Deep Learning

Álvaro Romo

Analytical Index

Analytical Index

1. Introduction to Deep Learning
2. Convolutional Neural Networks (CNN)
3. Recurrent Neural Networks (RNN)
4. Natural Language Processing with Deep Learning
5. Deep Generative Modelling
    a. Discriminative vs Generative
    b. Autoencoders
    c. Variational Autoencoders (VAEs)
    d. GANs: Generative Adversarial Networks
    e. Diffusion Models and Other Projects

5

Deep Generative Models

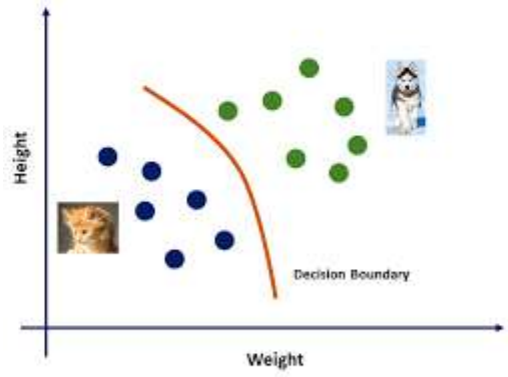
Introduction

5.1

Generative vs Discriminative

Generative vs Discriminative models

Discriminative

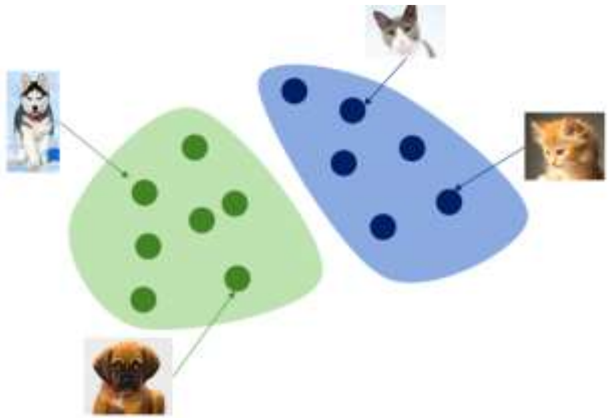


Model of the conditional probability of the target Y, given an observation x

Features $X \rightarrow Y$ Target

$$P(Y|X)$$

Generative

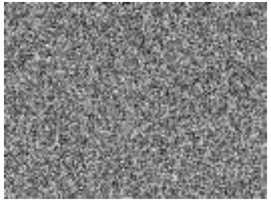


Target $Y \rightarrow X$ Features

$$P(X|Y)$$

Generative

Ruido a imagen



Texto a imagen

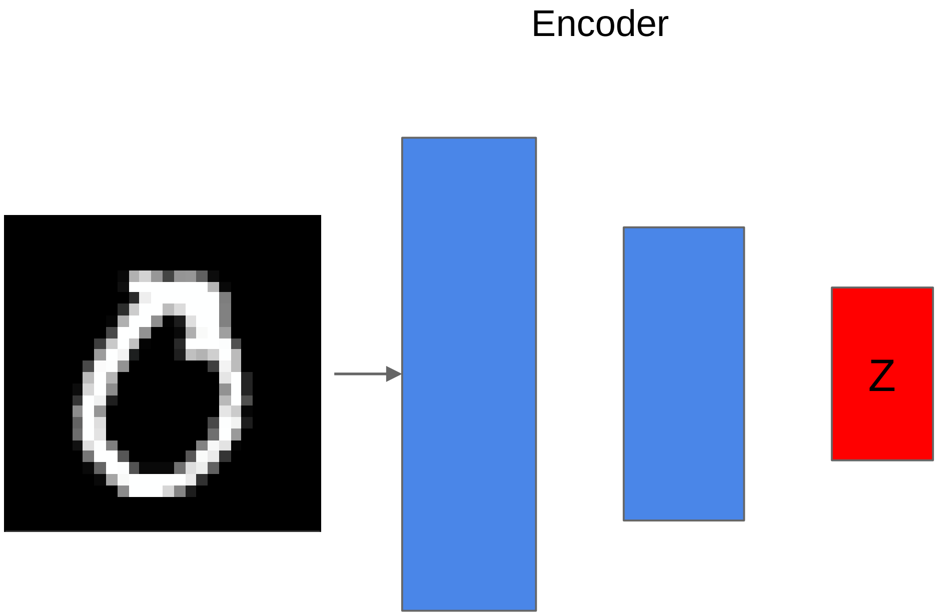
Genera un castillo con niebla en las montañas con un toque fantástico

Imagen a imagen

5.2

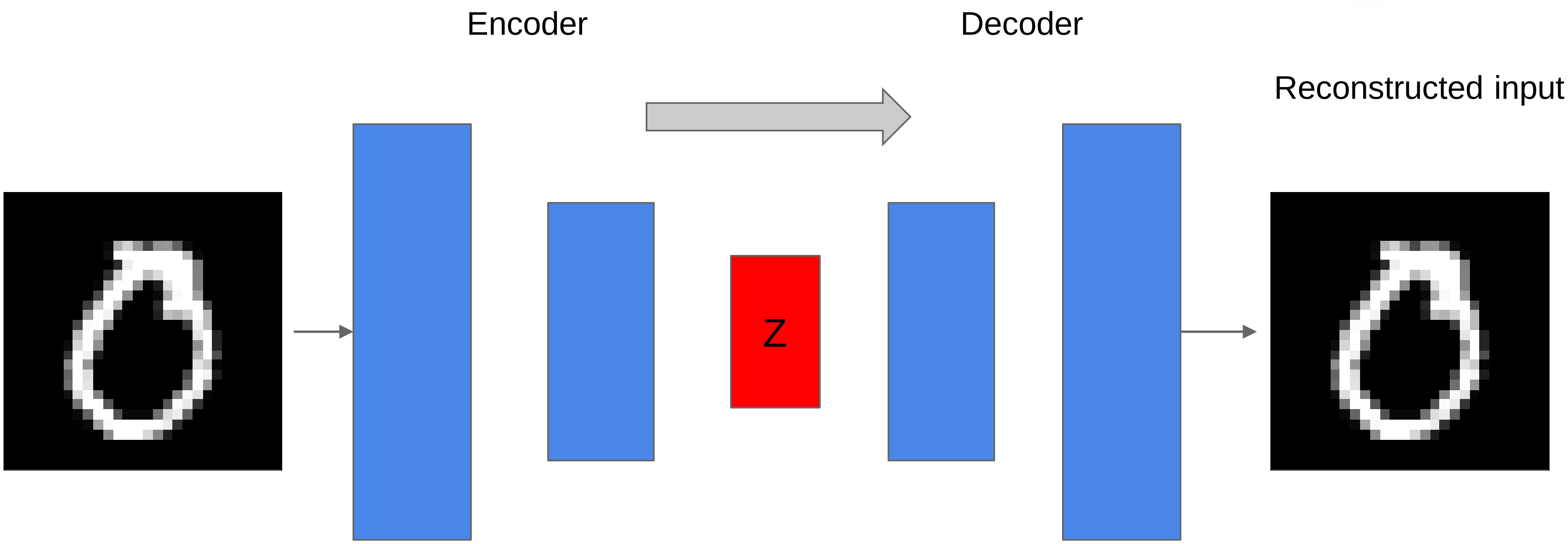
Autoencoders

Autoencoder Intuition



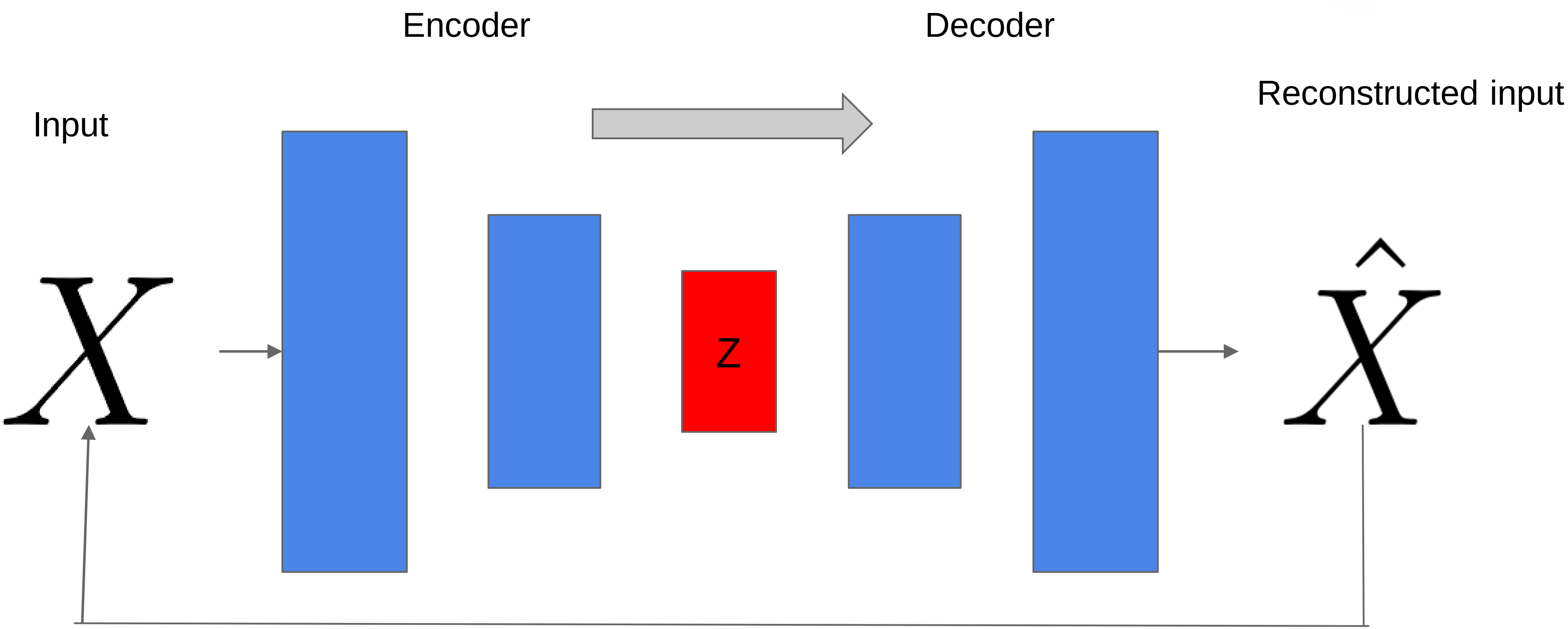
- We want to obtain a lower dimensional representation from the input data.
- Useful for dimensionality reduction and visualization.
- The encoder maps input data into a lower dimensional latent space (z)

Autoencoder Intuition



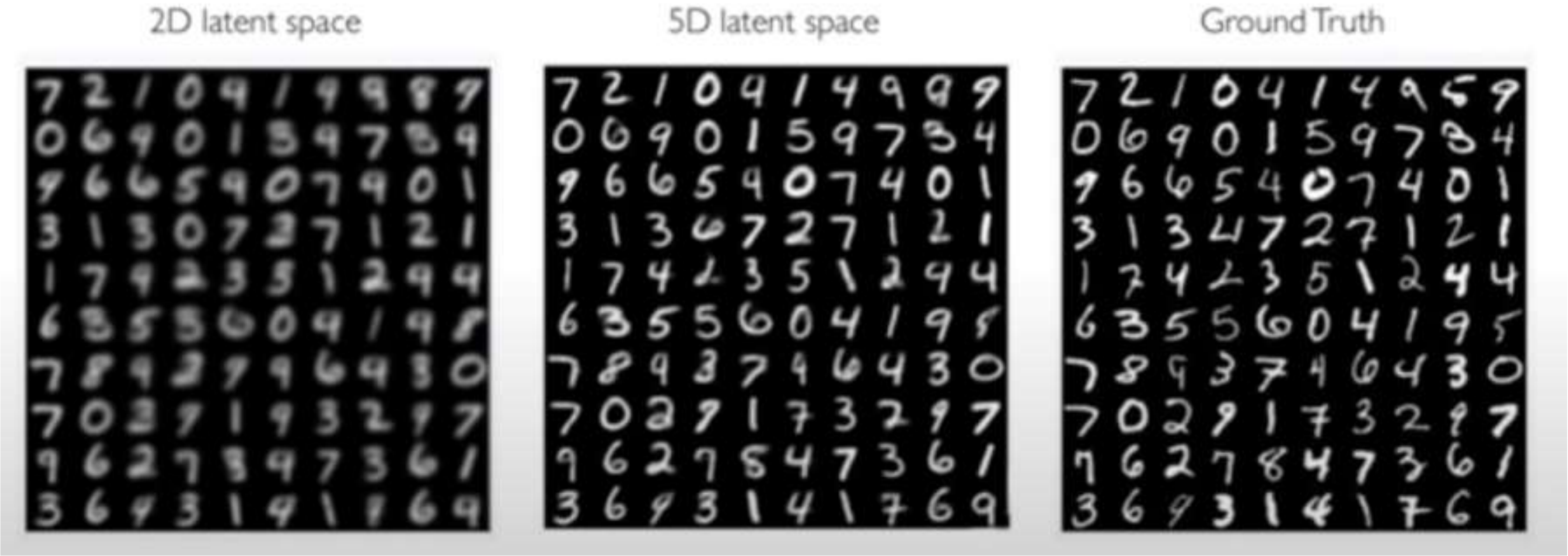
Train the model to reconstruct the original input data.
The decoder learns to reconstruct the data with the latent space.

Autoencoder: Reconstruction Loss

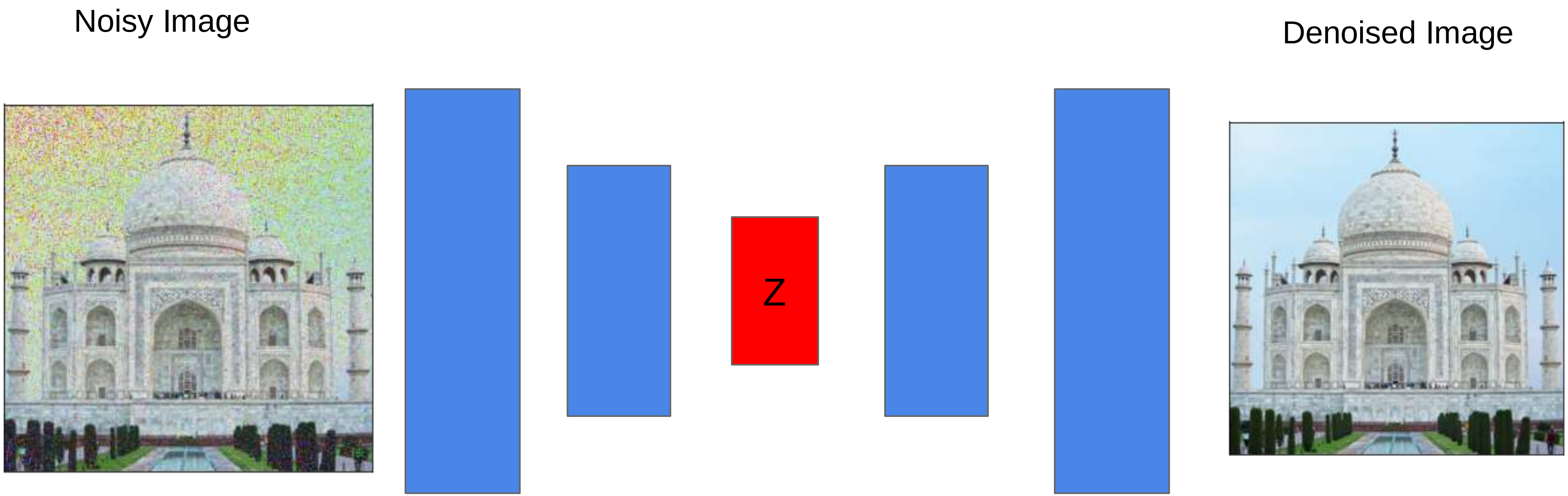


$$\mathcal{L}(X, \hat{X}, W) = ||X - \hat{X}||^2$$

Autoencoder Applications: Compression



Autoencoder Applications: Denoising

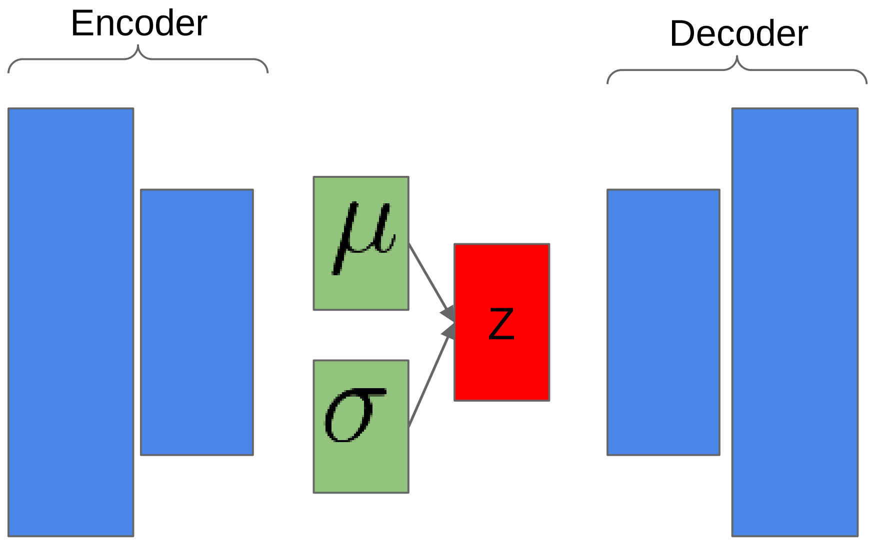


Autoencoder Applications: Colorizing Images

5.3

Variational Autoencoders

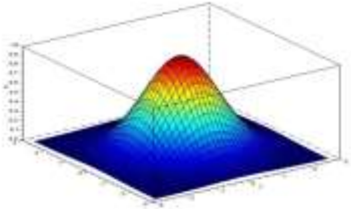
VAEs



Encoder computes: $q_\phi(z|x)$

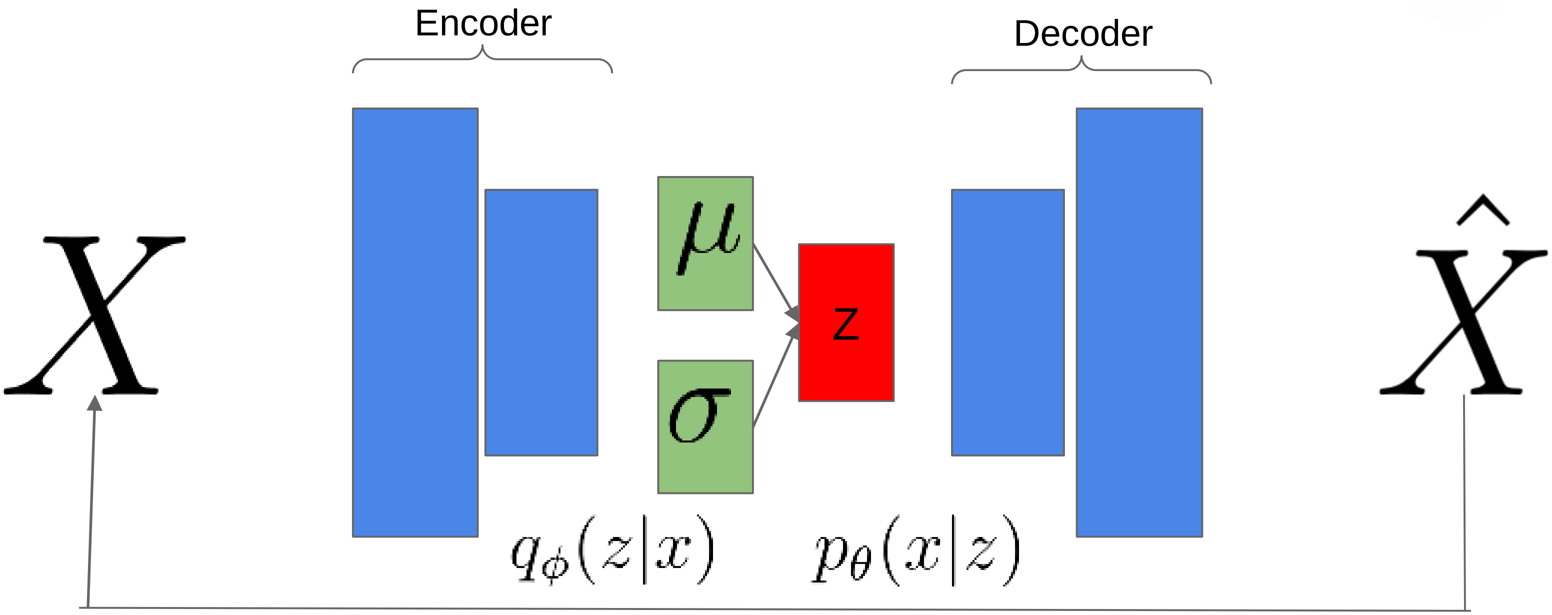
Decoder computes: $p_\theta(x|z)$

$x \xrightarrow{q_\phi(z|x)} z$



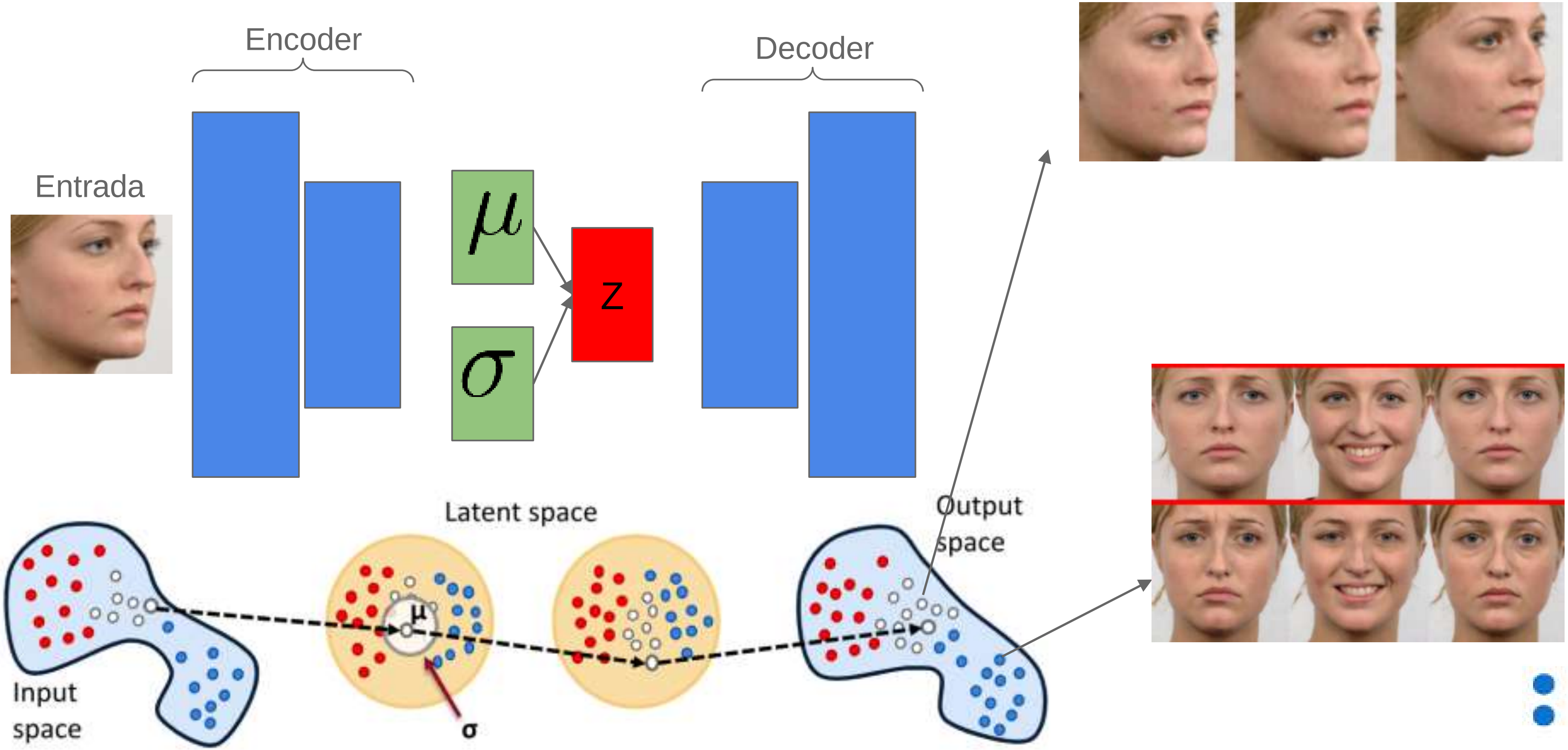
$z \xrightarrow{p_\theta(x|z)} \tilde{x}$

VAEs: Loss



$$\mathcal{L} = \underbrace{||X - \hat{X}||^2}_{\text{Reconstruction loss}} + D(\underbrace{q_\phi(z|x)||p(z)}_{\text{Regularization}})$$

VAEs



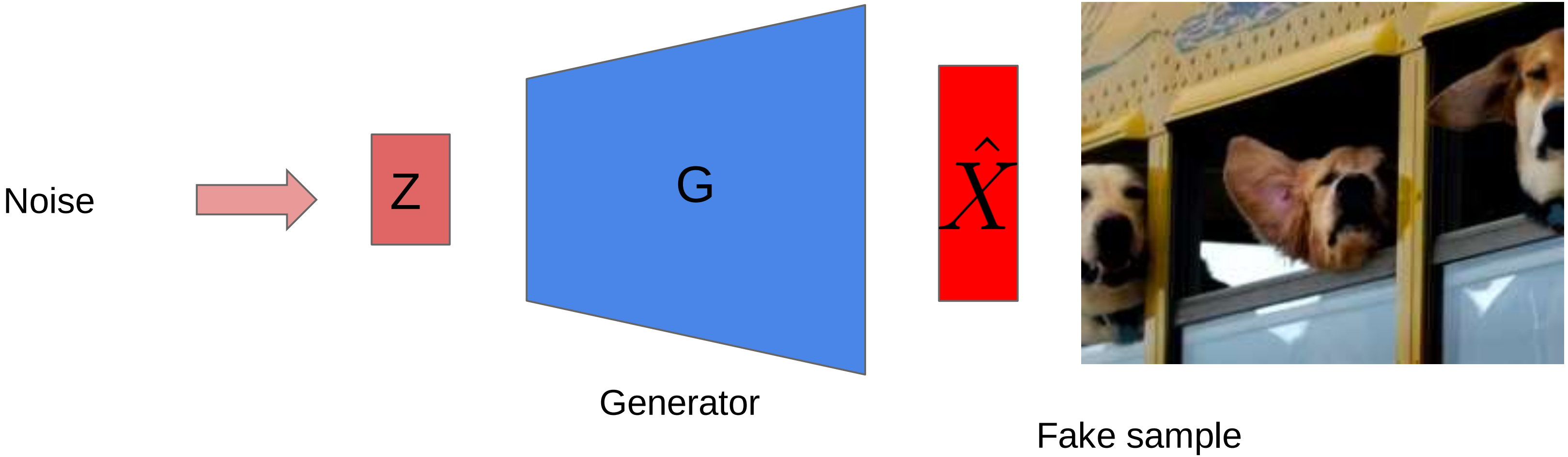
Encoder
Decoder
Entrada
μ
σ
z
Latent space
Input space
μ
σ
Output space

5.4

GANs: Generative Adversarial Networks

GANs: Why do we need to learn the distribution?

Instead of learning the density of the data, why don't we train a model that samples it directly?



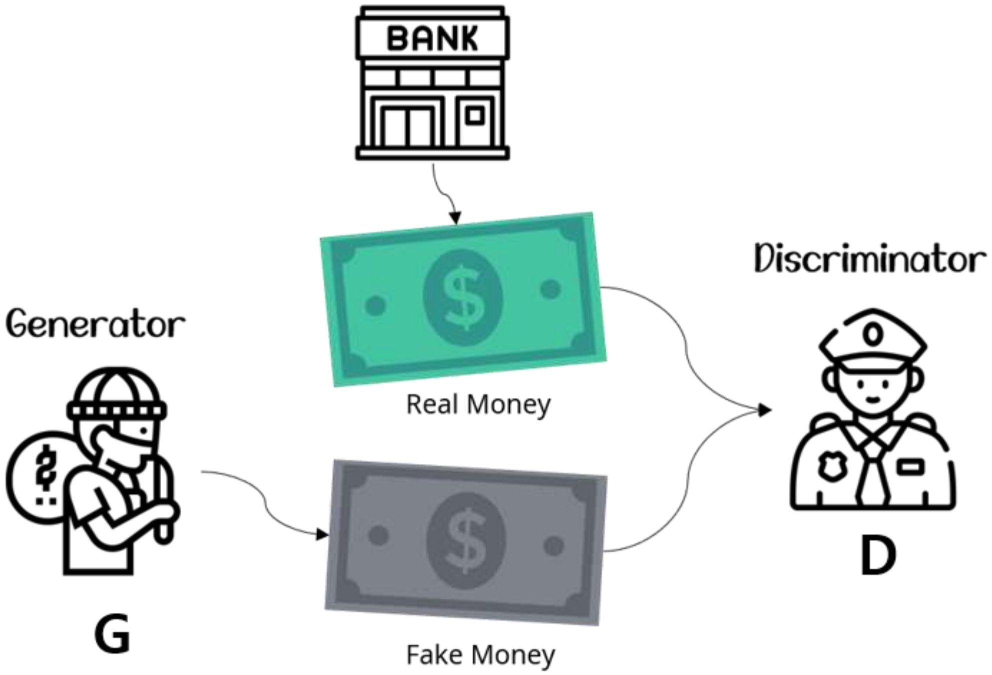
GANs: Generator & Discriminator Intuition

Generative ADVERSARIAL Network: dos redes neuronales compiten.

Generator

Aprende a muestrear datos generadas, que parezcan reales.

Podríamos pensar en el generador como si fuese un falsificador

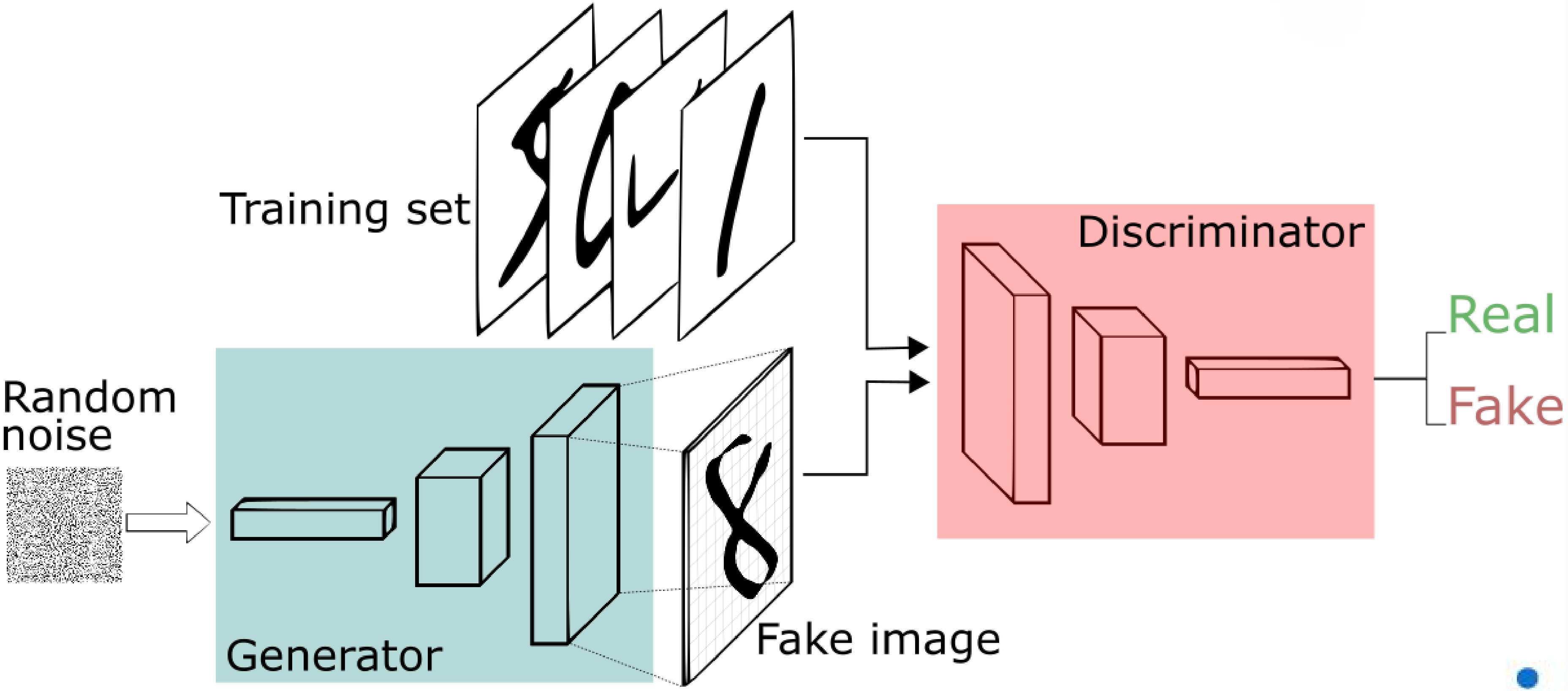


Discriminator

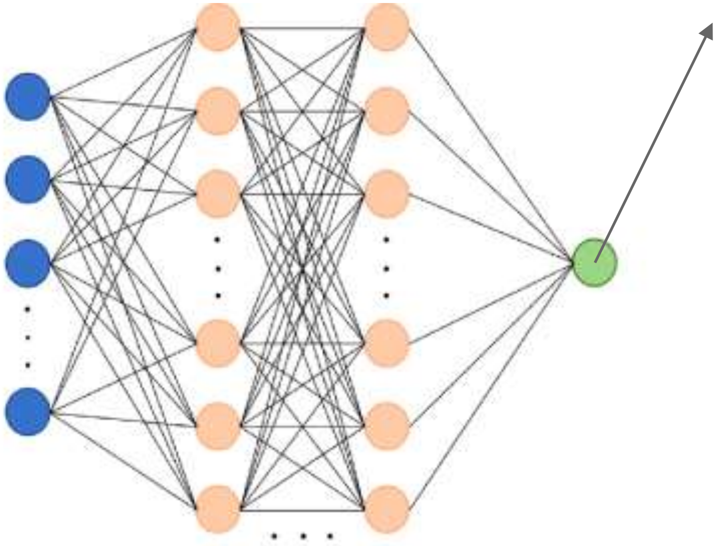
Aprende a distinguir los datos reales de los generados..

Podemos pensar como un en el discriminador como un policía.

GANs: Generator & Discriminator



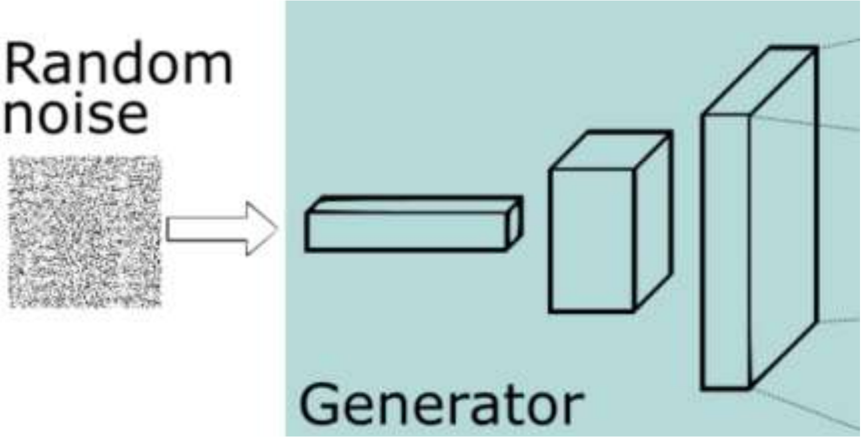
GANs: Discriminator



Fake or Real ??

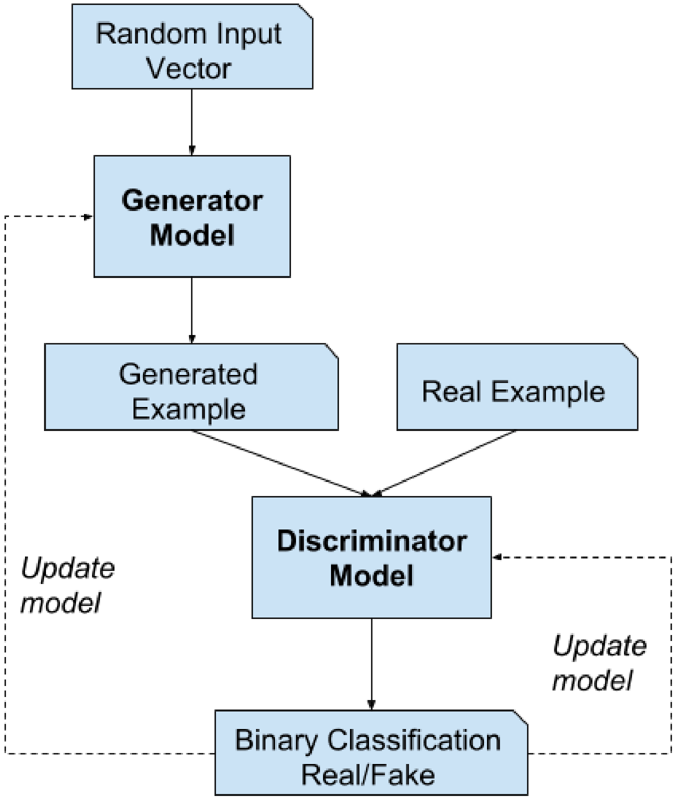
- Is a binary classifier.
- Learns the probability of class Y (**fake** or **real**) given input X.
- Probabilities are the feedback for the generator.

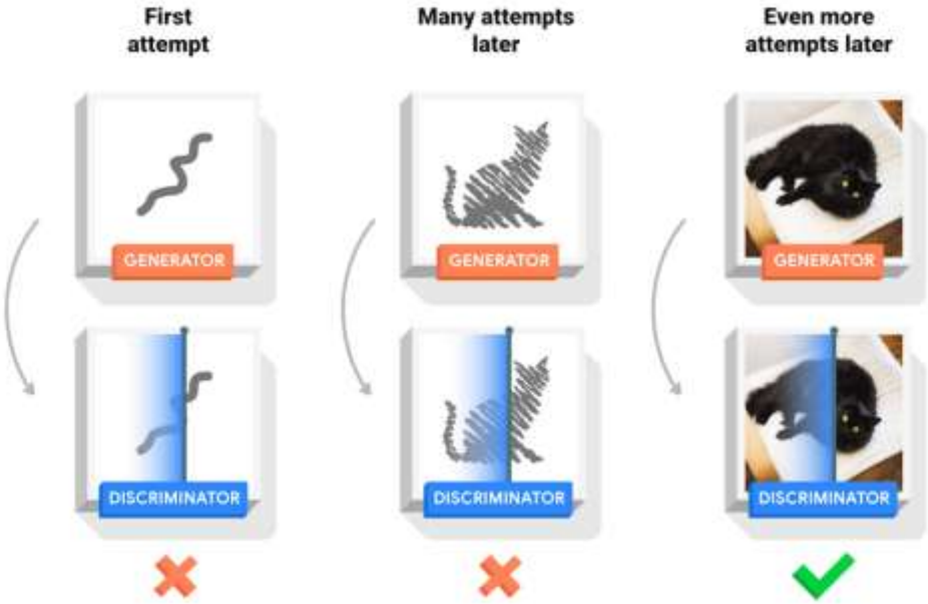
GANs: Generator



- Produces fake data.
- Learns the approximate distribution of inputs data.
- Takes noise as input

GANs: Training





Generator and discriminator, are trained together. The generator generates a batch of samples, and combined with real data are provided to the discriminator and classified as real or fake.

GANs: StyleGAN

StyleGAN is a novel generative adversarial network (GAN) introduced by Nvidia researchers in December 2018

GANs: StyleGAN

GANs: Pix2Pix Image to Image Translation

Paired translation: Train with pairs of images of the two domains

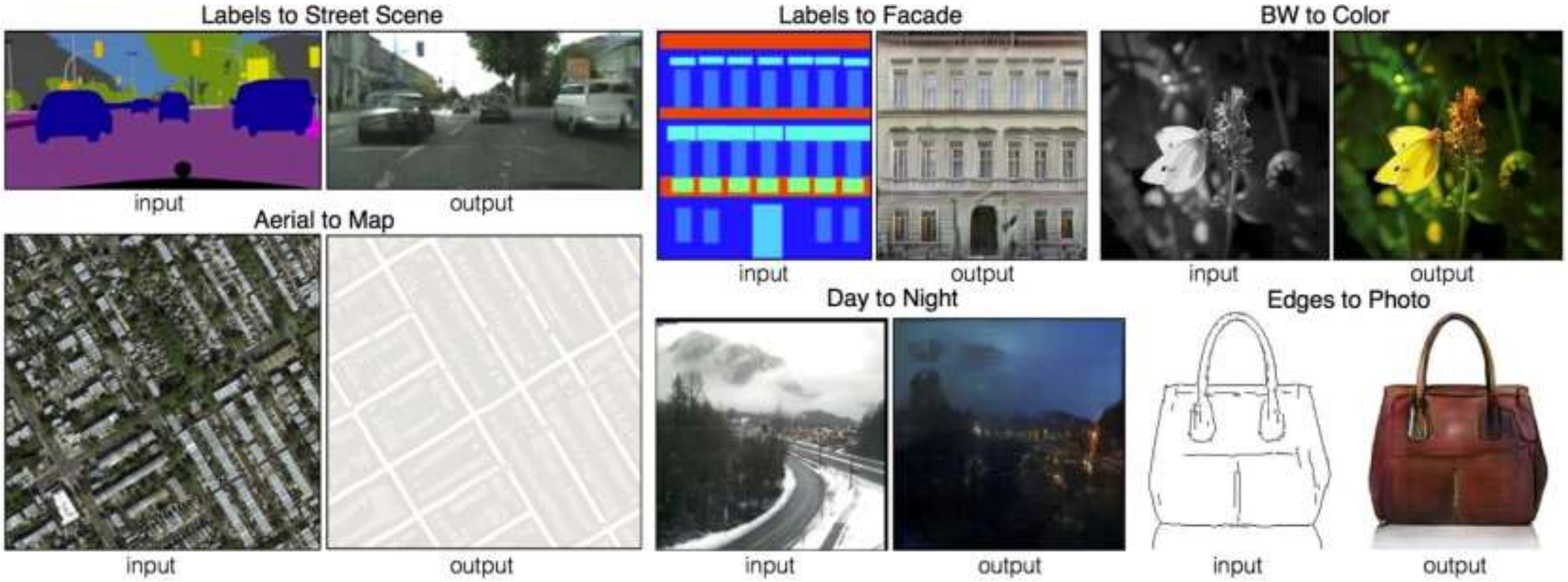


Image-to-Image Translation with Conditional Adversarial Networks

GANs: Face de-aging

https://blogthinkbig.com/deaging-juventud-hollywood-efectos

GANs: CycleGAN, Image to Image Translation

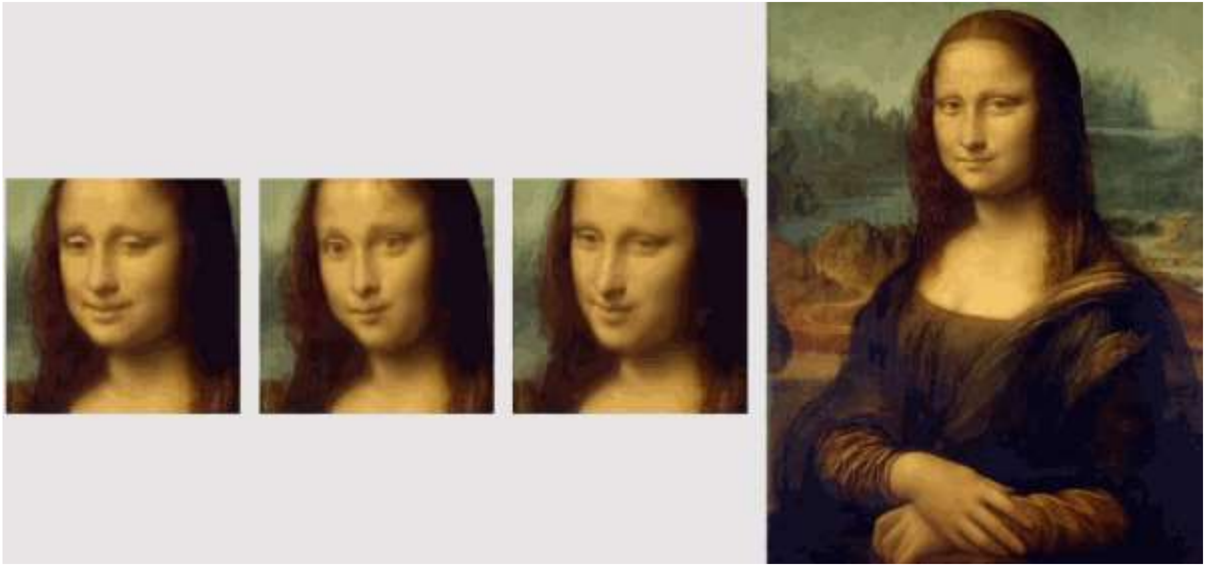
Unpaired image-to-image: Capture the characteristics of one image domain and figure out how these characteristics could be translated into another image domain.

GANs: Multimodal MUNIT

(a) edges ↔ shoes (b) edges ↔ handbags

https://github.com/NVlabs/MUNIT

GANs



Few-Shot Adversarial Learning of Realistic Neural Talking Head Models

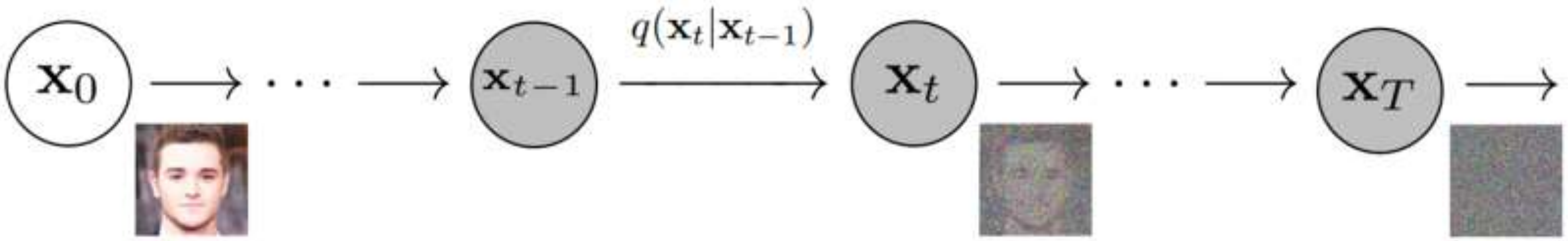
5.4

Diffusion Models and Other Projects

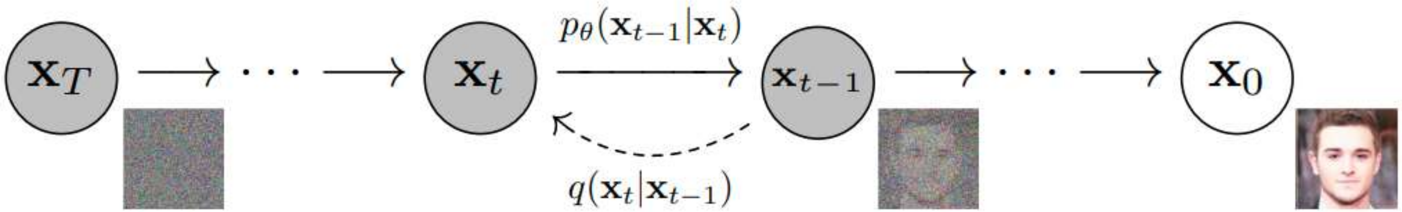
Diffusion models

The Diffusion Process

- Diffusion process: original data sample is progressively corrupted by adding noise.



- Reverse process (denoising): learn to recover the data by reversing this noising process.



[Denoising Diffusion Probabilistic Models Paper]

Sesgo en imagenes

https://www.bloomberg.com/graphics/2023-generative-ai-bias/

https://theconversation.com/ageism-sexism-classism

Midjourney

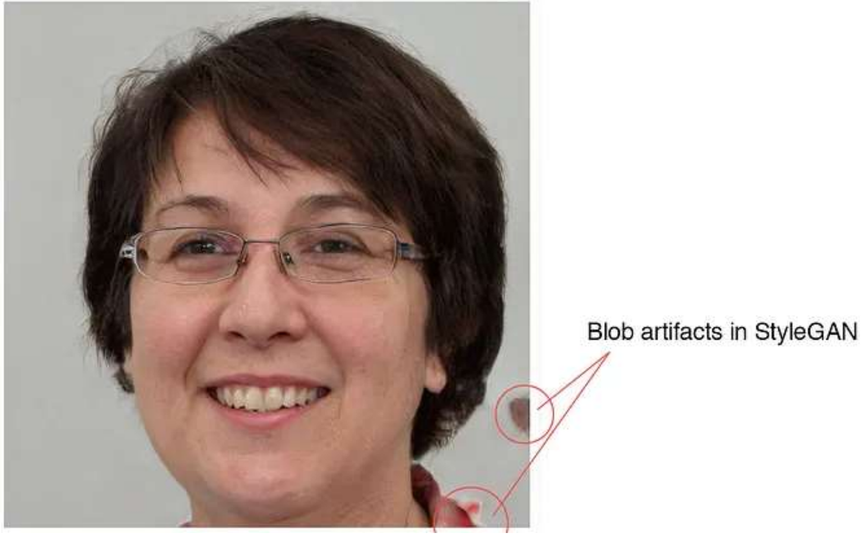
https://www.midjourney.com/home/

Fake news

Artificial intelligence is automating the creation of fake news, spurring an explosion of web content.

This content mimics factual articles by spreading false information about elections, wars and natural disasters.

AI generated image detection



In December 2019, Facebook deleted 682 accounts that allegedly used deceptive practices to push pro-Trump narratives to some 55 million users. As Facebook stated, some of these accounts used profile pictures generated by artificial intelligence

ntic master
revolucionamos la comunicación

UNIVERSIDAD COMPLUTENSE DE MADRID