

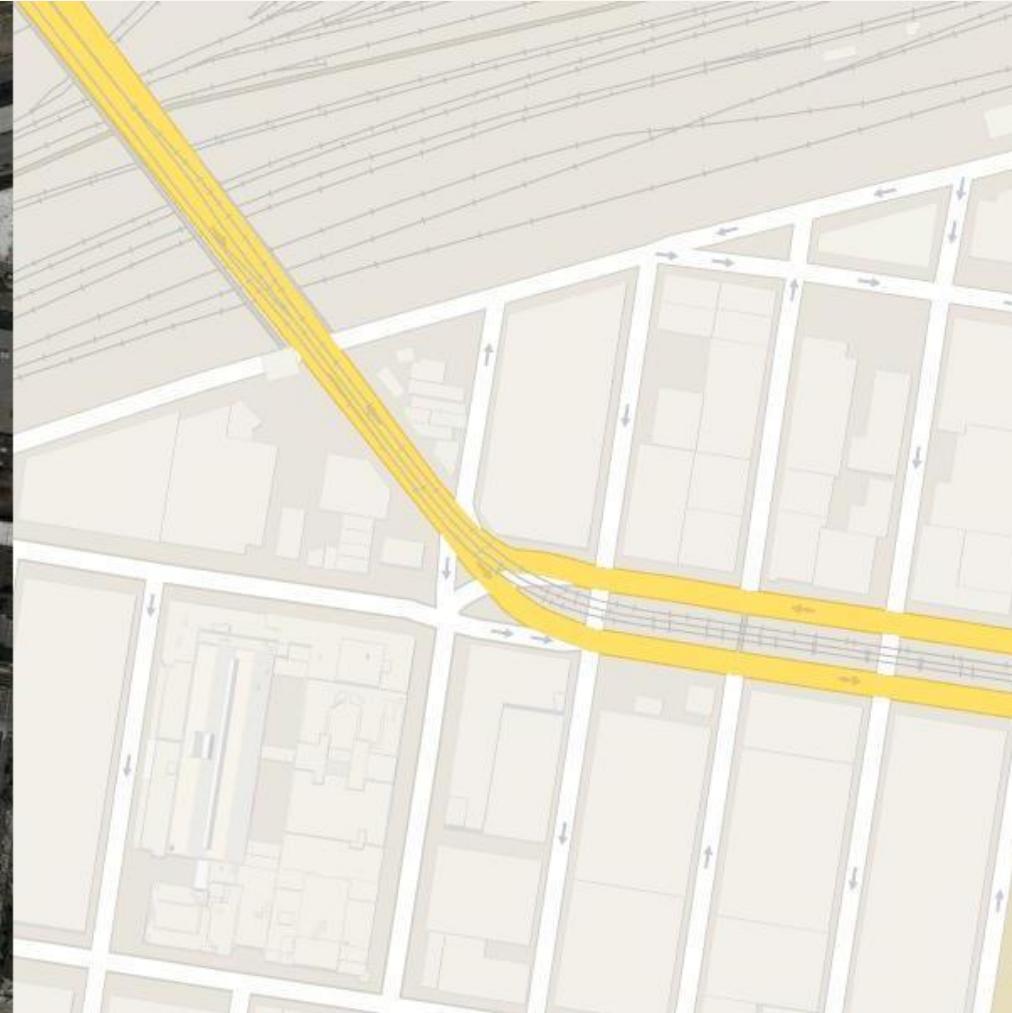
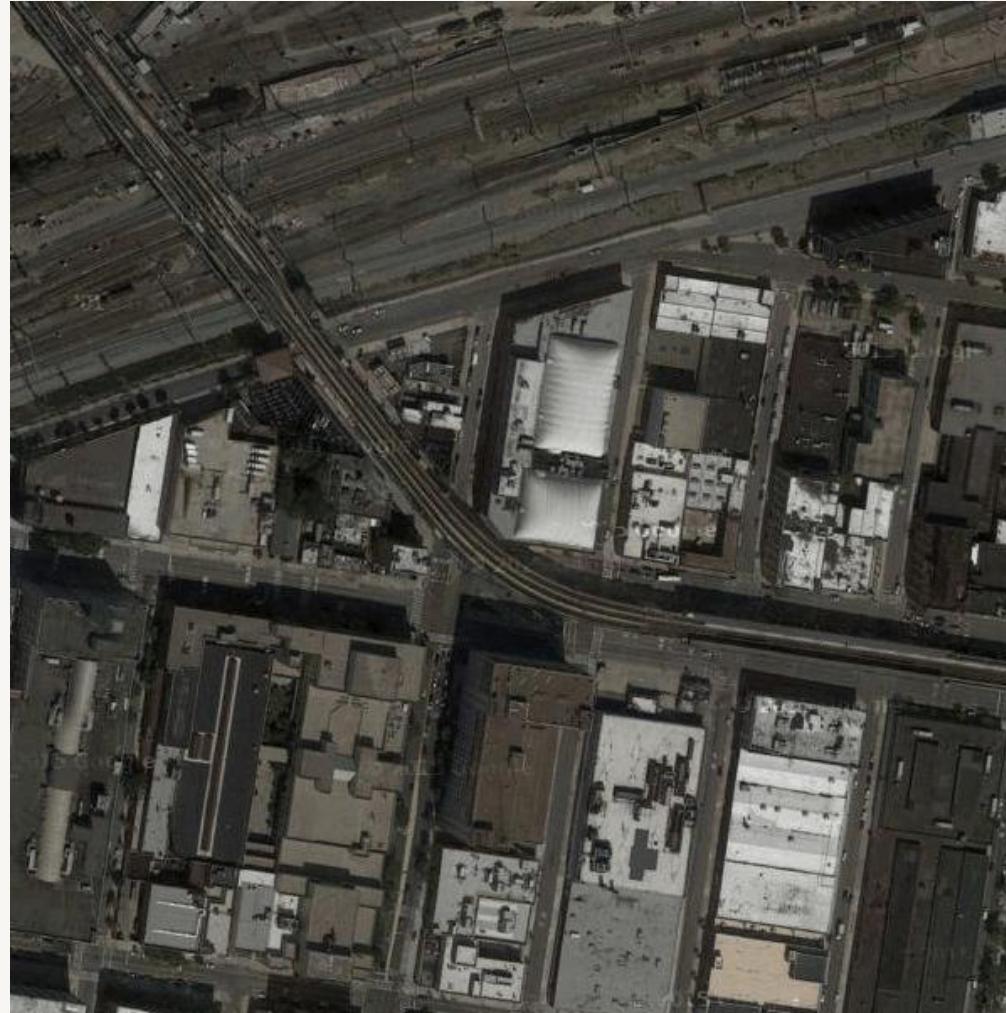


PIX2PIX

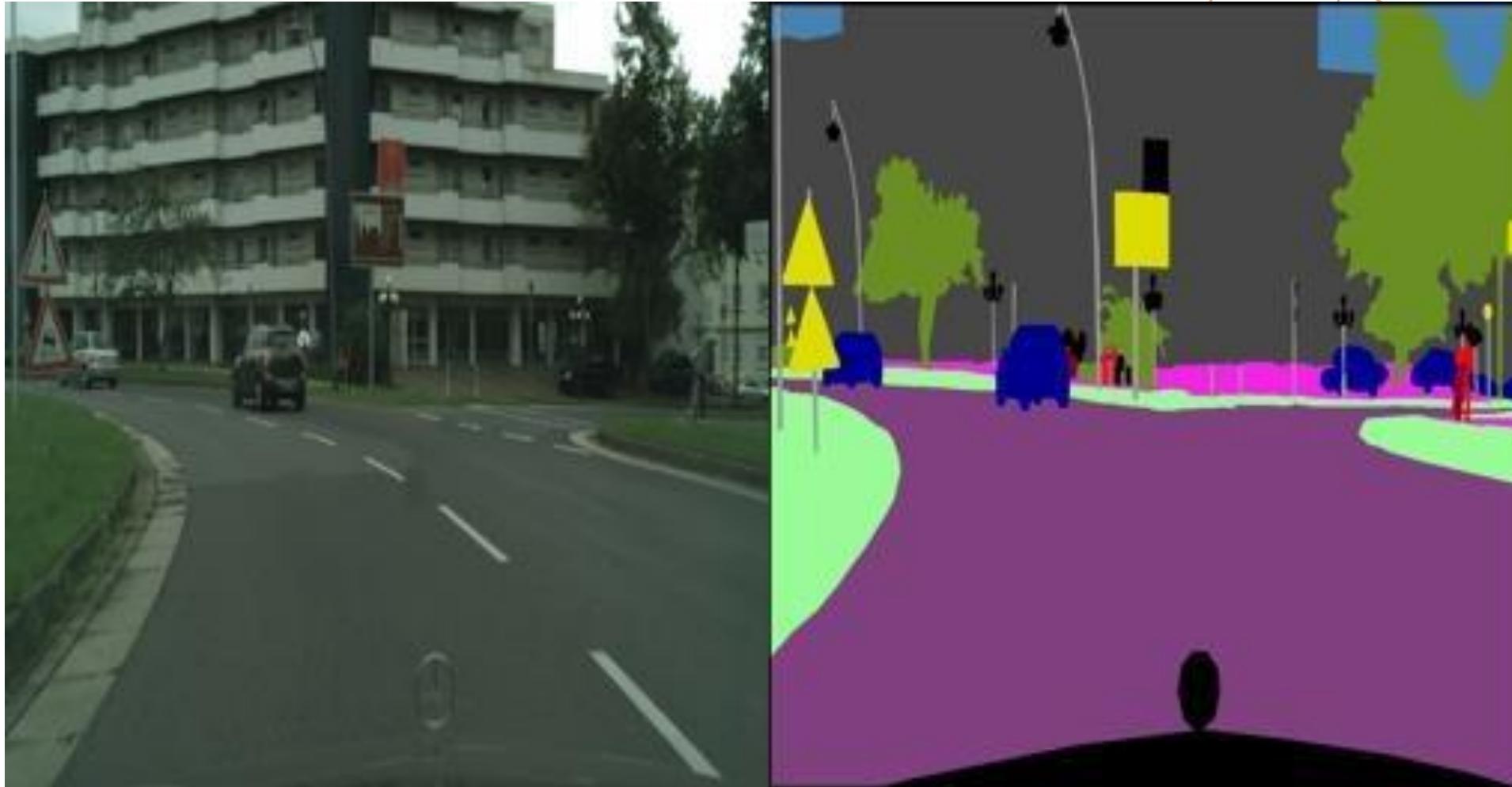
Image to image translation

Stanislaw Lorys

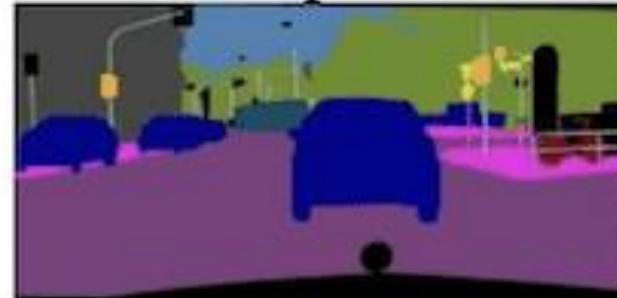
BERKLEY 2018



SEMANTIC SEGMENTATION



Input



Ground truth



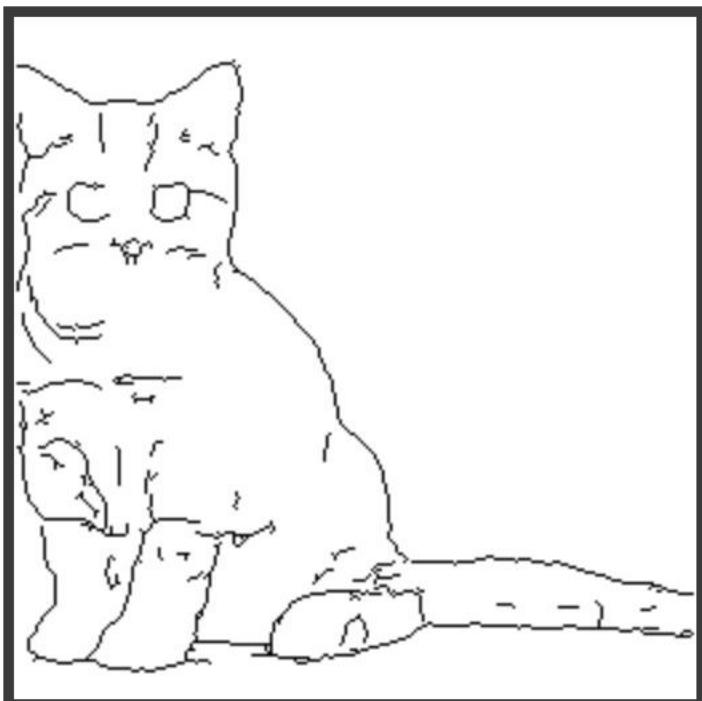
Output



OUTPUT
GENERATION

CAT SKETCH FILLING

INPUT



undo

clear

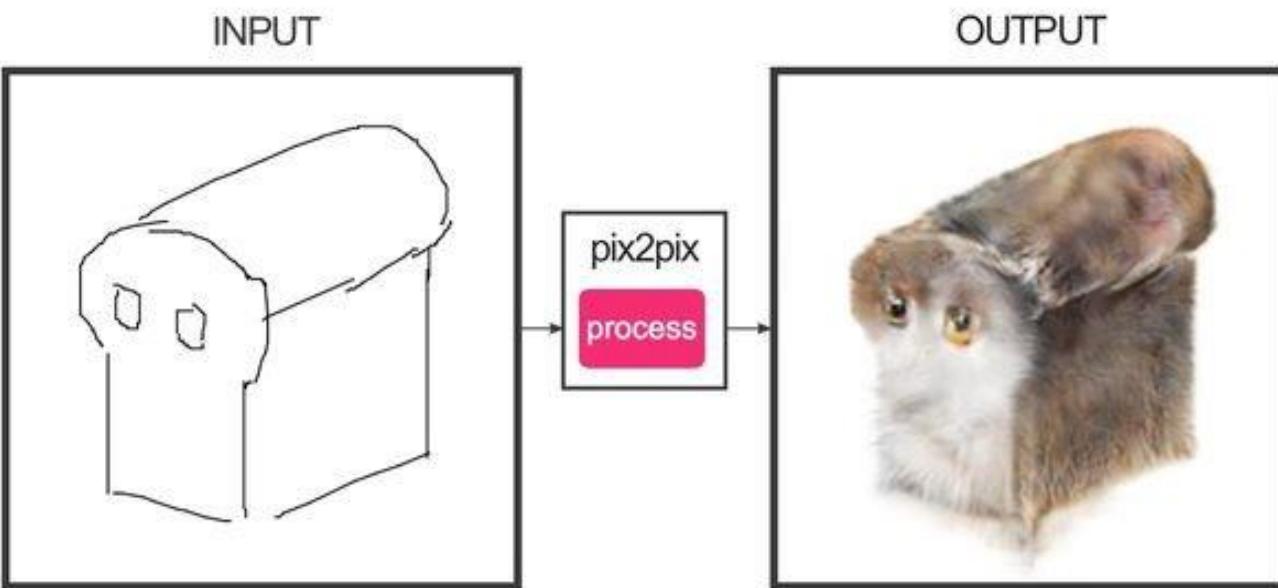
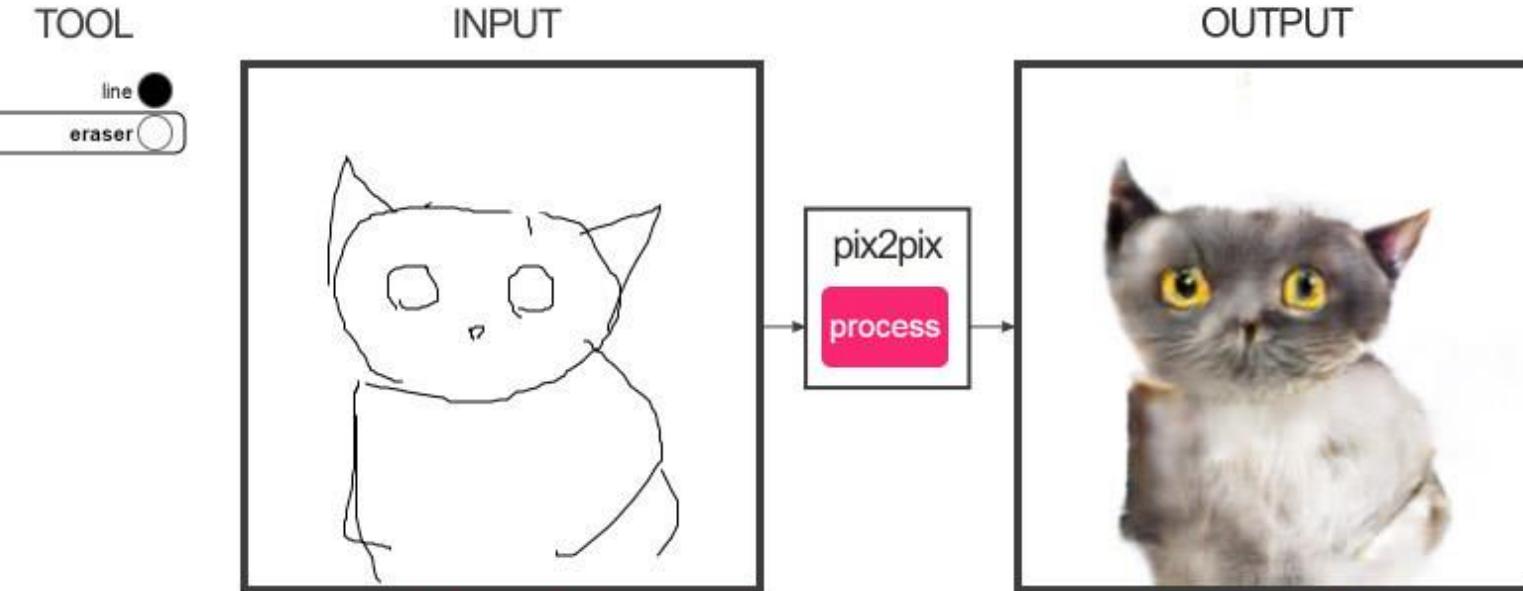
random

OUTPUT

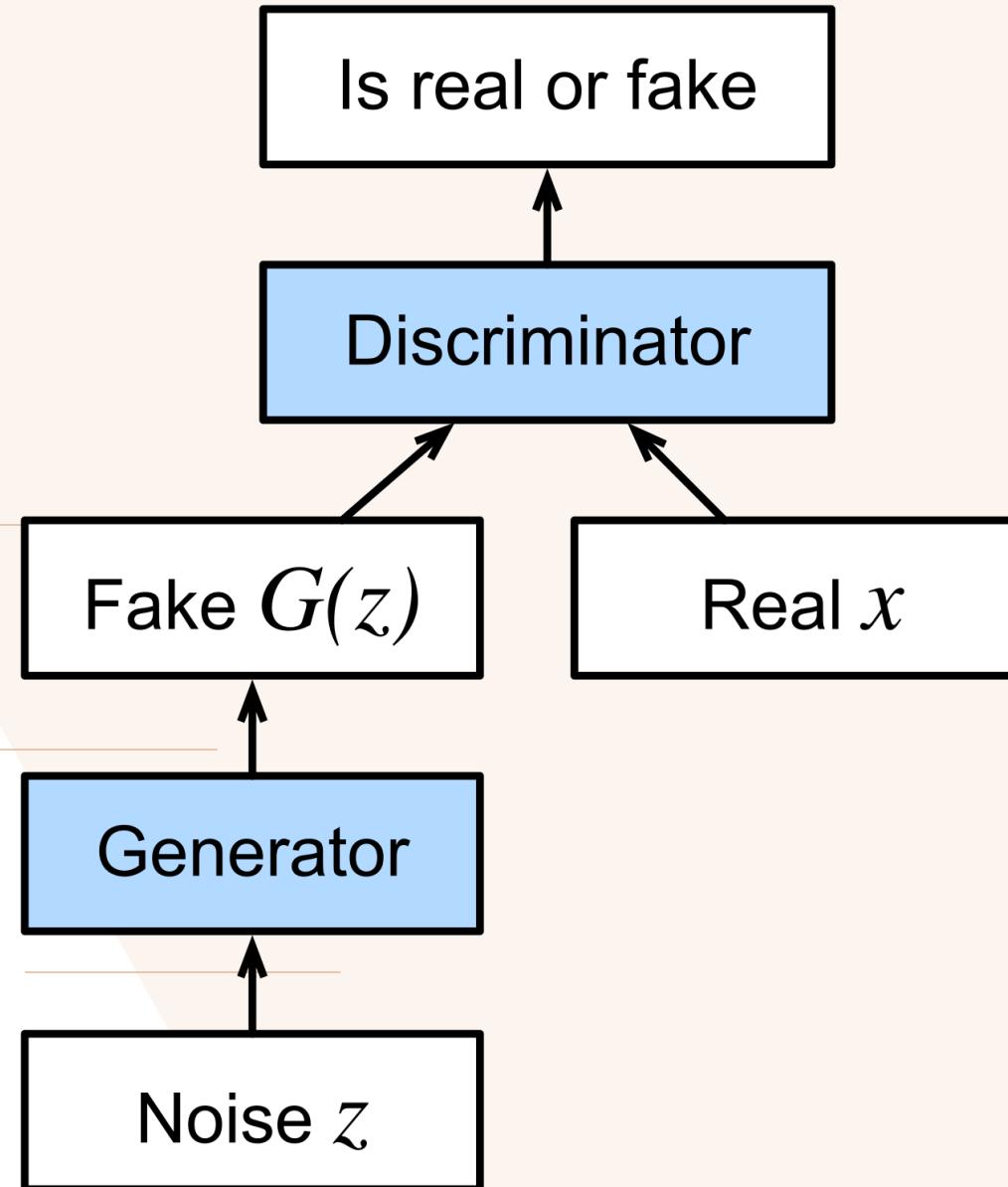


save

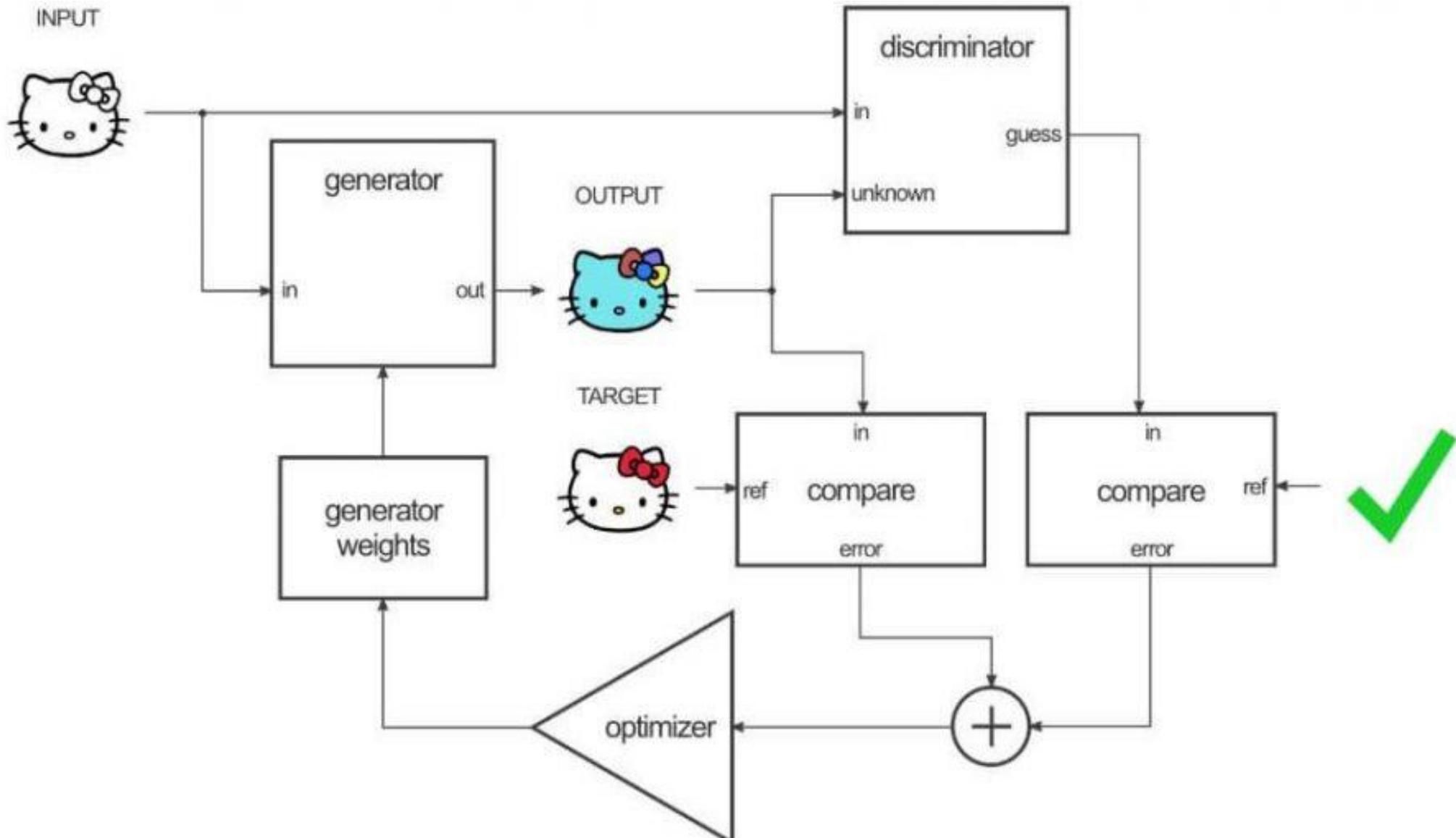
#EDGES2CATS



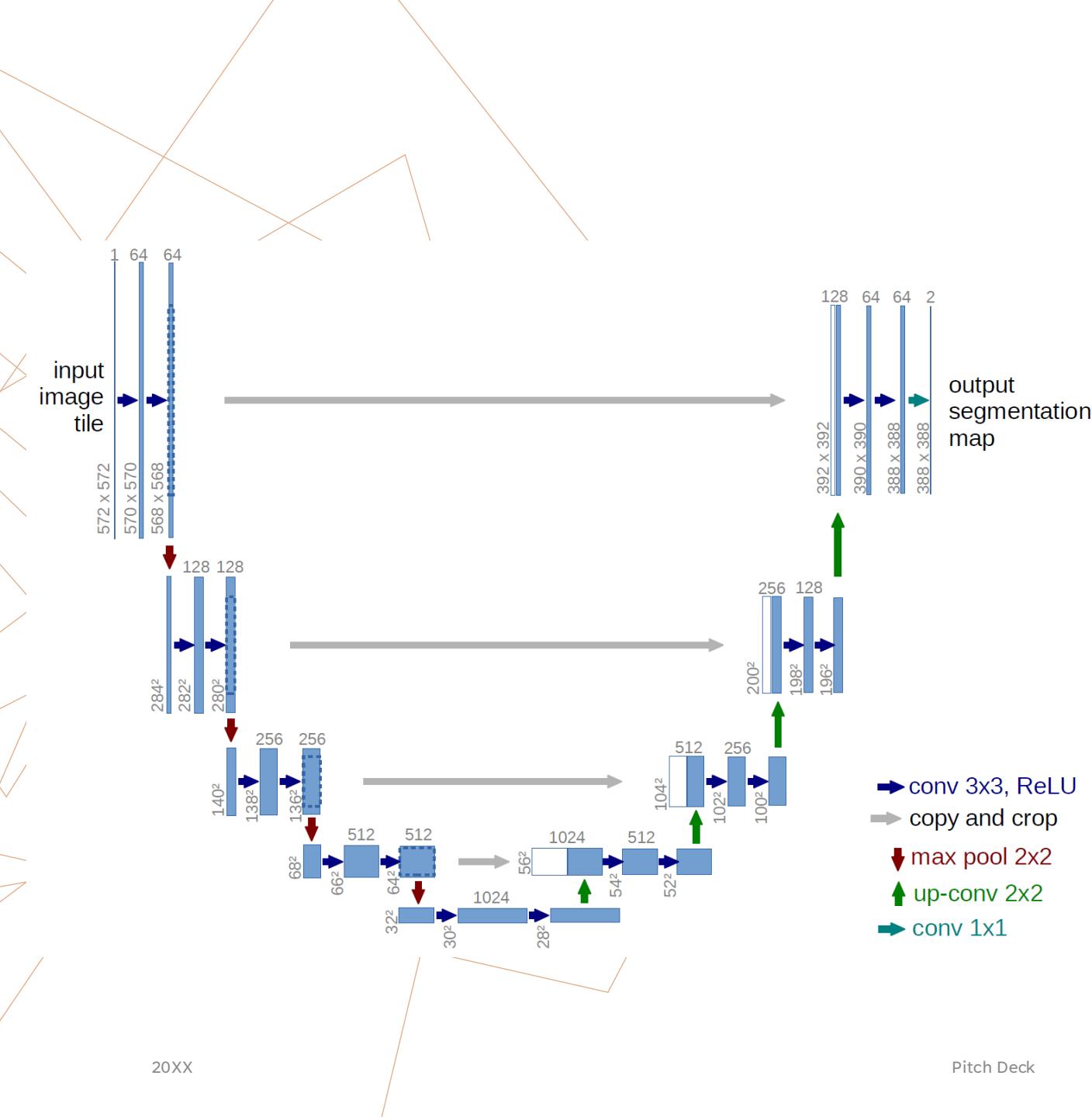
GENERATIVE ADVERSARIAL NETWORK



#EDGES2CATS GAD



GAD ARCHITECTURE



GENERATOR

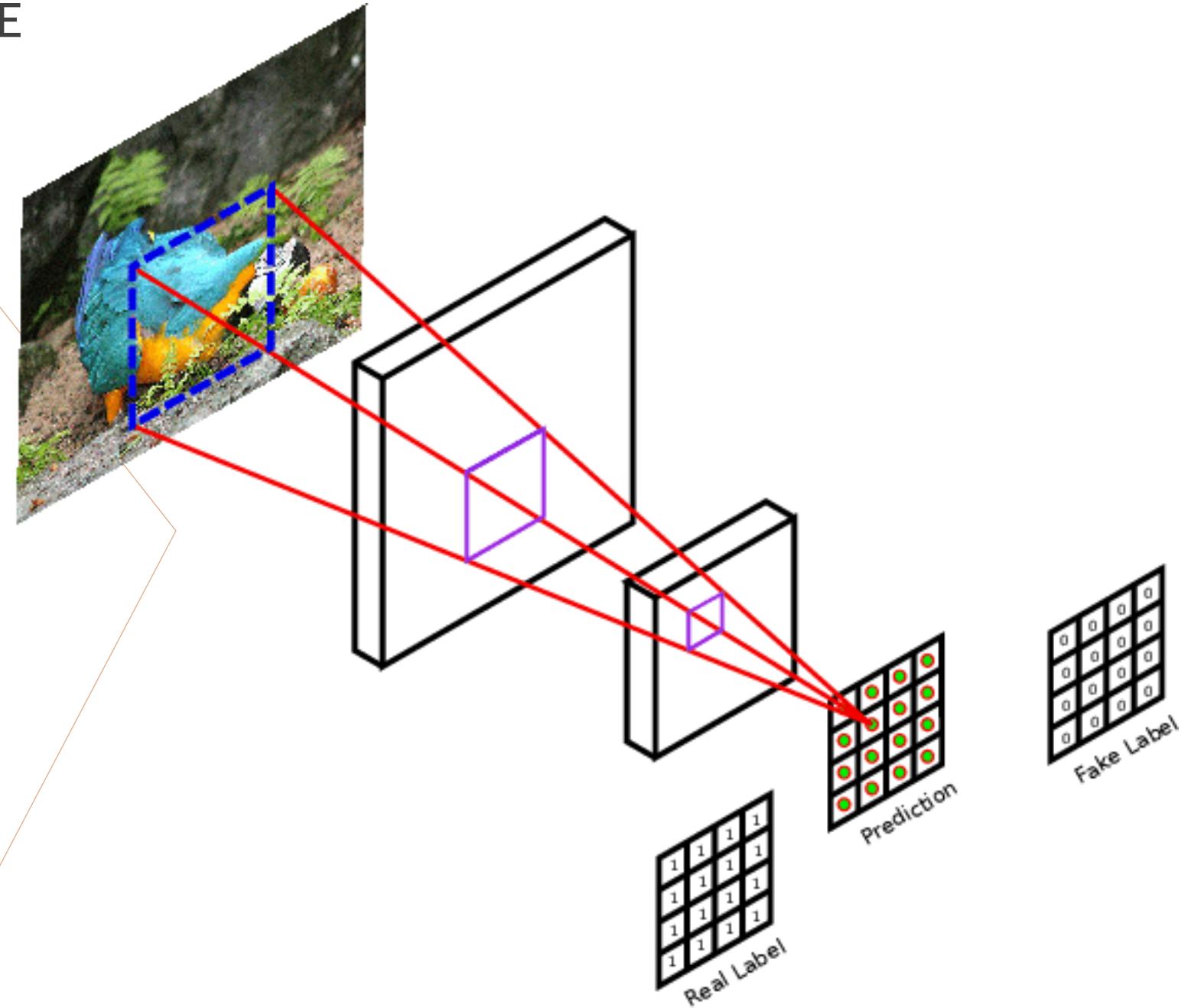
- U-net
- ResNet Generator
- SPADE (Spatially-Adaptive Normalization)
- Transformer-based Generators (e.g. ViT, Swin, DiT)

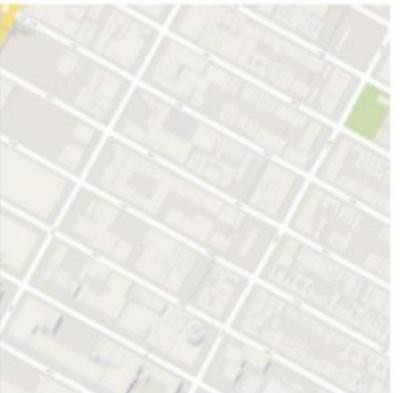
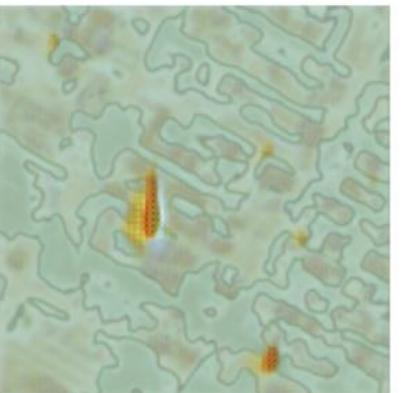
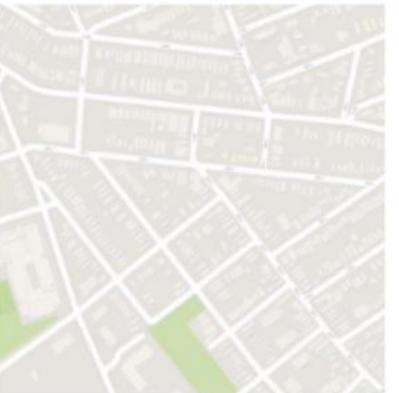
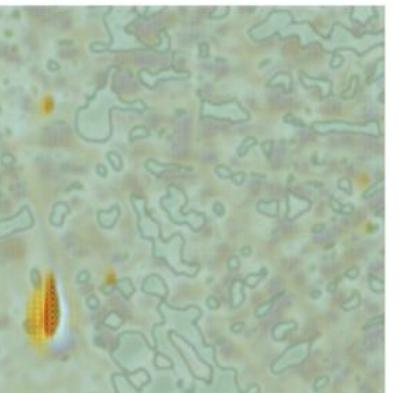
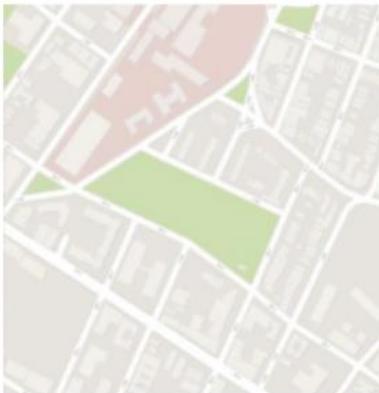
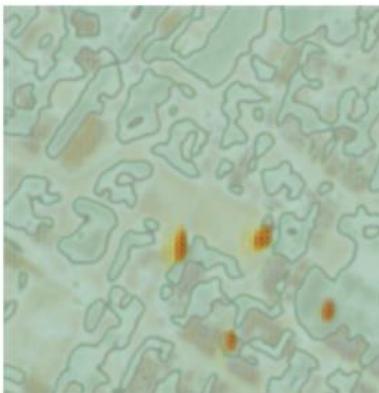
► conv 3x3, ReLU
→ copy and crop
↓ max pool 2x2
↑ up-conv 2x2
► conv 1x1

GAD ARCHITECTURE

DISCRIMINATOR

- PatchGAN
- Multi-Scale PatchGAN
- Feature Matching / Perceptual Loss Additions



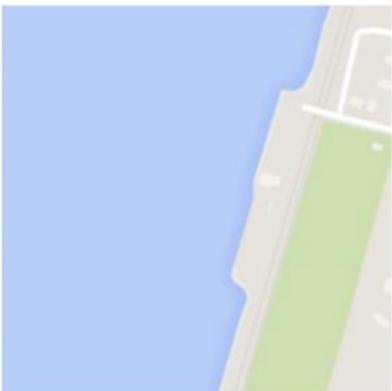
Input**Real****Fake****Input****Real****Fake****Input****Real****Fake**

20XX

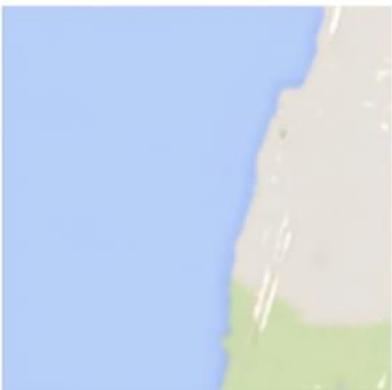
Input



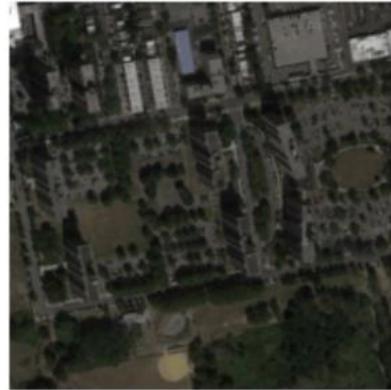
Real



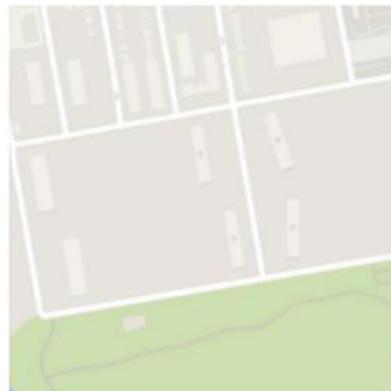
Fake



Input



Real



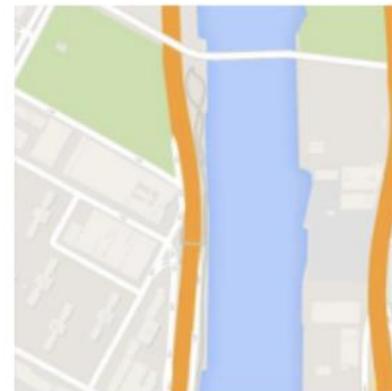
Fake



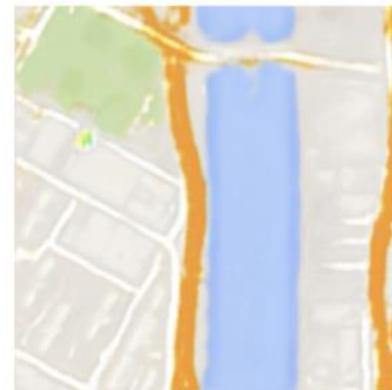
Input



Real



Fake





THANK YOU