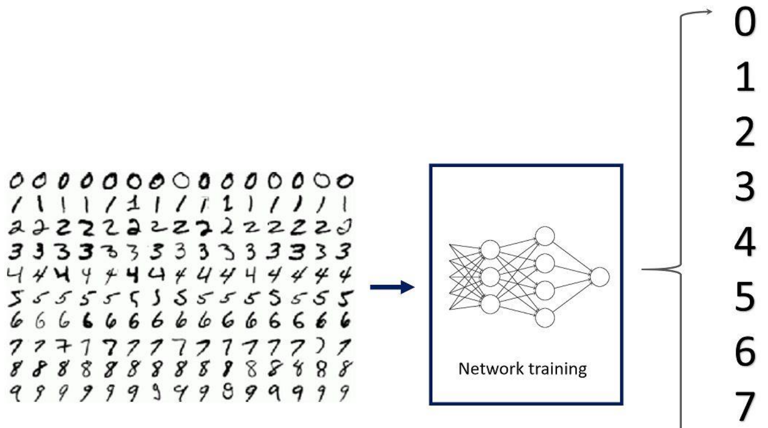
Generative Methods

GAN (Generative Adversarial Networks)

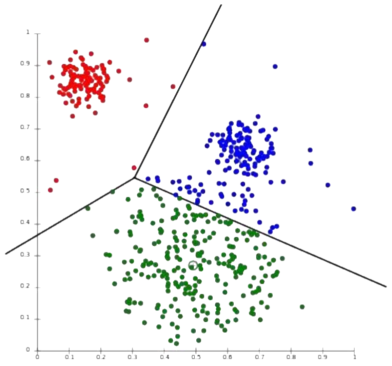
Supervised vs Unsupervised Learning

- Supervised - data + label
  - classification, regression, detection etc.
  - learning a function to map x to y



Data & Labels

- Unsupervised - Just data + no labels
  - clustering, dimensionality reduction, feature learning etc.
  - Learn some underlying hidden structure of the data



K-means clustering

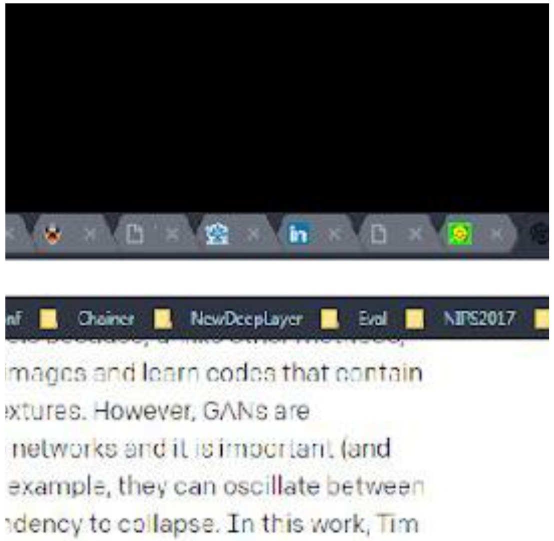
Generative modeling

- Given training data, generate new samples from the same distribution

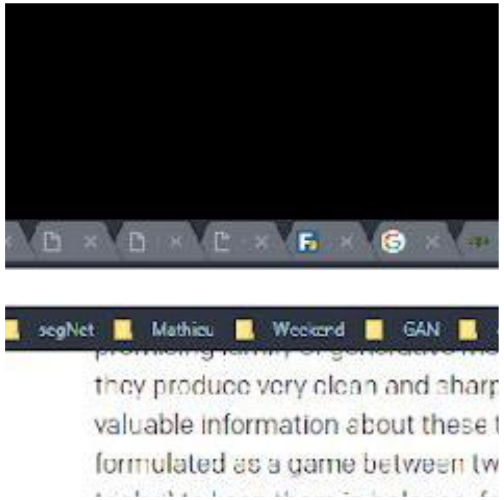
The 60 False-Positive pairs (1.00%) on LFW by DeepFace-ensemble.

Generative modeling

- Given training data, generate new samples from the same distribution



Training data ~ $p_{data}(x)$

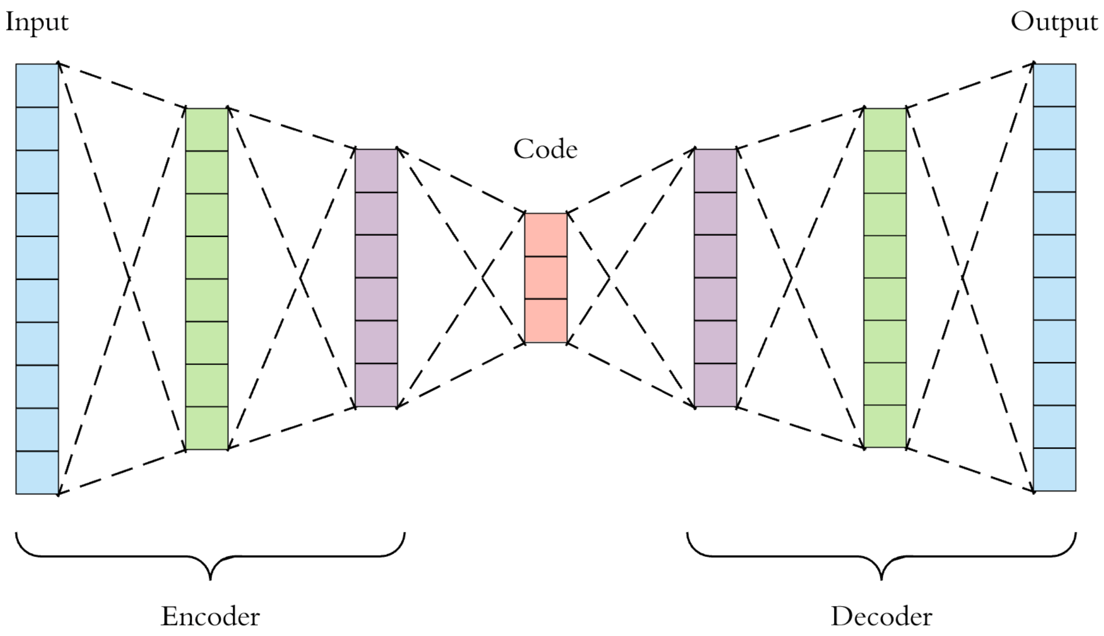


Generated samples ~ $p_{model}(x)$

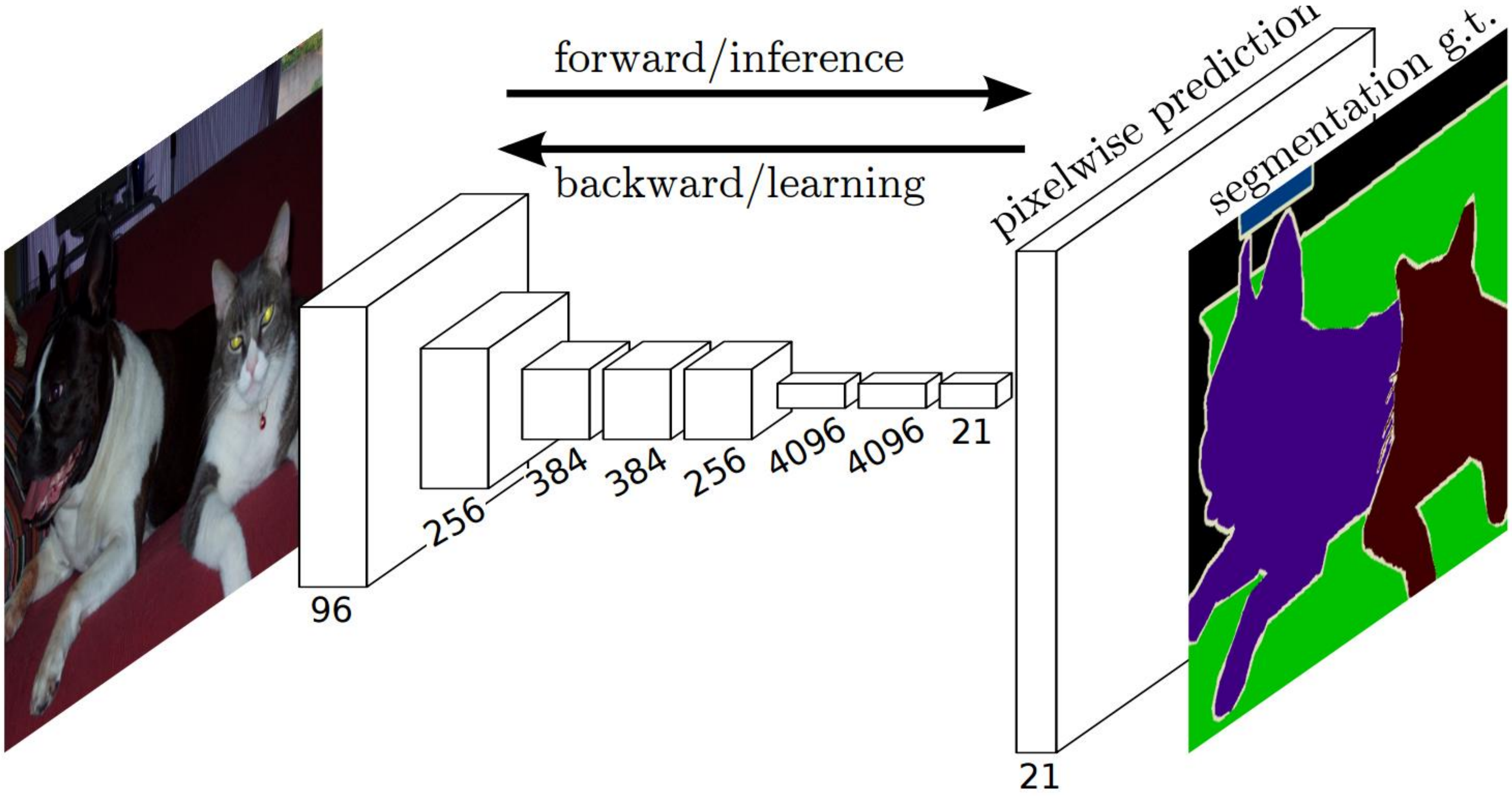
Want to: learn $p_{model}(x)$ similar to $p_{data}(x)$

Background - Auto encoders

- Encoder Decoder architecture
  - Use for compression
  - Or throw away the decoder after training use the compressed code for supervised learning tasks (particularly useful in problems with small training set)



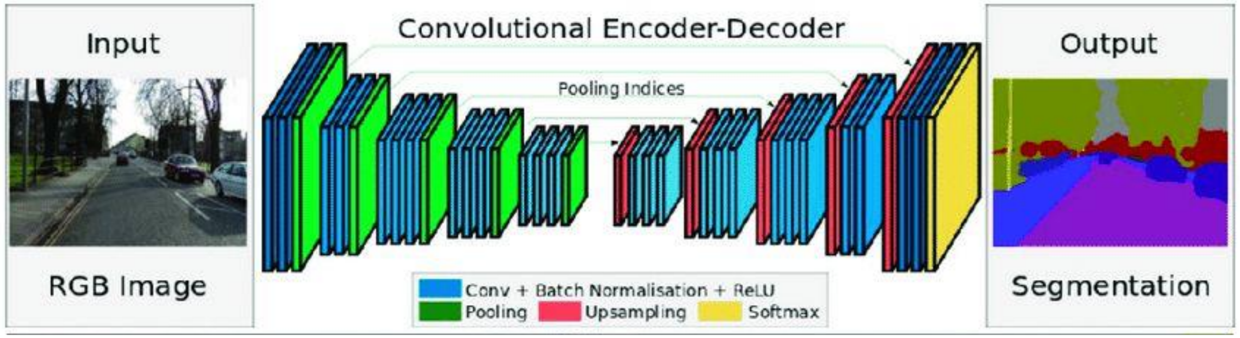
Background – Image to Image



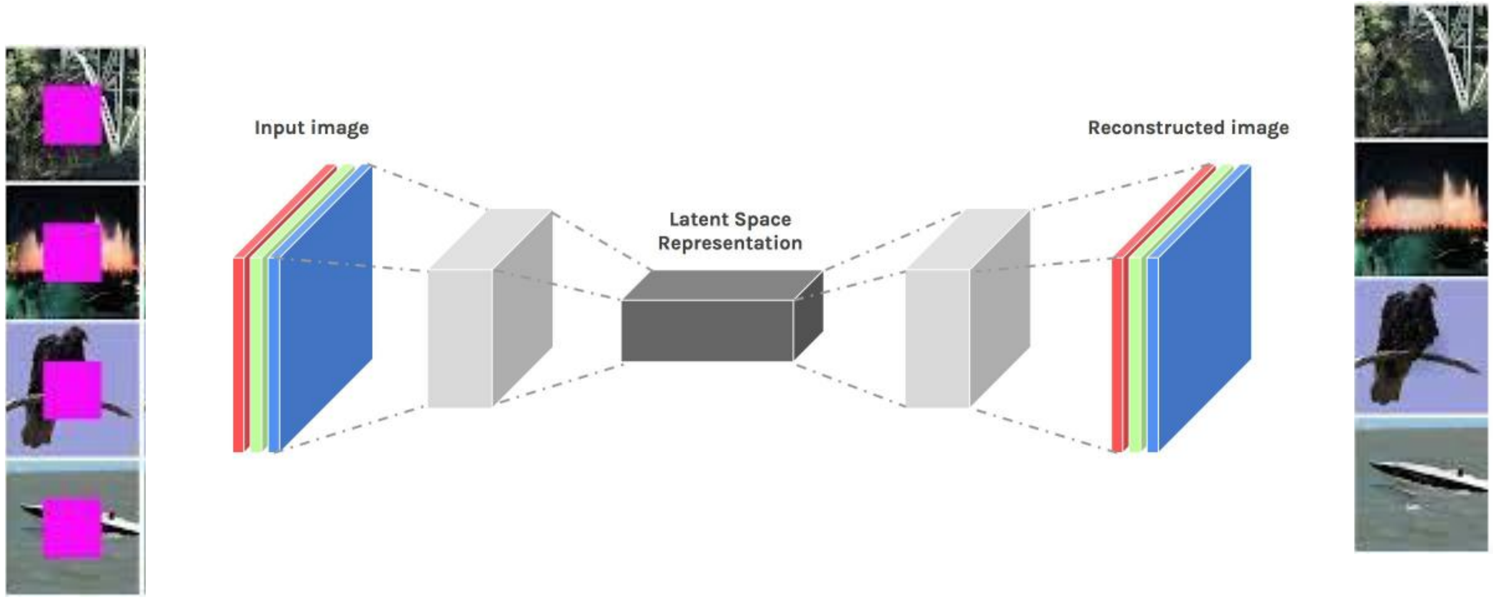
https://arxiv.org/pdf/1411.4038.pdf

Generation

Semantic Segmentation



Neural Inpainting



Generative Models

Increasing Role of Synthetic Data

- Code to synthesis “Real data” ;
  - Role in data augmentation
  - Capability to generate for Neural Networks

Can we detect generated?

- The “story” of Police and “fake notes”

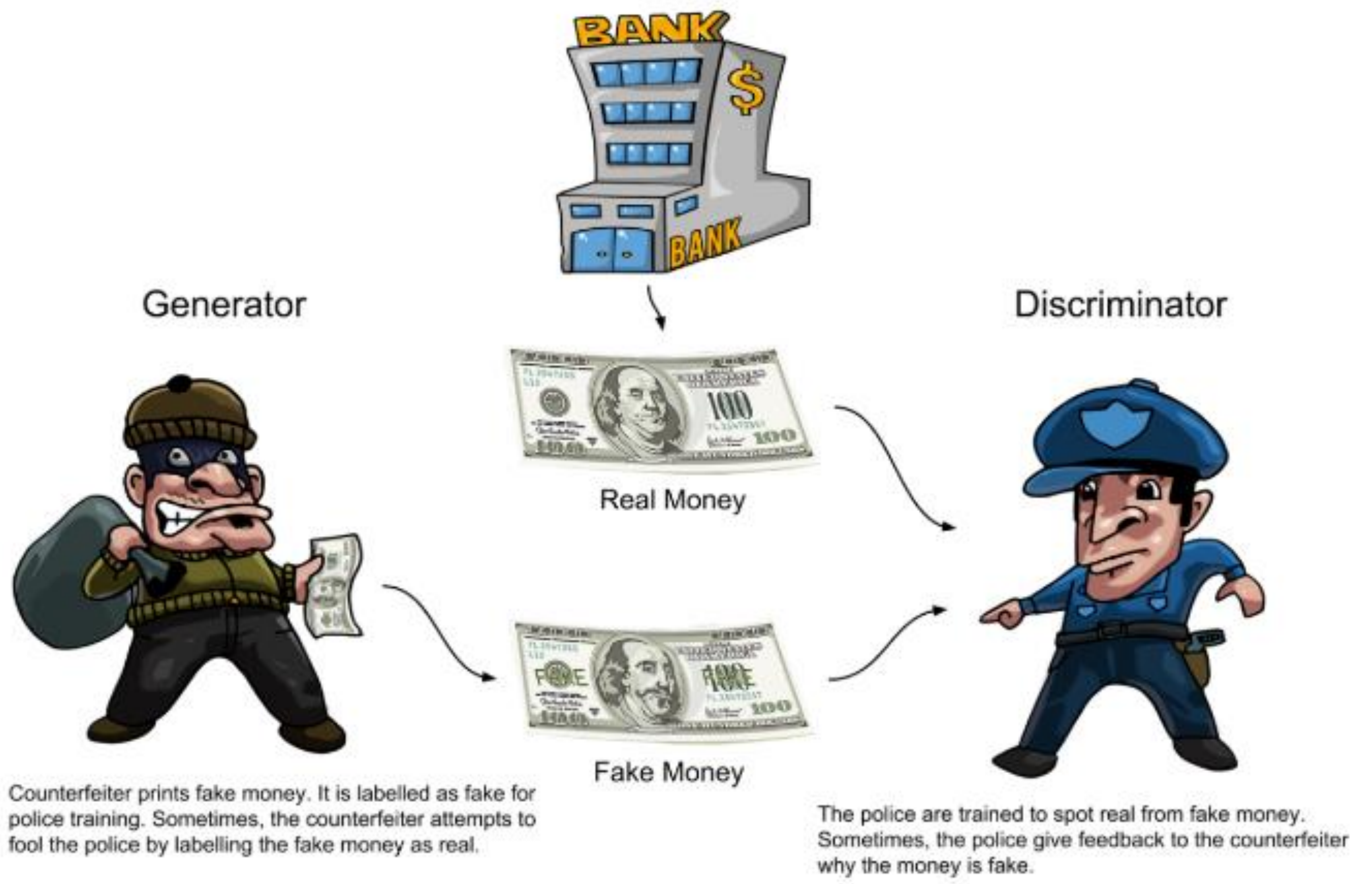
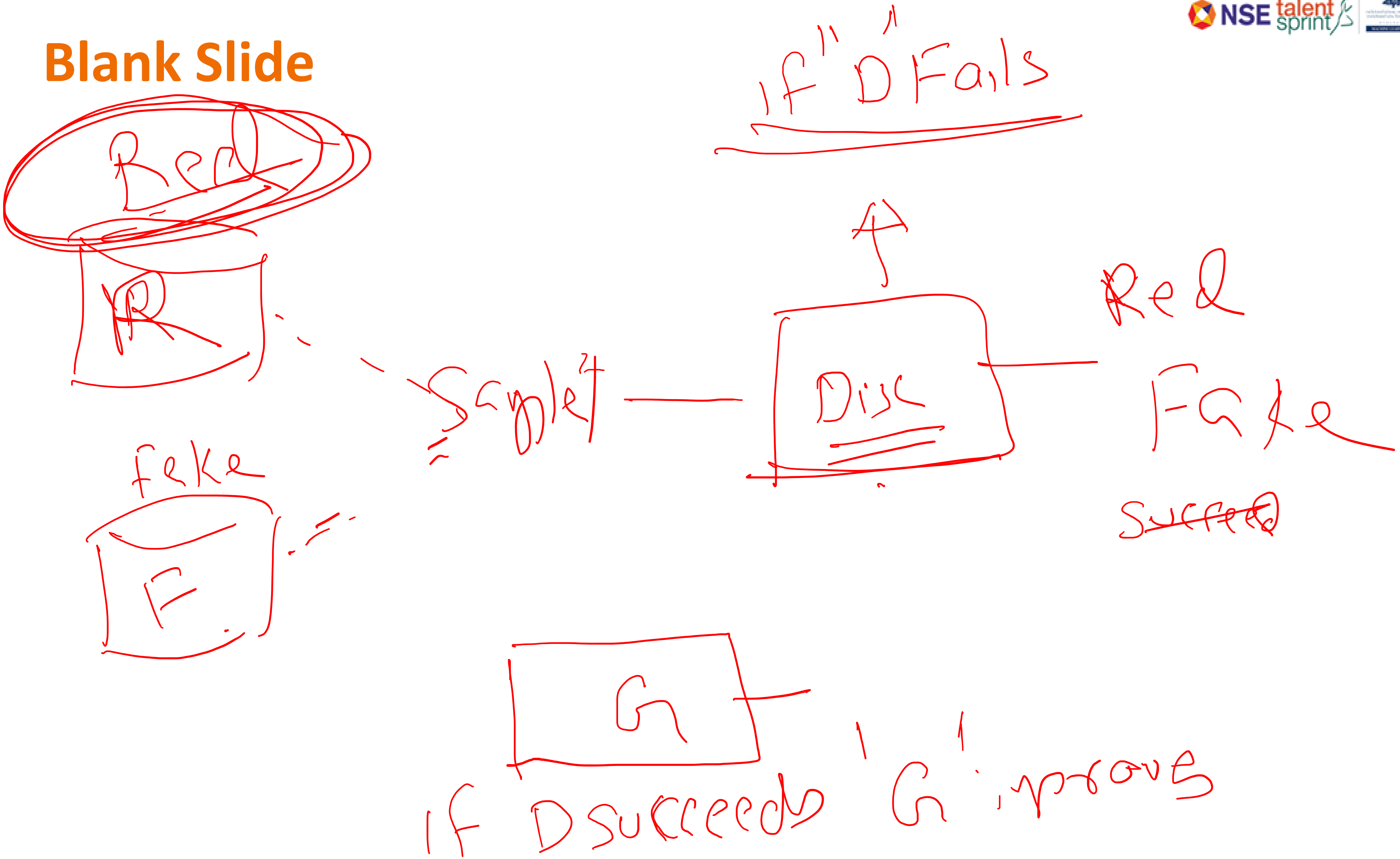


Image: Courtesy to Richard Gall.

Blank Slide

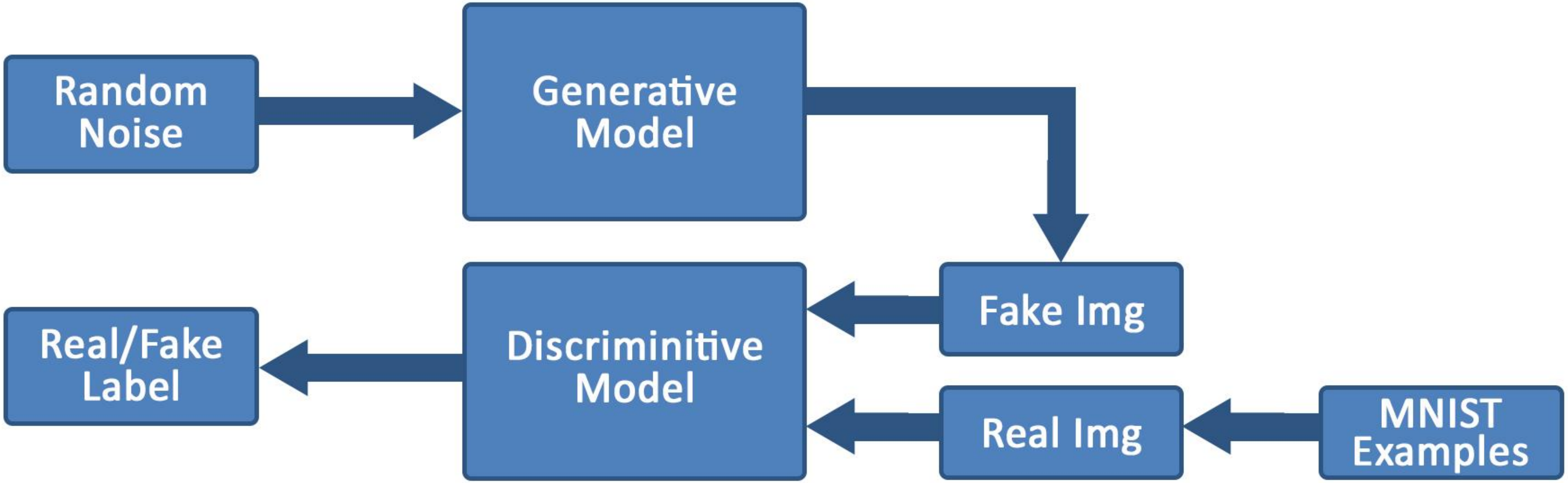
Blank Slide


If "D" Fails
Real
R
fake
F
Sample
Disc
Real
Fake
~~Succeed~~
G
If D succeeds 'G' improve

GANs

- Generative Adversarial Networks
- A game between two players:
  - Discriminator D
  - Generator G
- D tries to discriminate between:
  - A sample from the data distribution
  - A sample from the generator G
- G tries to "trick" D by generating samples that are hard for D to distinguish from data.

GAN: Network architecture



GAN: Network architecture

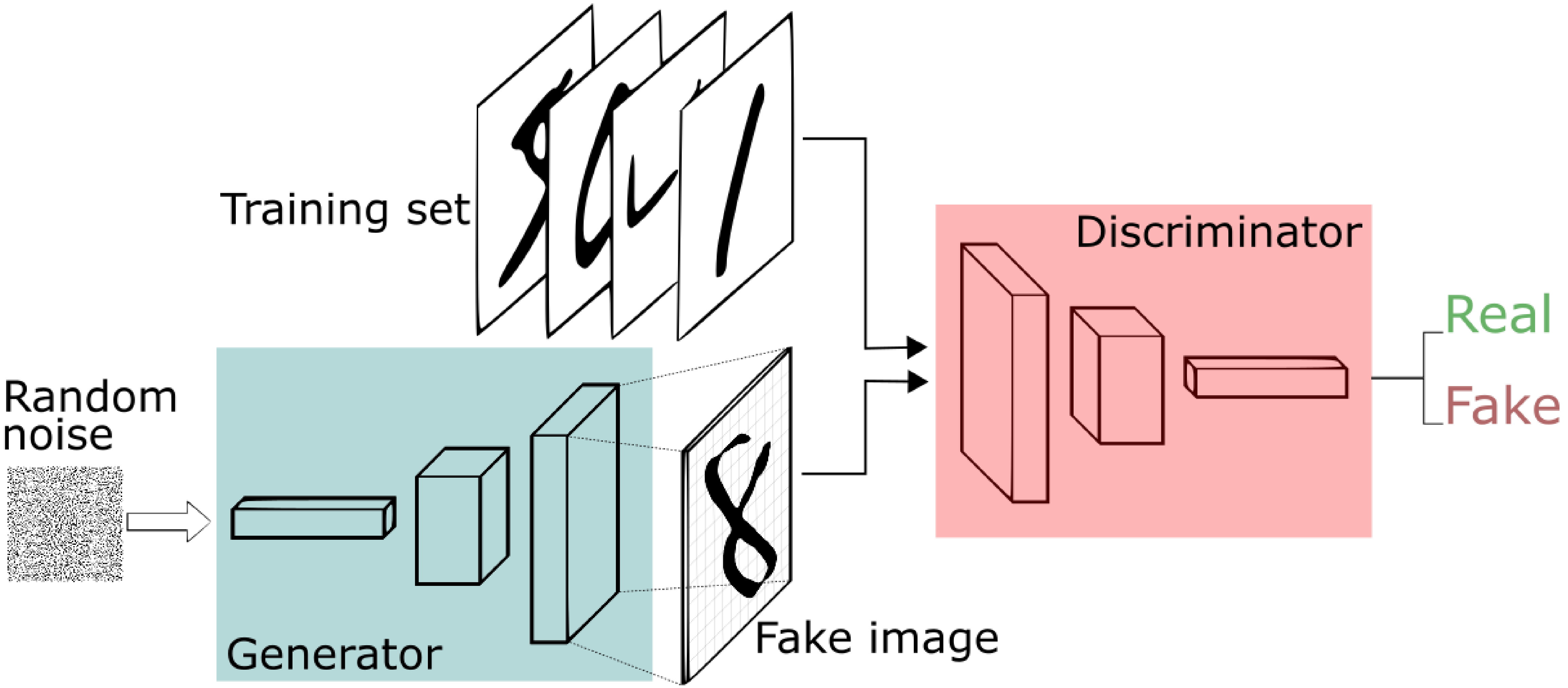
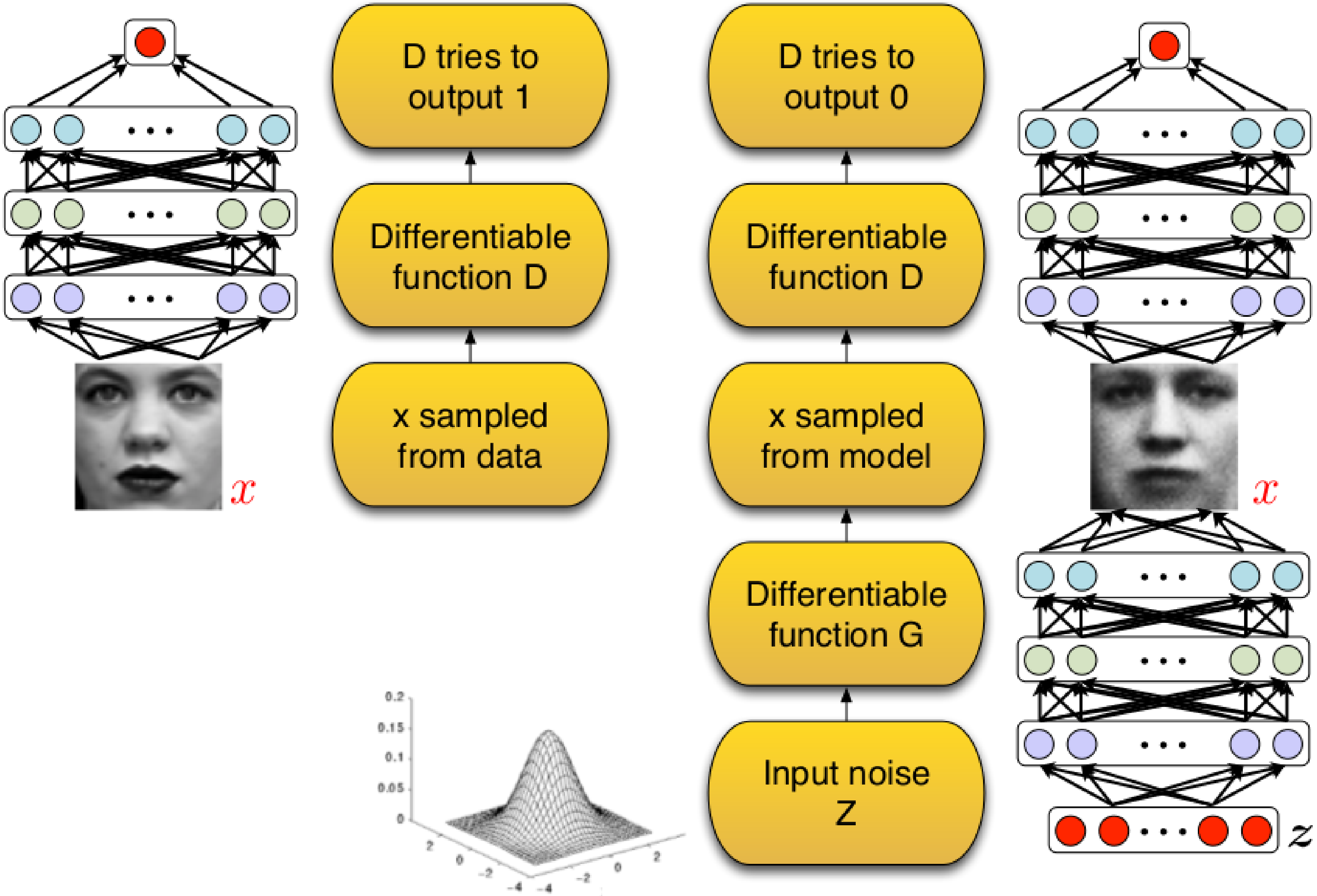


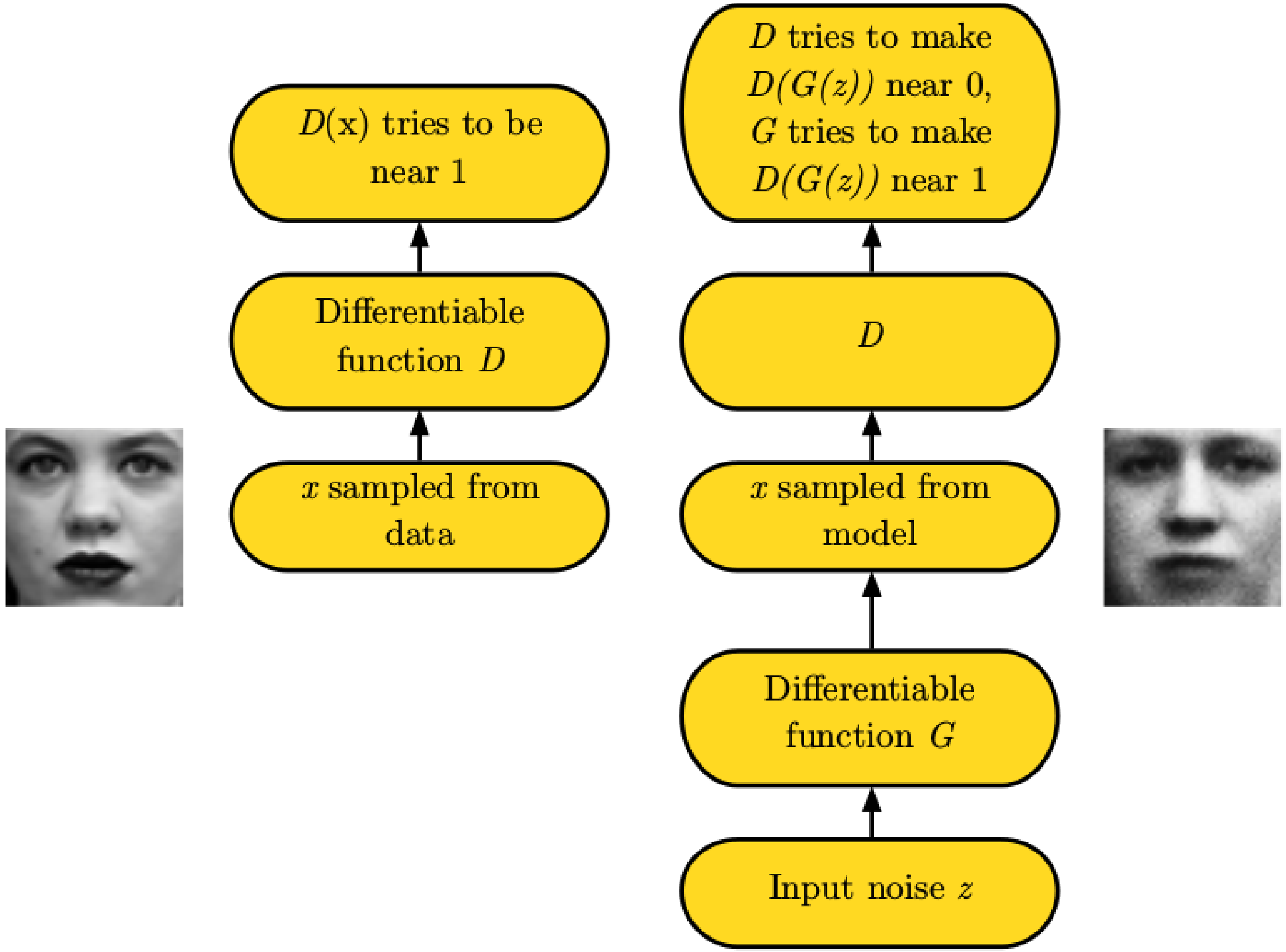
Image credit: Thalles Silva

GAN framework



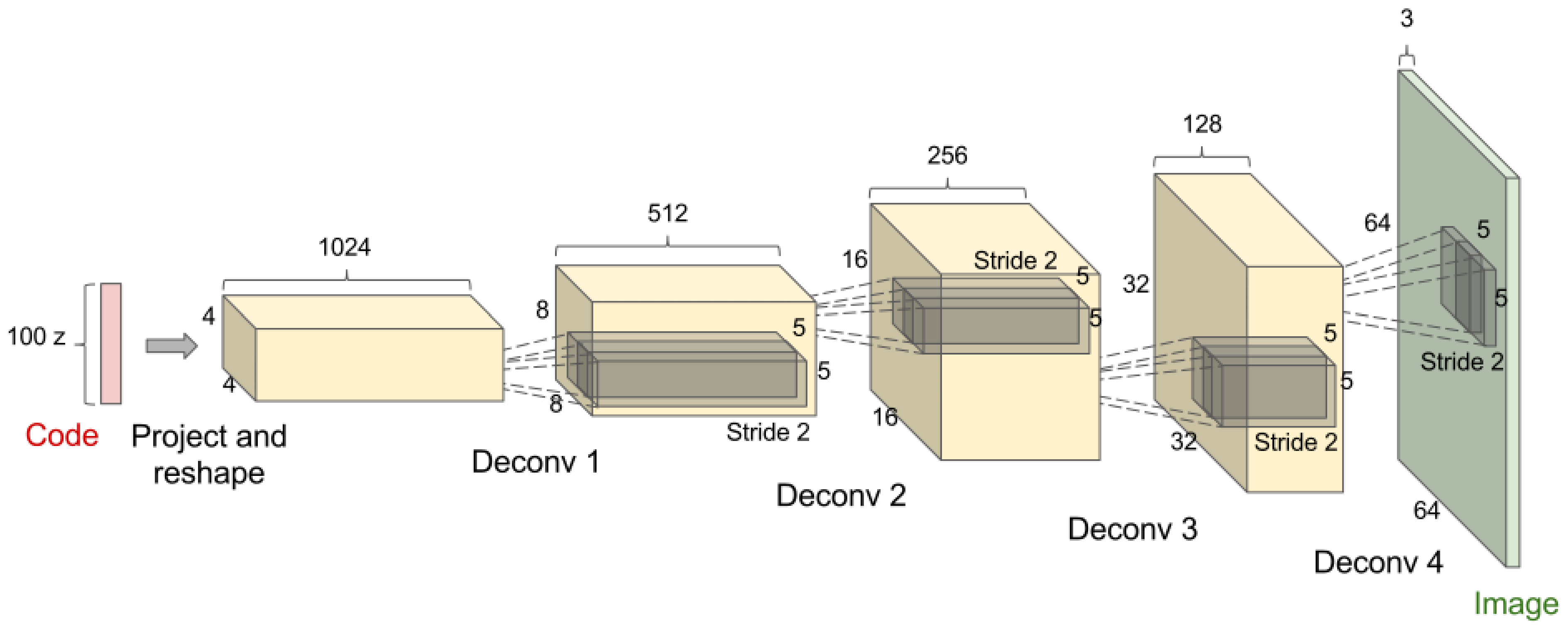
Slide credit: Ian Goodfellow

GANs



D(x) tries to be near 1
Differentiable function D
x sampled from data
D tries to make $D(G(z))$ near 0, G tries to make $D(G(z))$ near 1
D
x sampled from model
Differentiable function G
Input noise z

Extension: Deep Convolutional GAN (DCGAN)



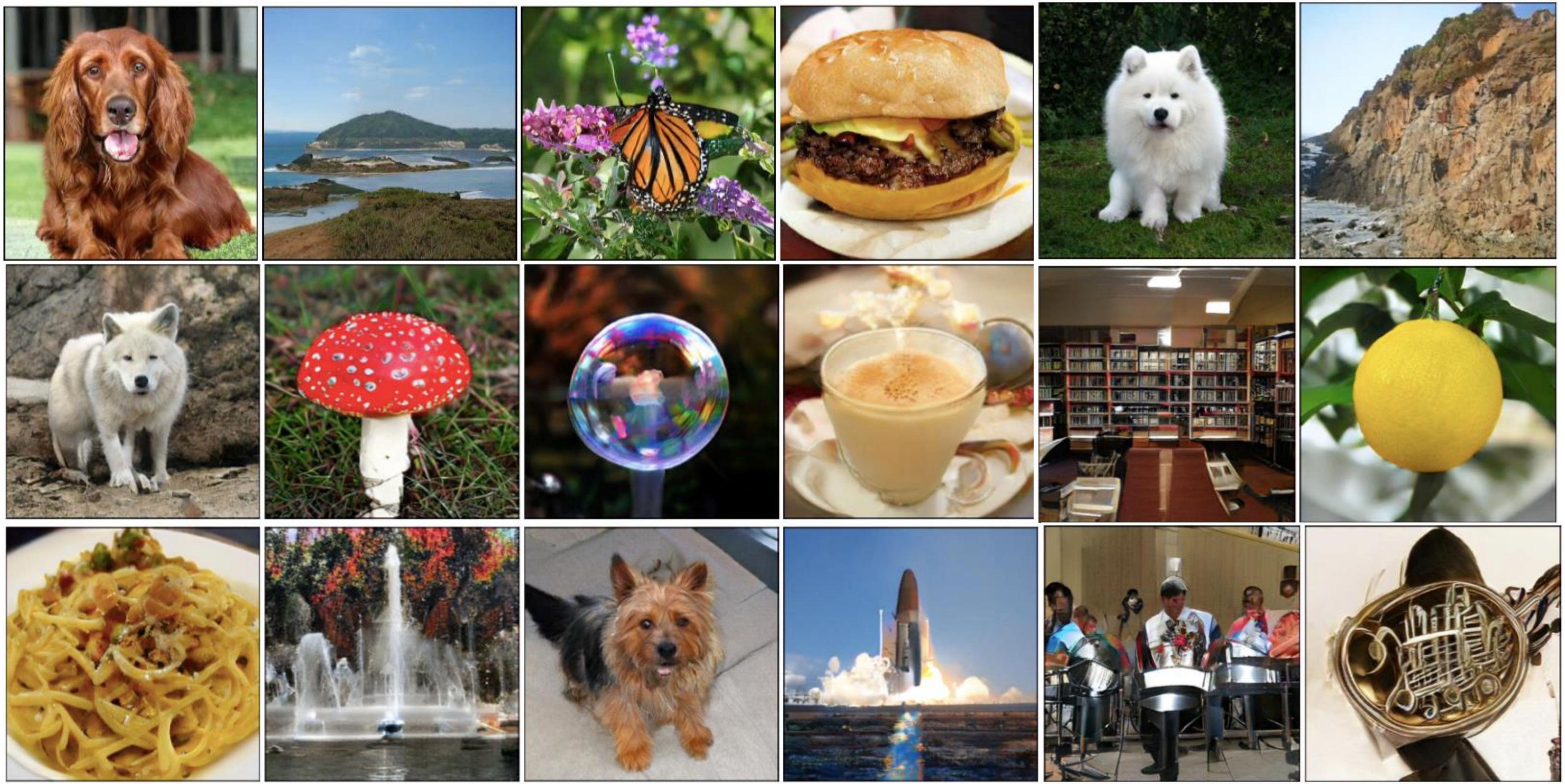
Radford et al ICLR 2016

Samples (Synthetic Images)

Results: Progressive GAN (ICLR 2018)

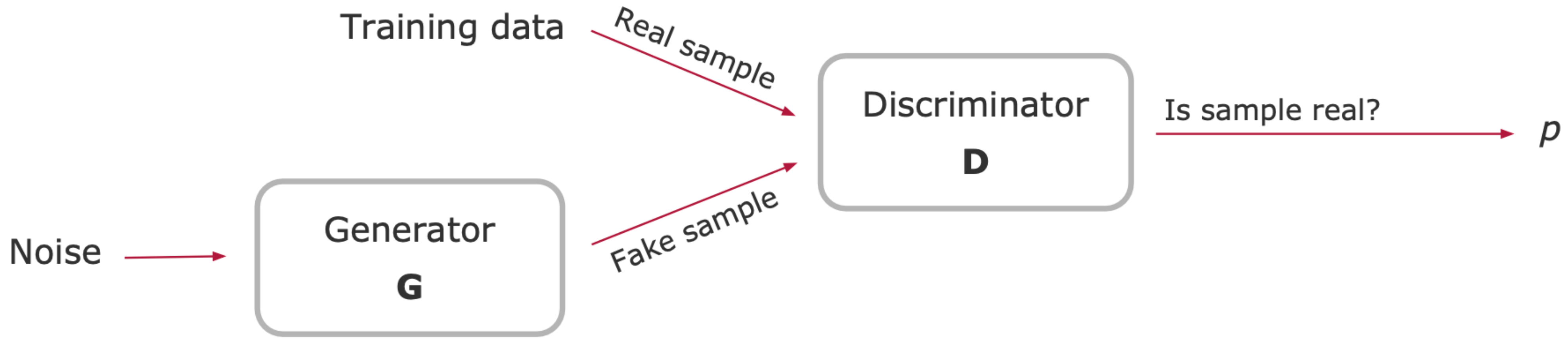
Results: Big GAN (ICLR 2019)

Results: Big GAN (ICLR 2019)



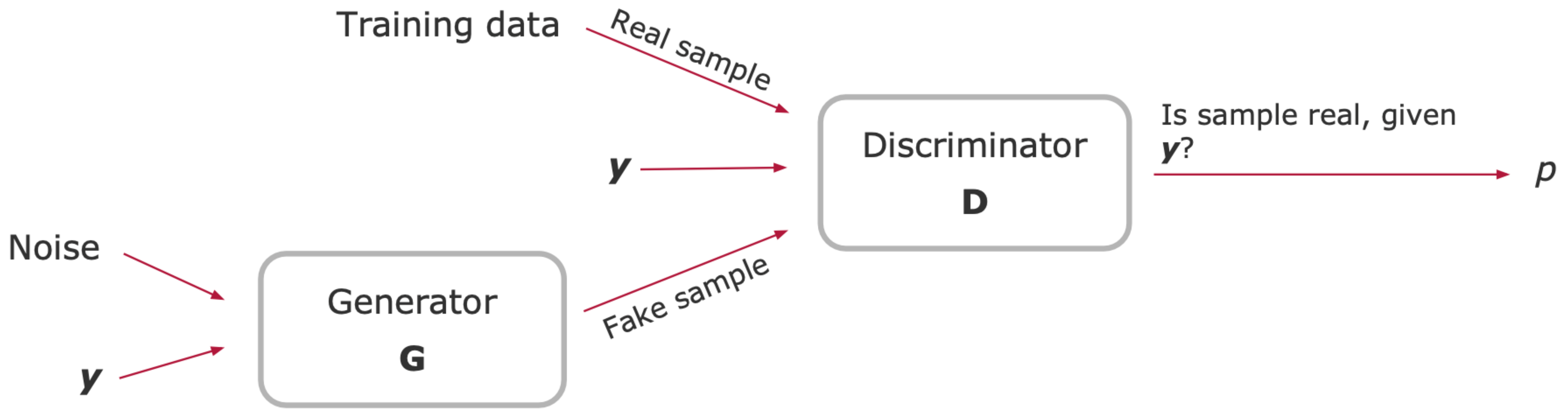
GANs

- **Generator (G)** that learns the real data distribution to generate fake samples
- **Discriminator (D)** that attributes a probability *p* of confidence of a sample being real (*i.e.* coming from the training data)



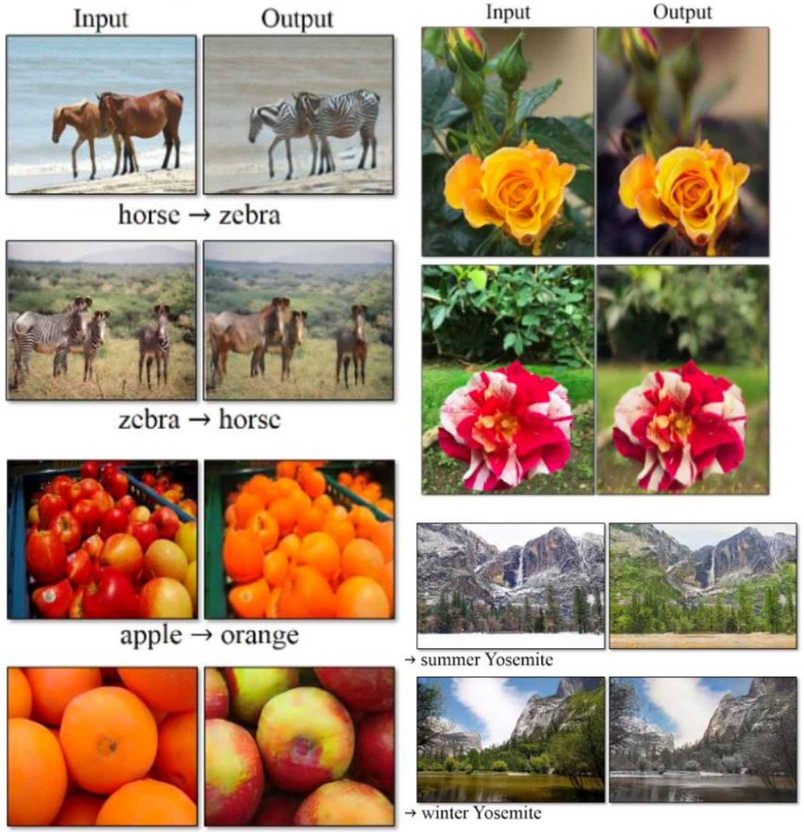
Conditional Generation

- G and D can be **conditioned by additional information *y***
- Adding *y* as an input of both networks will condition their outputs
- *y* can be external information or data from the training set



More GANs

Source->Target domain transfer



CycleGAN. Zhu et al. 2017.

Text -> Image Synthesis

Reed et al. 2017.

Pix2pix. Isola 2017. Many examples at https://phillipi.github.io/pix2pix/

Generation with “control

- Generator is a “black box”
- Separate “Style” and similar “Factors”
- Face
  - Higher Level attribute (Pose, Identity)
  - Fine attributes (freckles, hair, texture)

Results: Style Based Generator

Karras et al. A style-based generator architecture for generative adversarial networks arXiv preprint arXiv:1812.04948

Limitations of GANs

1. **Training instability**

   - Good sample generation requires reaching Nash Equilibrium in the game, which might not always happen

2. **Mode collapse**

   - When G is able to fool D by generating similarly looking samples from the same data mode

3. GANs were original made to work only with **real-valued, continuous data** (*e.g.* images)

   - *Slight changes* in **discrete data** (*e.g.* text) are impractical

How to Evaluate Generated Outputs?

- What makes a good generative model?
  - Each generated sample is indistinguishable from a real sample

  - Generated samples should have variety

FID Score

A Number of Advanced Generative Models

- GANs vs VAEs
- Style GANs and Controllable Generation
- Generation of Multiple Modalities
- Conditional Generation
- Capability to Edit/Process Compact codes
- Diffusion models
- Etc.
- Concern of generation as “fakes”.
  – An emerging ethical concern

Case study

Image to Image Translation

Image to Image Translation with GANs

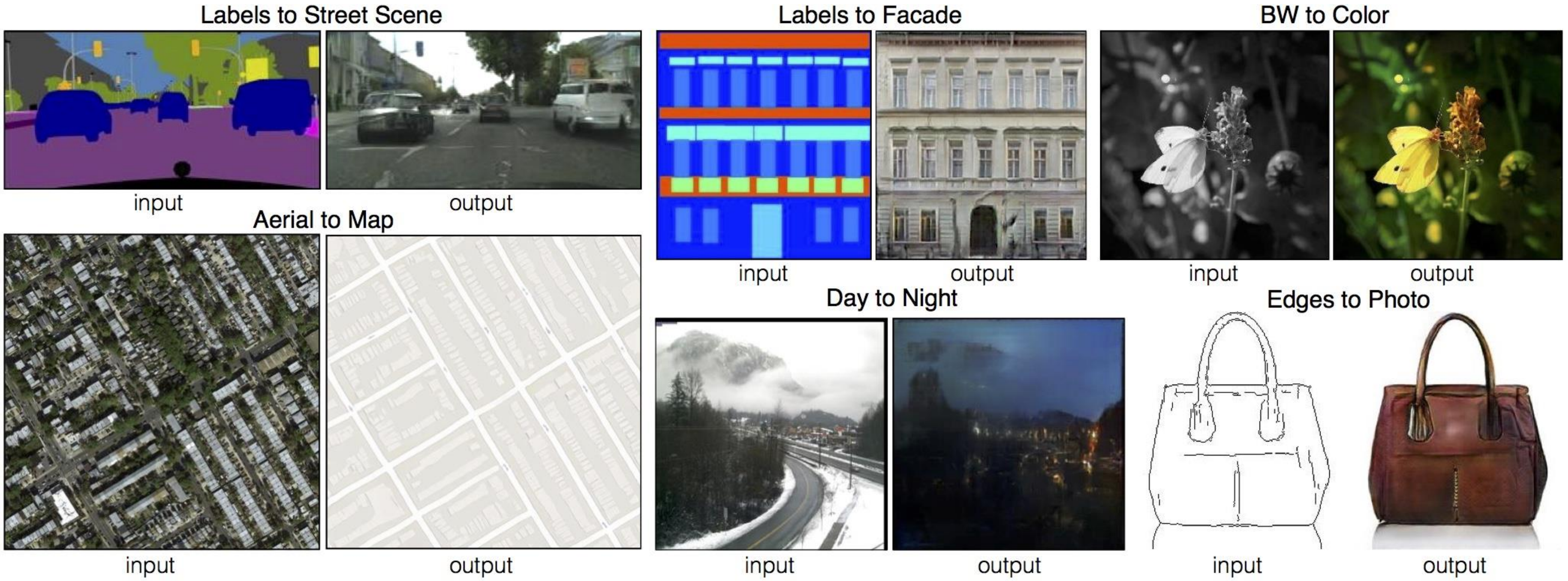
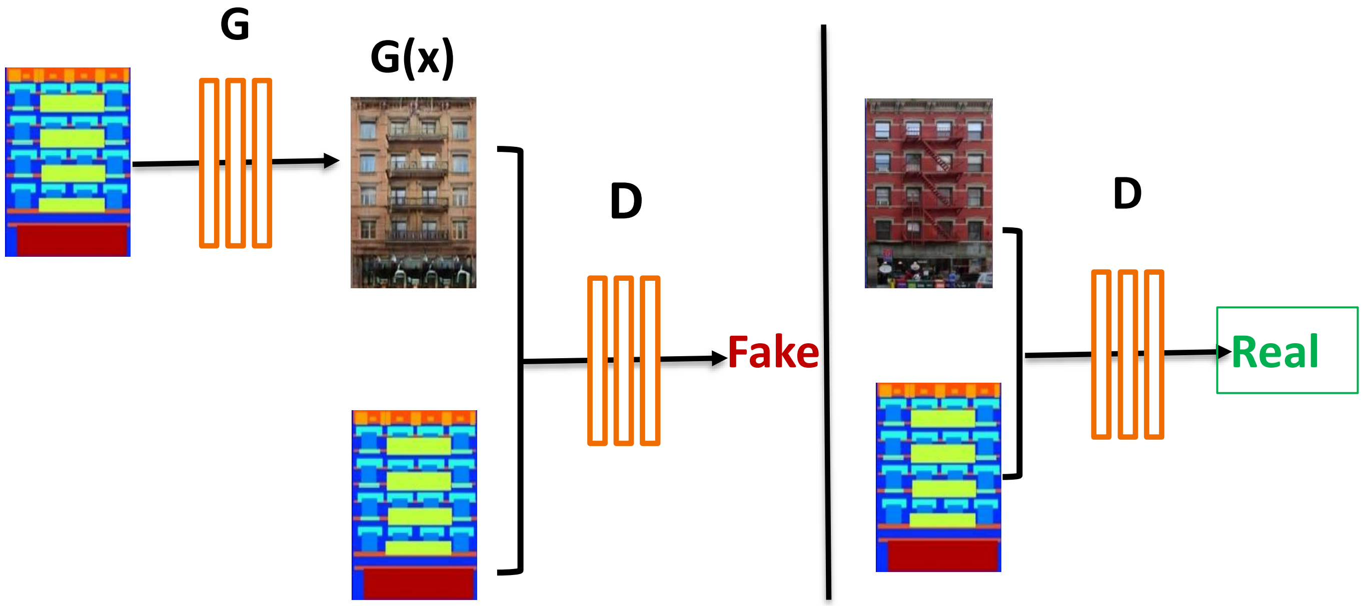


Image-to-Image Translation with Conditional Adversarial Networks
Phillip Isola Jun-Yan Zhu Tinghui Zhou Alexei A. Efros

Key Idea

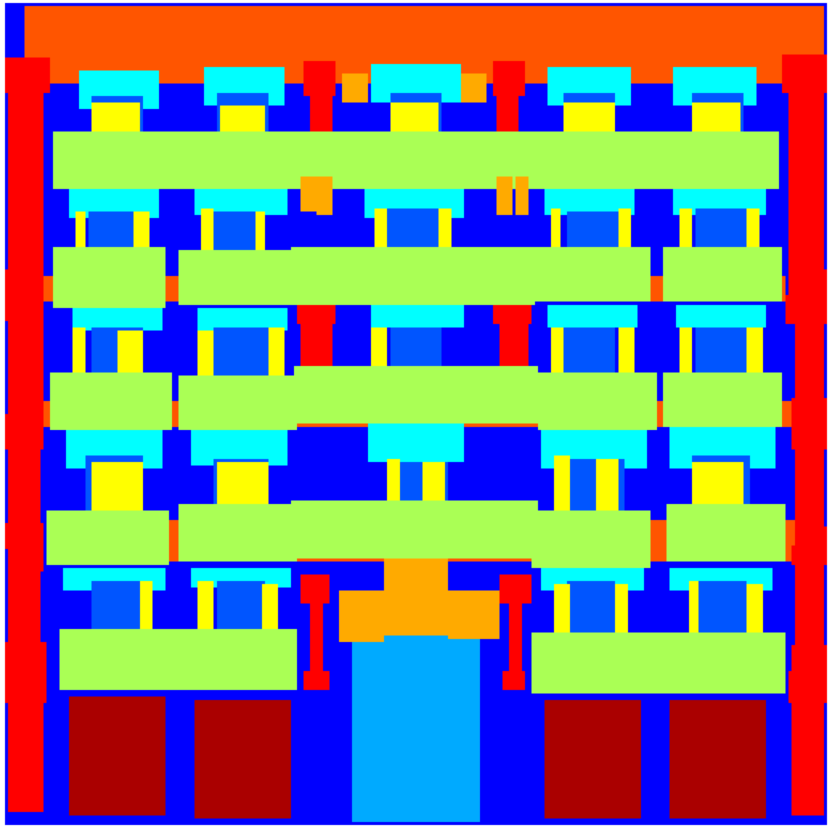
- Training a conditional GAN to map Facade → photo
- The discriminator, D, learns to classify between fake and real {edge, photo} tuples
- The generator, G, learns to fool the discriminator
- Unlike an unconditional GAN, both the generator and discriminator observe the input edge map



Code Available

Results

Input



Output

Data Driven Understanding: *Aka* Deep Learning

AGE = “40”
Location: (x,y)

Regression

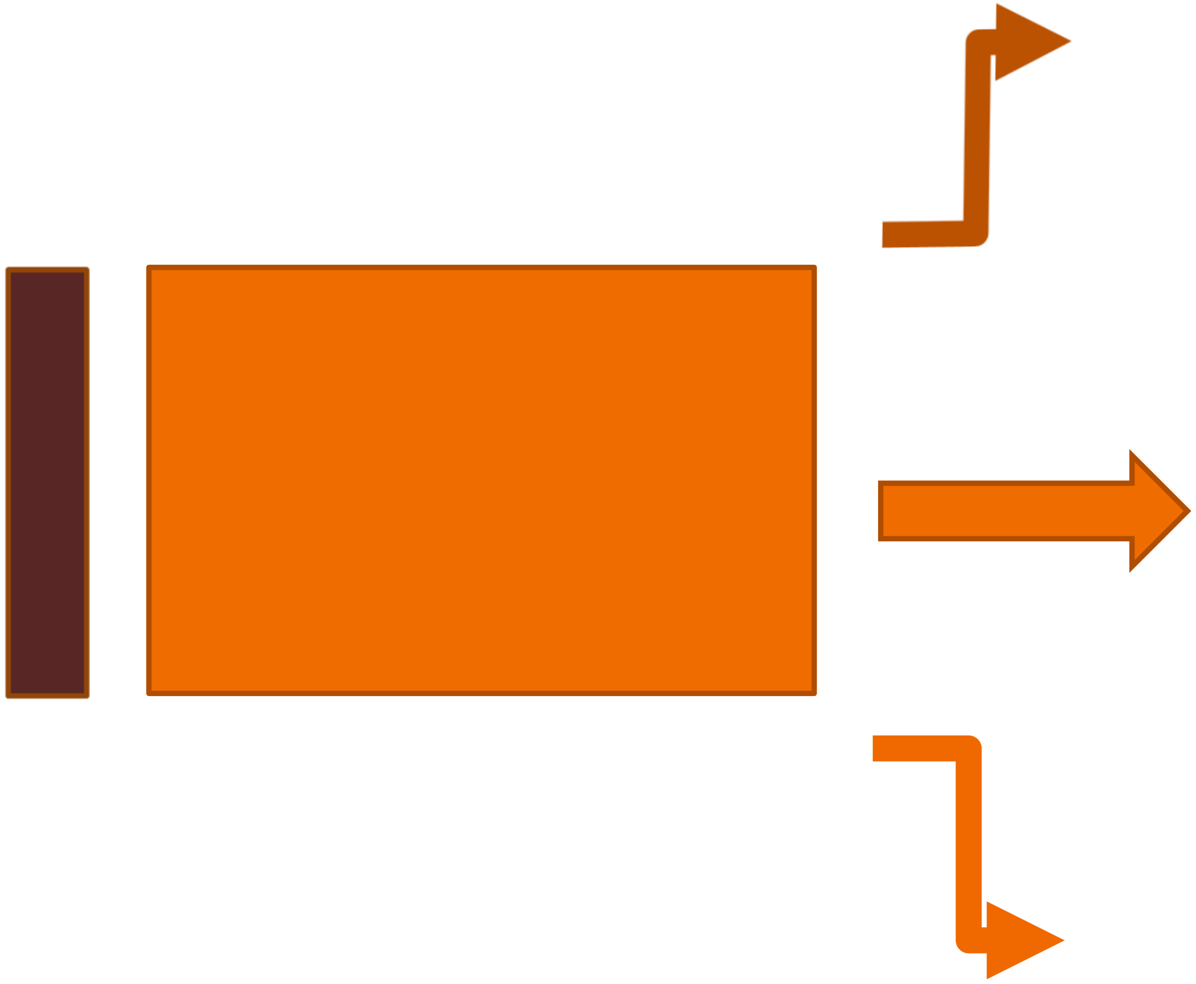
“WALKING” (Action)

“Male” (Gender)

Classification

Structured Prediction

Generative Techniques



To prove study we see that $\mathcal{F}|_U$ is a covering of $\mathcal{X}'$, and $\mathcal{T}_i$ is an object of $\mathcal{F}_{X/S}$ for $i > 0$ and $\mathcal{F}_p$ exists and let $\mathcal{F}_i$ be a presheaf of $\mathcal{O}_X$-modules on $\mathcal{C}$ as a $\mathcal{F}$-module. In particular $\mathcal{F} = U/\mathcal{F}$ we have to show that

$$\widetilde{M}^\bullet = \mathcal{I}^\bullet \otimes_{\mathrm{Spec}(k)} \mathcal{O}_{S,s} - i_X^{-1}\mathcal{F})$$

is a unique morphism of algebraic stacks. Note that

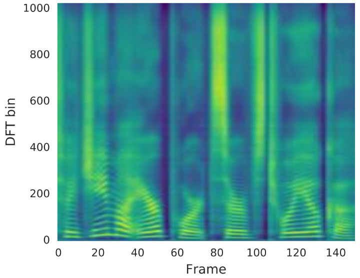
$$\mathrm{Arrows} = (Sch/S)^{opp}_{fppf}, (Sch/S)_{fppf}$$

and

$$V = \Gamma(S, \mathcal{O}) \longmapsto (U, \mathrm{Spec}(A))$$

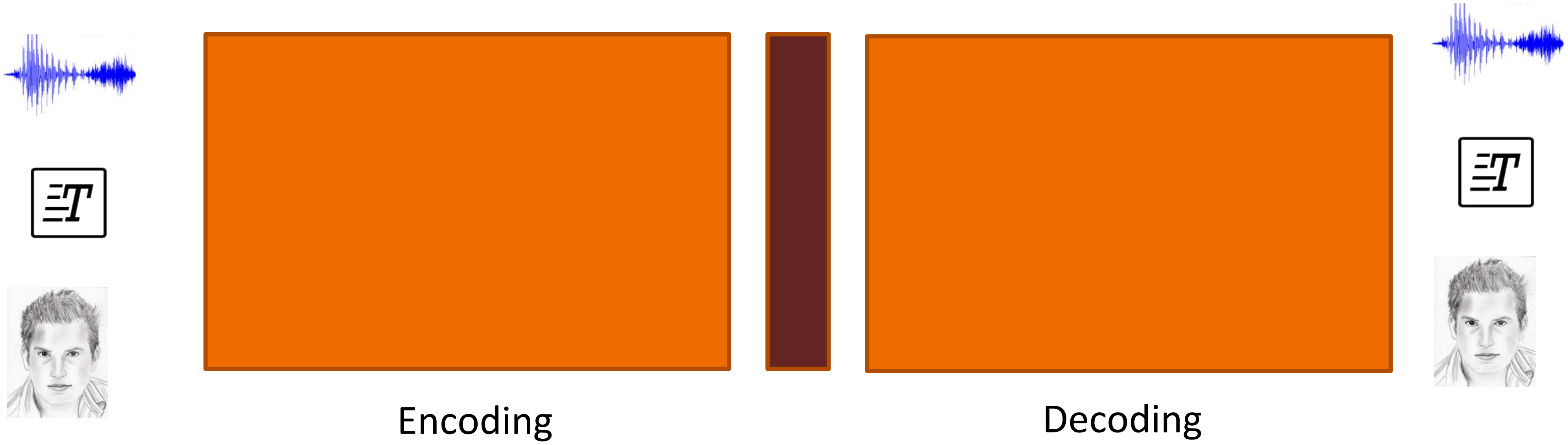
is an open subset of X. Thus U is affine. This is a continuous map of X is the inverse, the groupoid scheme S.

Proof. See discussion of sheaves of sets. $\square$



Andrej Karpathy, “RNN” Ian Goodfellow, “GAN” LSTM/Conv1D

Powerful Combination: Translational Models



Inherently multimodal and cross-modal

Summary

- A new training paradigm
  - Adversarial Training
- A set of new methods for generation
  - GANs, VAEs, variants
- Applications
  - Understanding ML, Data, Algorithms
  - Practical Utility vs Ethical Concerns: Race that can never end?
- Cross Modality
  - Transfer and Translation

Thanks!!

Questions?