Computer Vision

Lecture 6

Machine Learning Decal

Hosted by Machine Learning at Berkeley

Agenda

Computer Vision Introduction

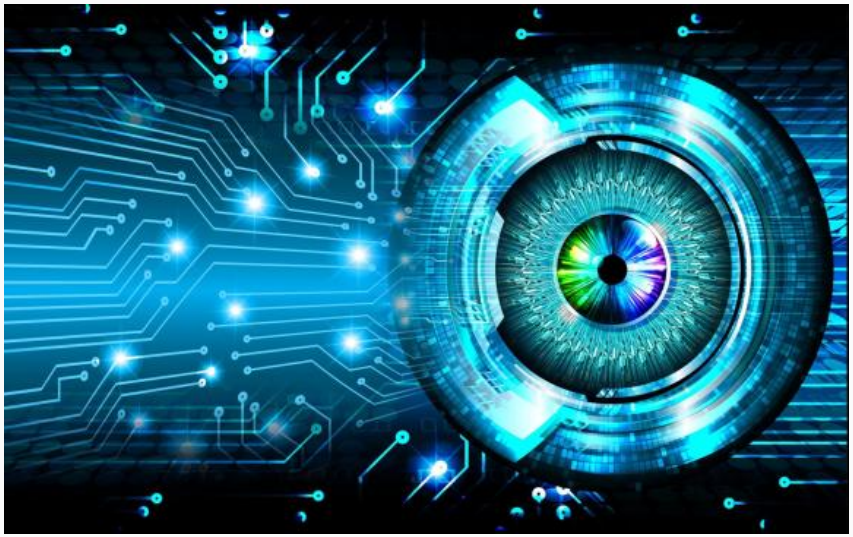
Style Transfer

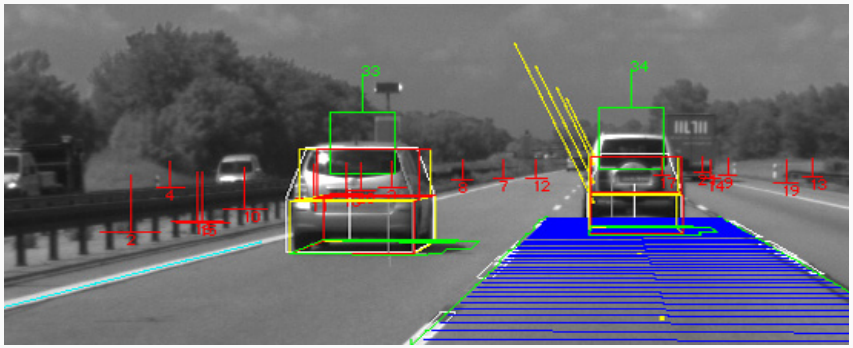
Generative Adversarial Network (GAN)

Questions

Computer Vision Introduction

Computer Vision

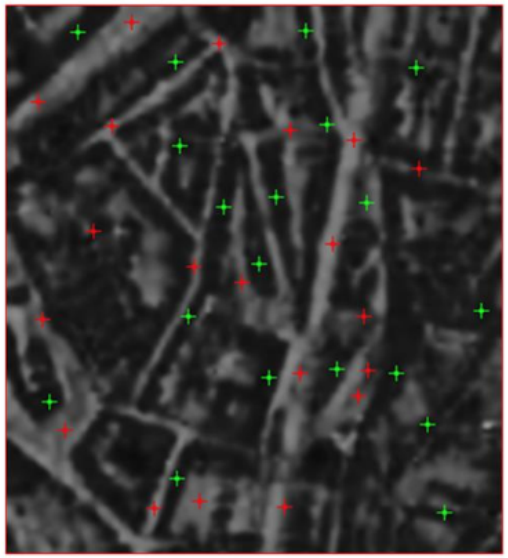
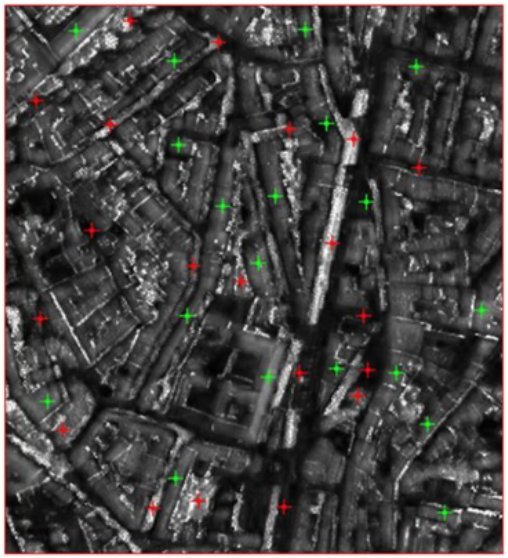




- Sight to gain information about the world
- Make sense of what is happening

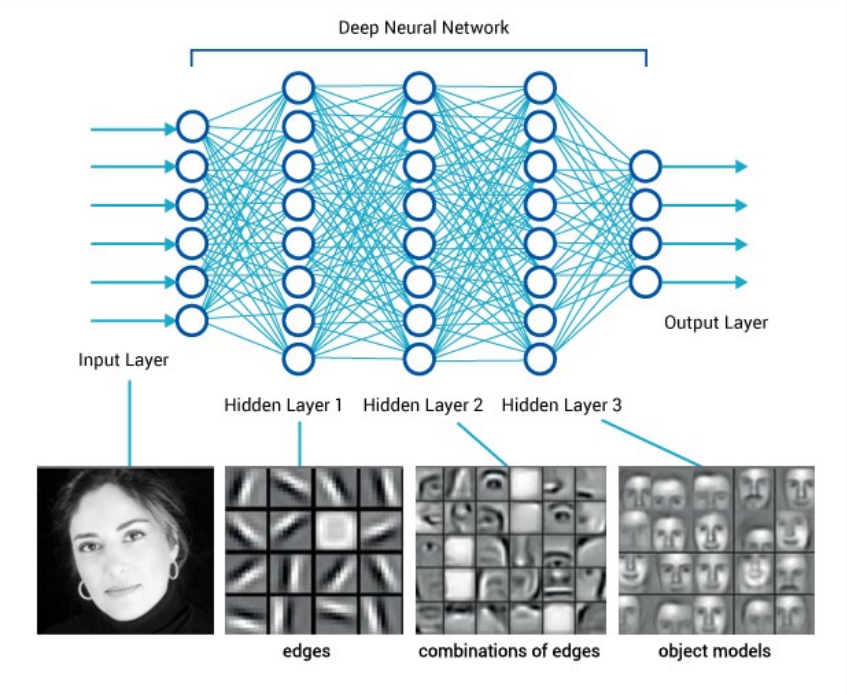
Preprocessing Images

- Scaling
- Cropping
- Background Removal
- Rotations
- Changing pixels (i.e. color to black and white, changing lighting)

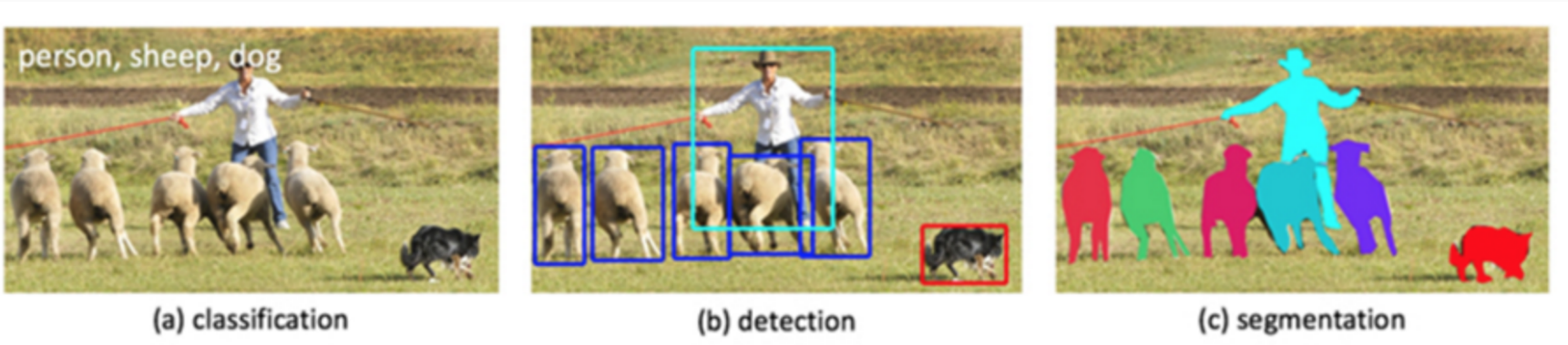




CNN's and CV

- CNN's use less weights, so practical for images which can be $64x64 = 4096$ which would be ridiculous for fully connected layers
- CNN's don't care where, spatially, something happens, for example classification of a dog doesn't care where in the frame the animal is
- CNN's good at learning features, like how first layer can act as an edge detector



Applications



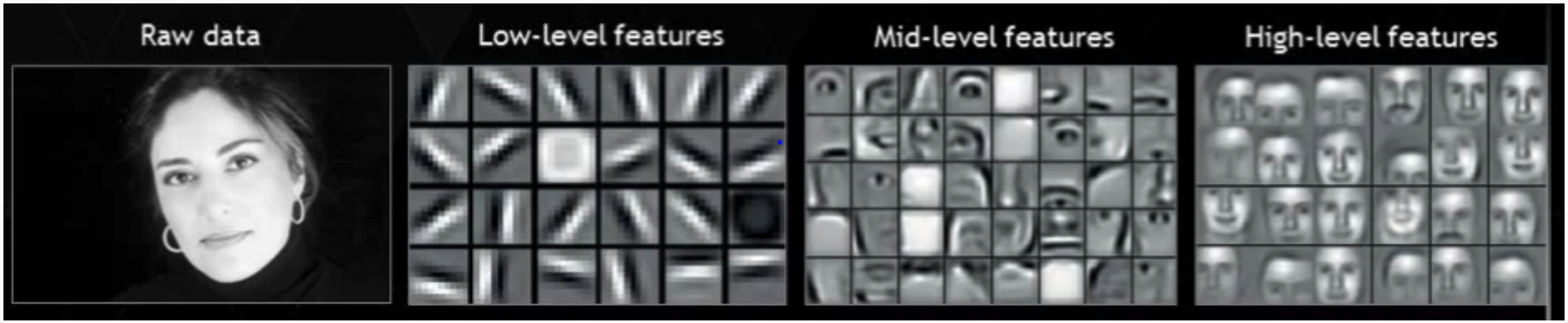
Style Transfer

Style vs Content

- Separate out the style and content of each image
- Content is represented by features derived from the pixels
- Style is represented by Gram matrix

Content

- Represented by features learned by CNN
- Feature loss

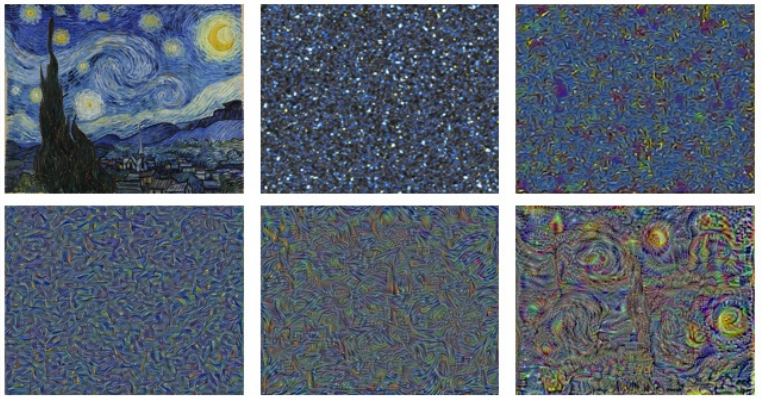


Style

- Gram matrix can be thought of like non-normalized correlation matrix for filters of that layer
- Select multiple layers to calculate so that different sized receptive fields and details are captured in the style

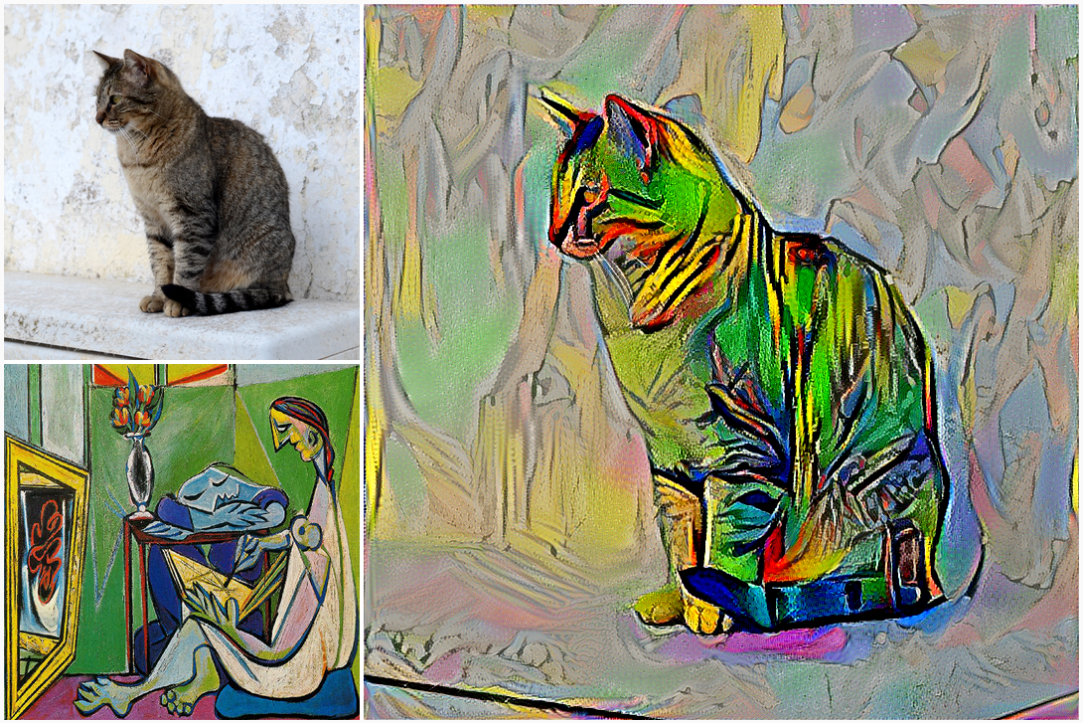
$$G^l_{i,k} = \sum_k F^l_{i,k} F^l_{j,k}$$

Gram Matrix Equation



Loss

- Loss is computed by adding both components of content loss and style loss
- Typically weight with 1 for content, 10000 for style (due to orders of magnitude in Gram matrix vs pixels)



DEMO

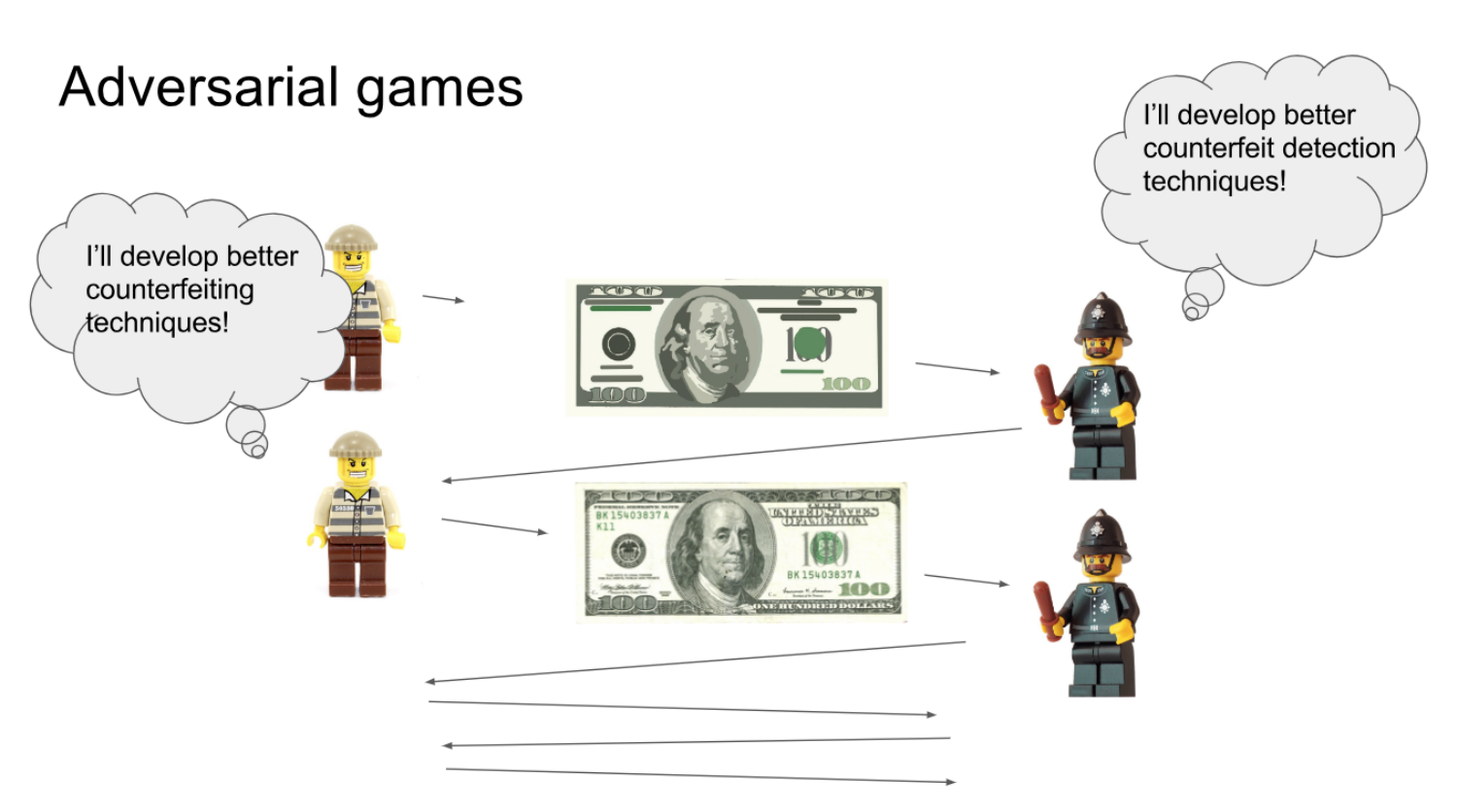
Generative Adversarial Network (GAN)

Motivation

- "What I cannot create, I do not understand" -Richard Feynman
- Ability to generate new, never before seen images

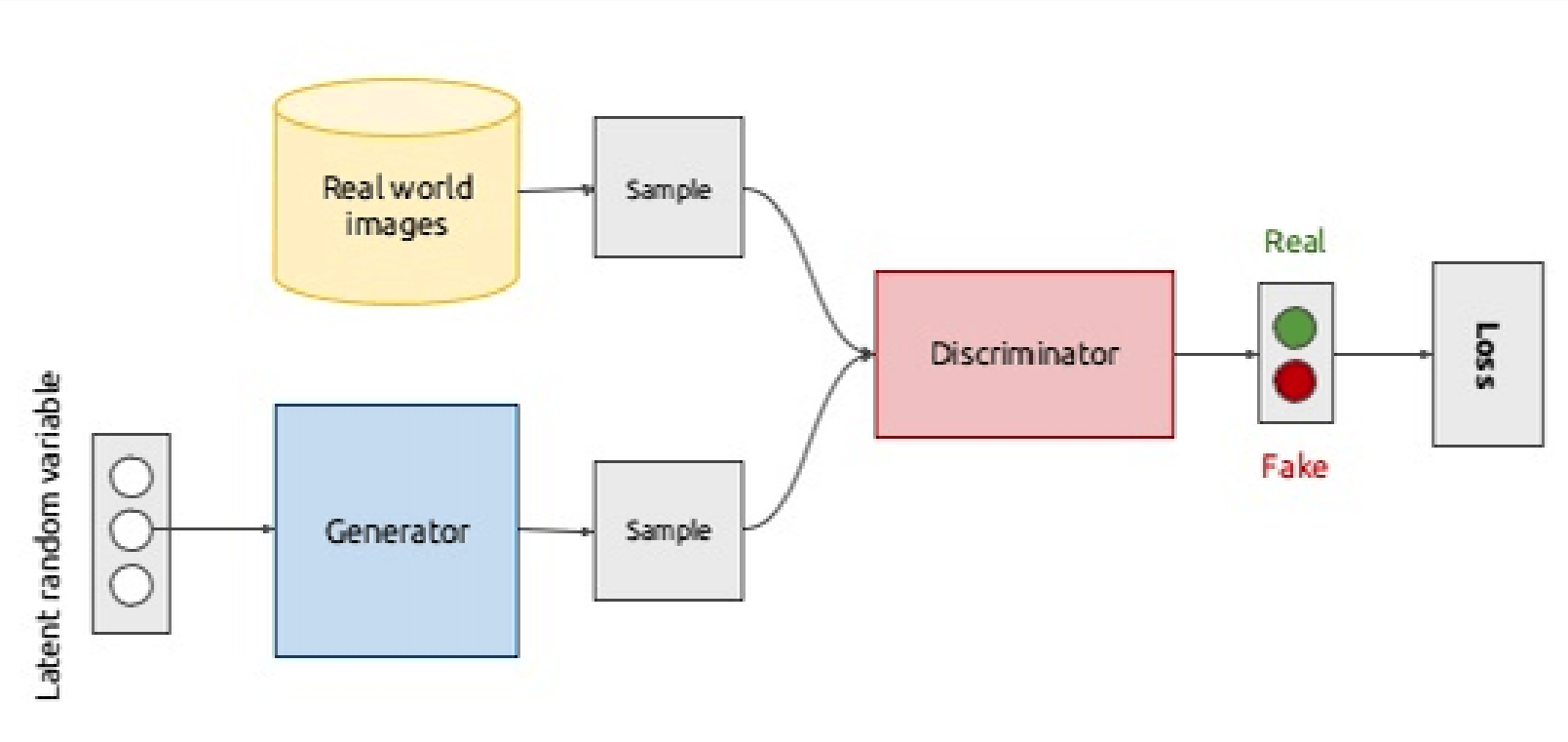
Counterfeit Money Analogy

- Counterfeiter always trying to make better fake currency
- Authorities always getting better at detection
- Both get better by practicing against each other



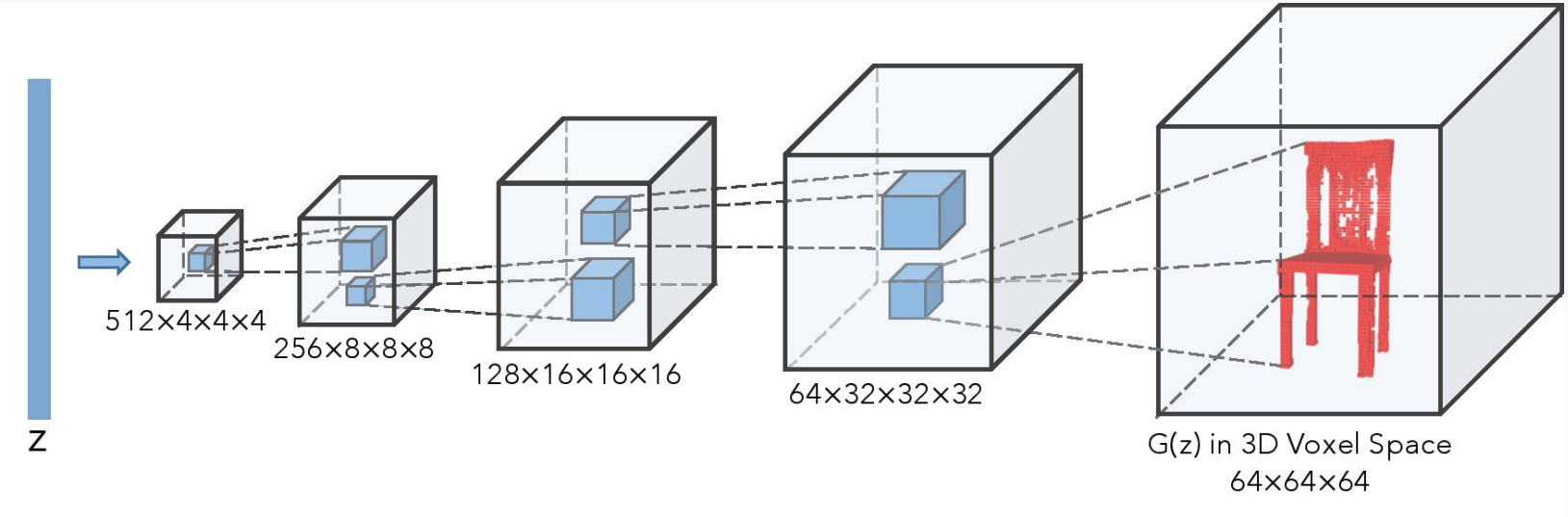
Model Overview

- Generator that generates images
- Discriminator that keeps trying to guess if a generated image is real or fake, while also being trained on real images
- Minimax game



Generator

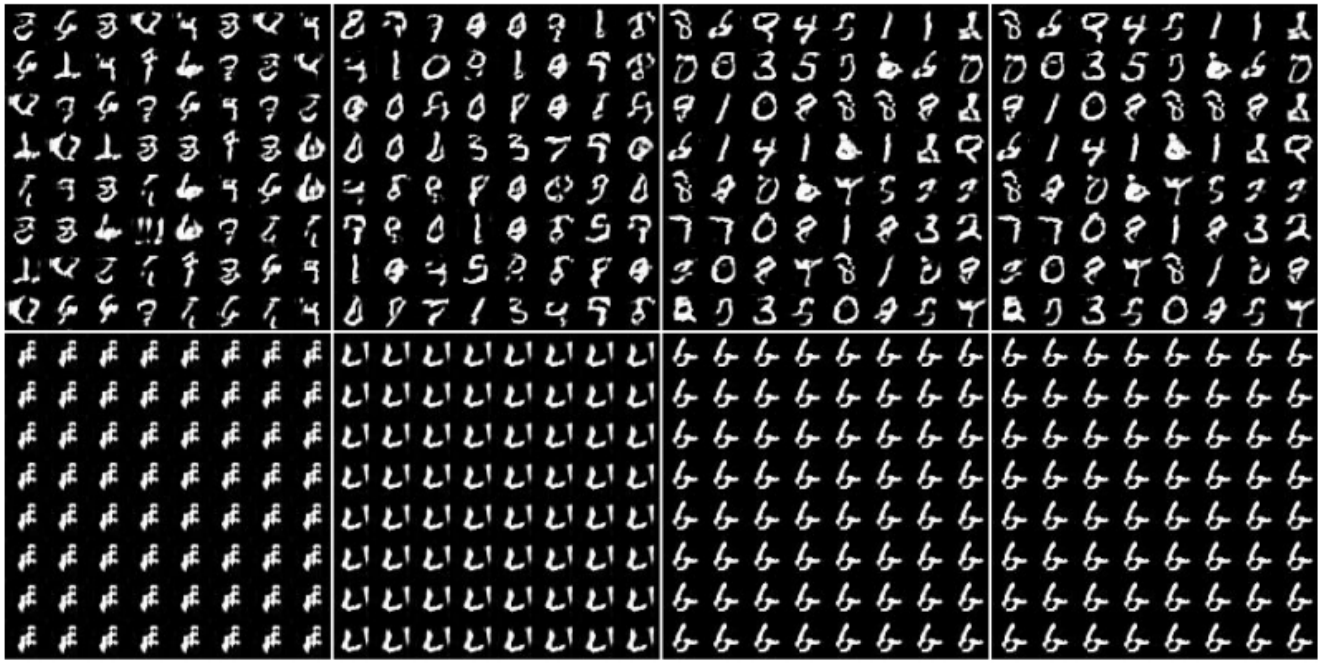
- Takes in latent random variable z as noise that provides stochasticity
- Uses layers called deconvolutions (as opposed to the convolutional layers that we know) to generate



DEMO

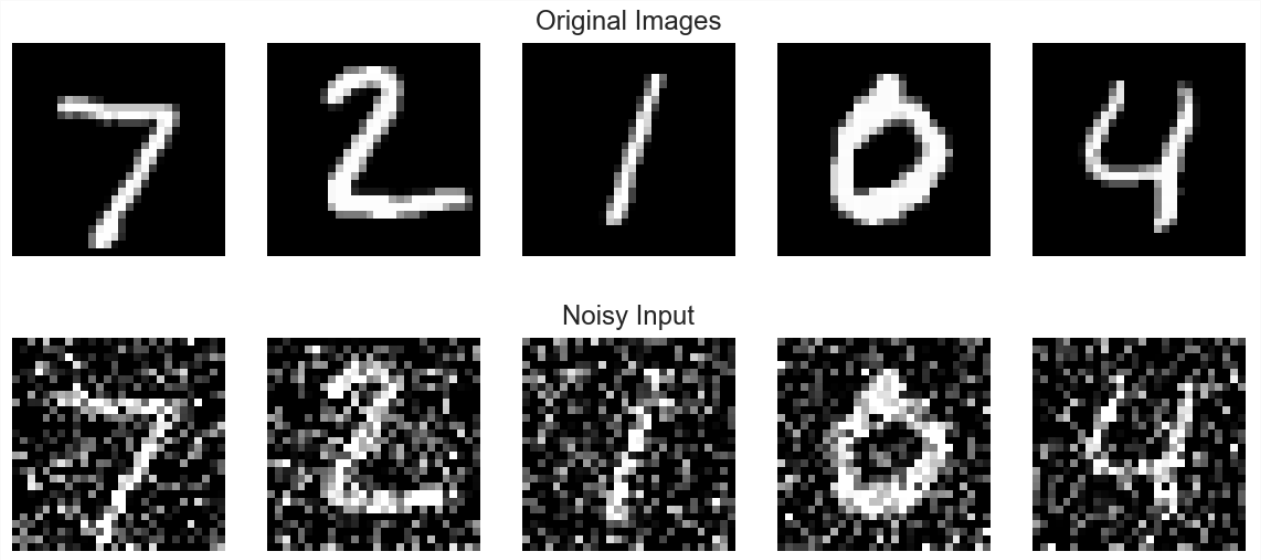
Difficulties and Problems

- Mode collapse
- Hard to tell when convergence is reached and done training
- Random seed has surprisingly big impact



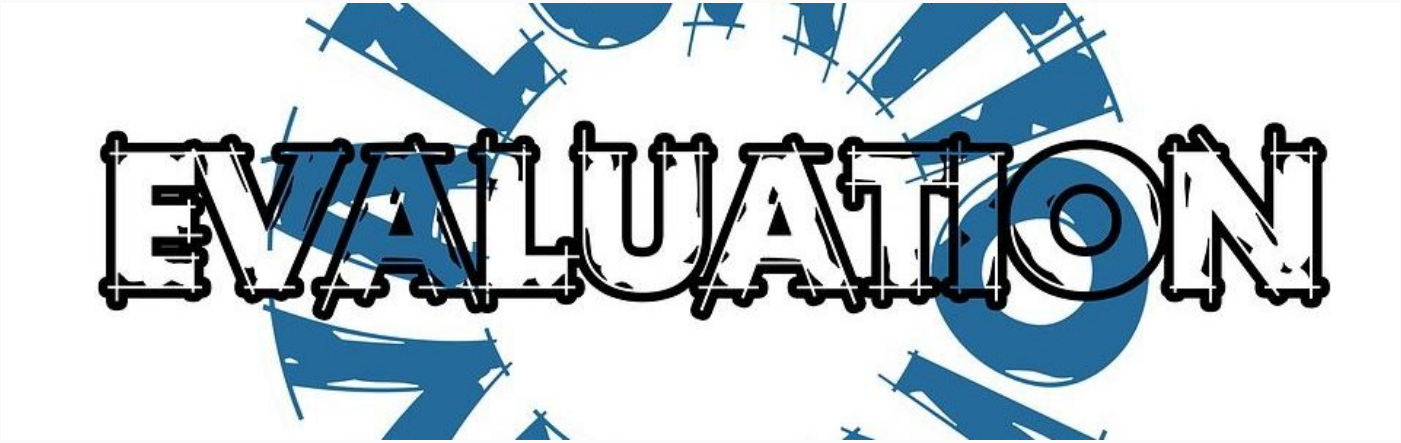
Training Strategies

- Label smoothing (instead of 1 for real, .7-1.0 for real)
- Train discriminator for more steps than generator, or vice versa
- Adding noise to real input (makes it less "perfect" and easier to replicate)



Quantitative Evaluation

- Generative Adversarial Metric (GAM) is $\frac{p(y=1|x;D_1)p(x;G_2)}{p(y=1|x;D_2)p(x;G_1)}$
- Inception Score uses Inception v3 Network pre-trained on ImageNet by calculating a statistic based on the KL divergence (look up more details if interested)



Qualitative Evaluation

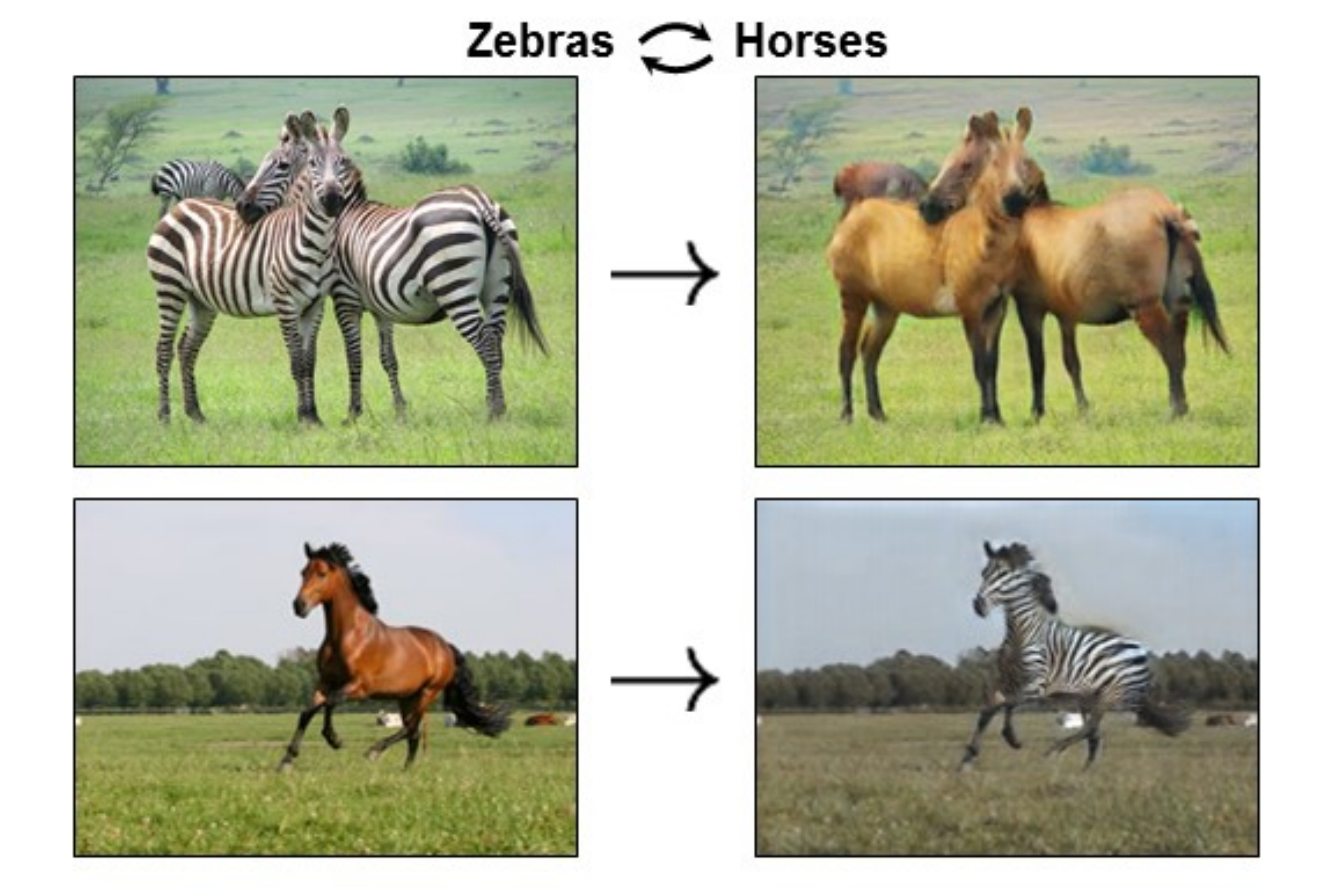
- Just look

Ethical Concerns

- Deep Fakes
- We have trouble trusting news, but now even photos and videos can be fake too?

Extra Readings for those Interested

- Wasserstein GAN
- CycleGAN
- SAGAN



Questions

Questions?