THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING

SIPL
Signal and Image Processing Lab

TECHNION Israel Institute of Technology

Image Manipulation with GANs Spatial Control

Karin Jakoel and Liron Efraim, Supervised by Tamar Rott-Shaham

Introduction

- We suggest a new approach that enables **spatial** editing and manipulation of images using Generative Adversarial Networks (GANs).
- This ability is possible thanks to a test-time spatial normalization that uses the trained model as is and does not requires any fine tuning. Therefore our method is significantly fast and does not required further training. We demonstrate the new approach for the task of class hybridization and saliency manipulation.

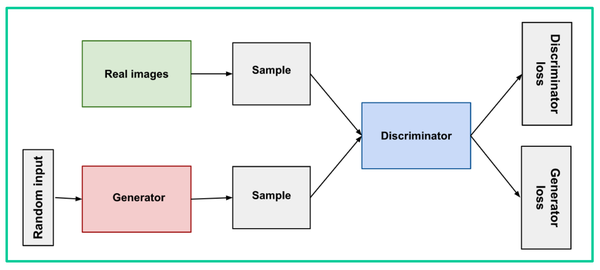
Goals

- In this project we want to control the generation by manipulating spatial regions.

Challenges

- BigGAN structure is duplicating the noise to all image so spatial changes are difficult to achieve.

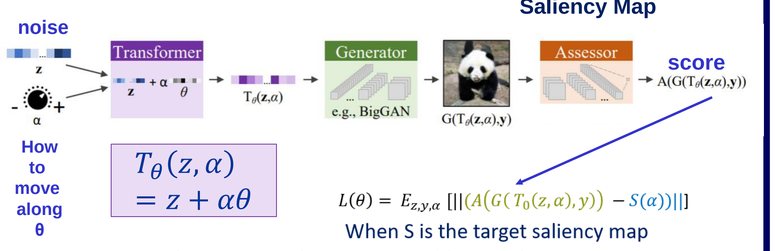
GAN



- Generative Adversarial Network (GAN)
- The core idea of a GAN is based on the "indirect" training through the discriminator, which itself is also being updated dynamically.
- This basically means that the generator is not trained to minimize the distance to a specific image, but rather to fool the discriminator.
- This enables the model to learn in an unsupervised manner.

GANalyze

- Study cognitive properties of image, as memorability and emotional valence.
- Find directions in BigGAN latent space that control these attributes and navigate according to it.



GANalyze Network structure with our changes

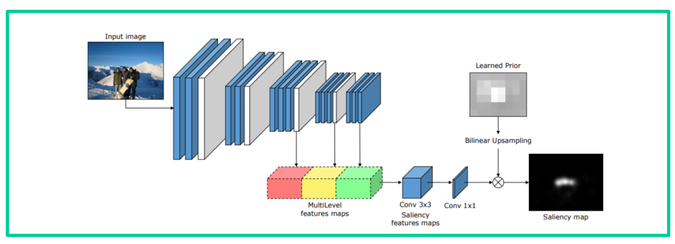
- We Replaced the:
  - Assessor to Saliency map detector
  - Loss Function
  - Added new transformation

  We also train on single class and single image since saliency is a property which is unique to image.

Saliency Detection

- What is saliency in image?
- - Identifying the fixation points that human viewers would focus on at first glance.
- - Highlighting the most important object regions in an image

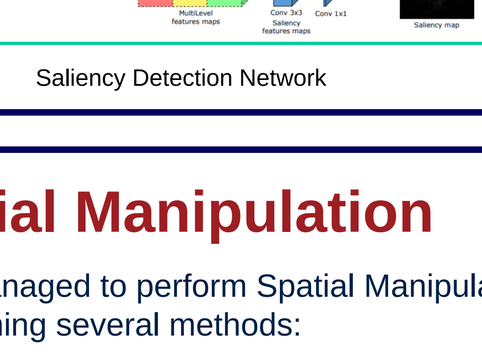
- Features detected:
- - lower-level features (color, texture, contrast, location)
- - higher-level concepts (faces, people, text, horizon)



Saliency Detection Network

Spatial Manipulation

- We managed to perform Spatial Manipulation combining several methods:
  - **Controlling Z Chunks** - The Z input to bigGAN is combined of 6 chunks where each chunk controls a property of the generated image. We chose to freeze / limit the chunks that control the translations.
  - **Duplicating noise and control regions by mask** - We chose to mask the noise in “W space” in order to achieve spatial changes. Instead of looking at duplicated Z, we choose a “Z map” - specific regions of Z.



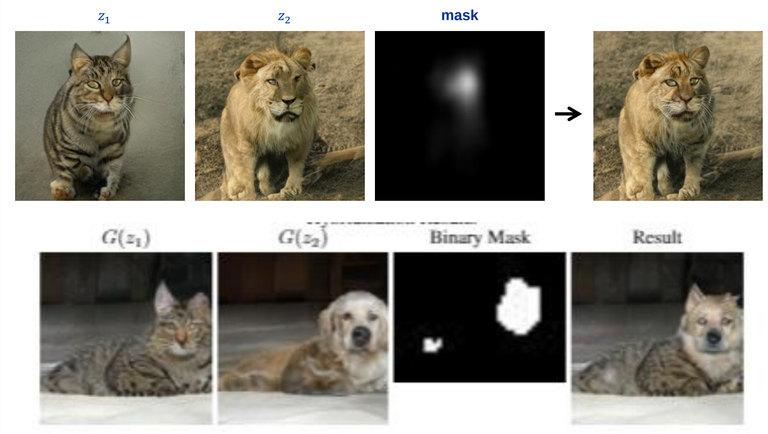
Noise Masking to control different noise to different regions in image

- **Optimizing 2 different weight vectors -** instead of training only 1 vector that changes in 1 part of the mask, we trained 2 weight vectors (and got 2 different Z) for each part of the mask.

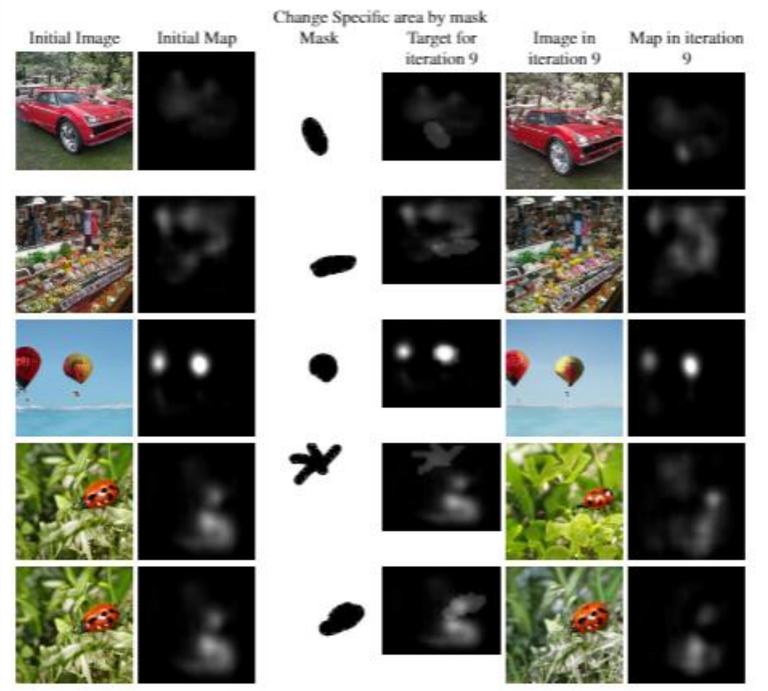
Class Hybridization

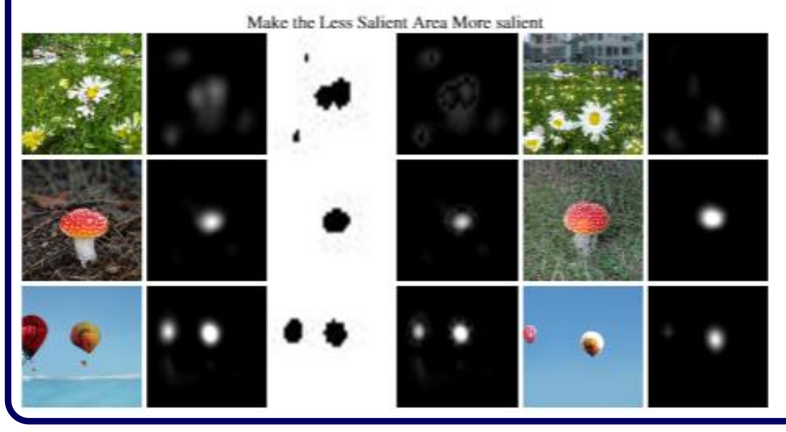
- We create an image that combines two different classes according to a guidance mask.
- The class representation is embedded in the generation process as part of the latent code. That is, the class number is concatenated to z in each of the model's scales. We therefore keep z fixed along all spatial locations and combines only the entries corresponded with the classes.

Class Hybridization Results



Spatial Manipulation Results





Conclusions

- Spatial control over bigGANs can be reached by using separate Z in different spatial regions in the image.
- Saliency manipulation can be done by manipulating salient regions different than other regions.
- One application can be class hybridization, by using same Z with different class in different image regions.

June 2021