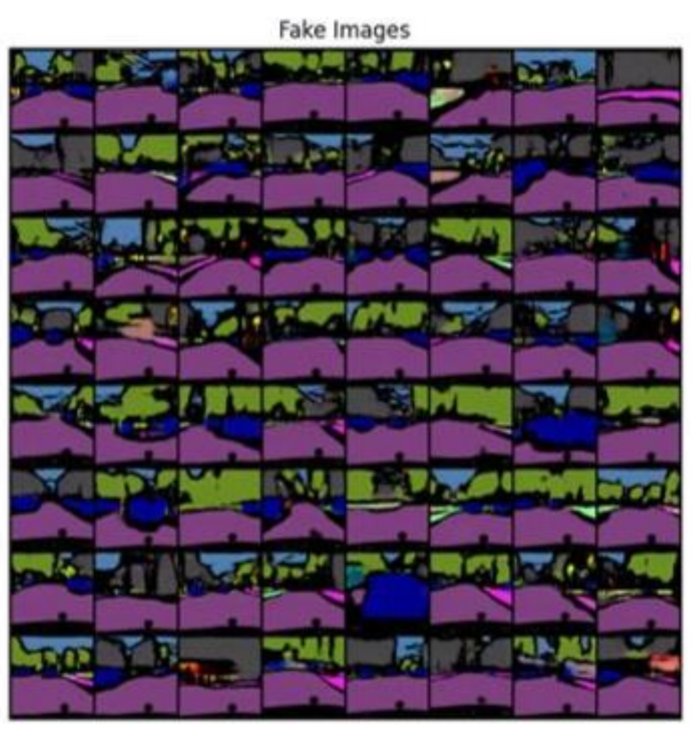
- The project goal is to create segmentation maps using unconditional GANs in order to use those images in another GAN which will create a realistic street image.
- Instead of generating street images directly from one GAN, our training process will have two stages:
  1. We will create a segmentation images of streets using styleGAN.
  2. We will use those images as an input for existing GAN (SPADE) that create and add a layer of texture to a segmentation images to create a realistic picture.



Project scheme

DCGAN

- In order to understand and learn about GANs, we used existing DCGAN that works for 3x64x64 face images and tried it on our data.

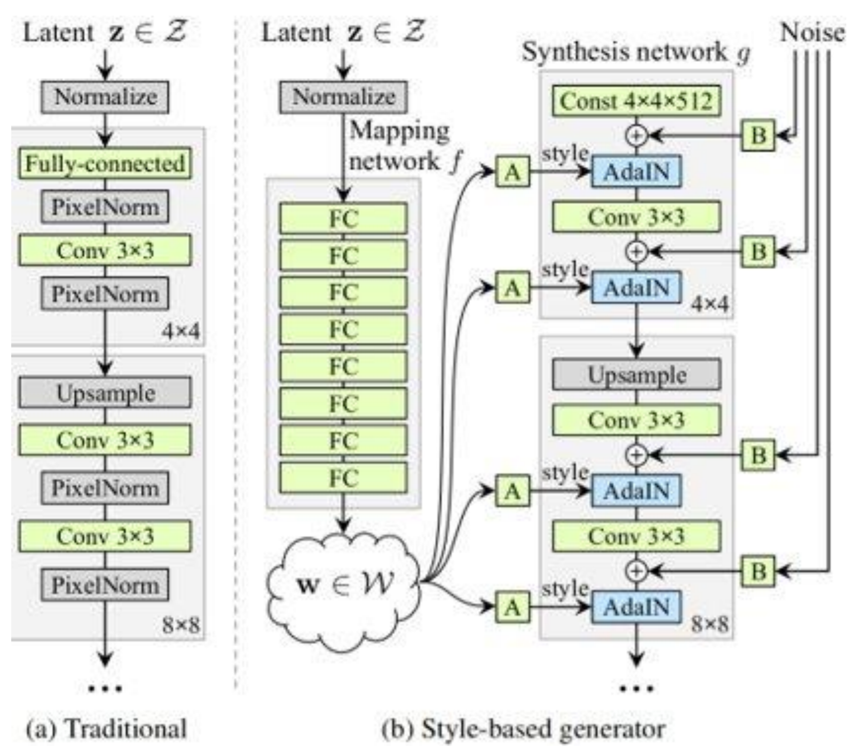


DCGAN results

- Problems with DCGAN Results:
  1. Distorted dashboard symbol
  2. No human are shown, only “redish” stains
  3. Places where the color is not uniform

StyleGAN

- StyleGAN not only allows for a better understanding of the generated output, but also produces images that look more authentic than previously generated images.



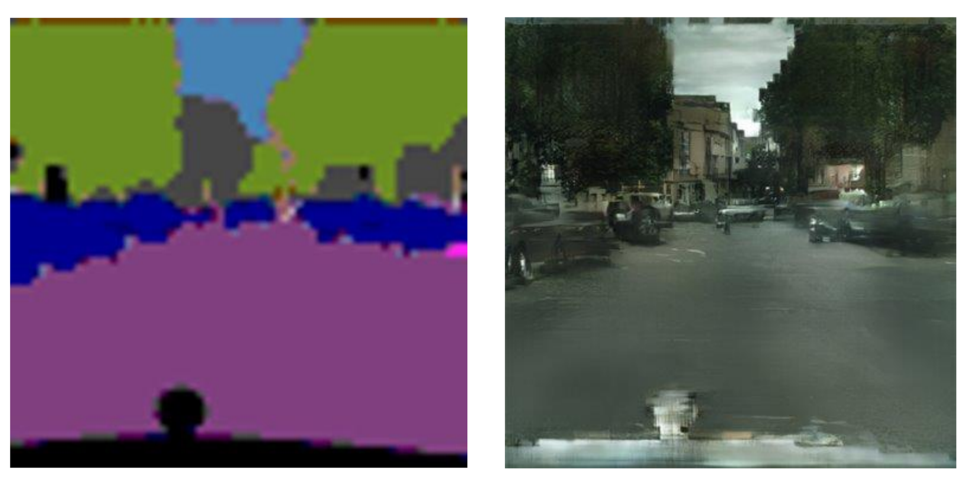
Traditional generator vs styleGAN generator

- <u>The Changes to the model include:</u>
  - Progressive growing.
  - The use of a mapping network to map points in latent space to an intermediate latent space.
  - The use of the intermediate latent space to control style at each point in the generator model.
  - The introduction to noise as a source of variation at each point in the generator model.

StyleGAN results

SPADE

- SPADE stands for spatially-adaptive normalization, a simple but effective layer for synthesizing photorealistic images given an input semantic layout.
- In order to use SPADE we had to do some adjustments:
  - Conversion of each pixel to the closest pixel value among the segmentation map options.
  - Create Label Map from the segmentation map.



StyleGAN and SPADE results

Comparison

- We compare our results to the use of only unconditional gan by using the same styleGAN model on the realistic images.

Unconditional gan results

FID

- For the evaluation of the performance of GANs at image generation, we will use the “Frechet Inception Distance” (FID) .
- FID captures the similarity of generated images to real ones better than the Inception Score.
- A lower FID indicates better-quality images; conversely, a higher score indicates a lower-quality image and the relationship may be linear.

$$\mathrm{FID} = |\mu - \mu_w|^2 + \mathrm{tr}(\Sigma + \Sigma_w - 2(\Sigma\Sigma_w)^{1/2}).$$

FID equation

Results

		FID
Our segmention map	original segmentation map	245.989
Our images after SPADE	Original realistic images	179.463
Images styleGan generation	Original realistic images	247.257

FID comparison table

Conclusions

- We created successful segmentation images of streets using styleGAN.
- We used these images as an input for existing GAN (SPADE) that created and added a layer of texture to a segmentation images to create a realistic picture.
- Our two-step process produced better results from the unconditional GAN alone.
- Two-step process simplifies the image generation process.

September 2016