Redes Adversárias Generativas (GANs)

Prof. André Hochuli

Sumário

- Modelos Discriminativos e Generativos
- Redes Adversárias Generativas
- Arquiteturas e Aplicações
- Tutoriais

Modelos Discriminativos e Generativos

- **Discriminativos:**

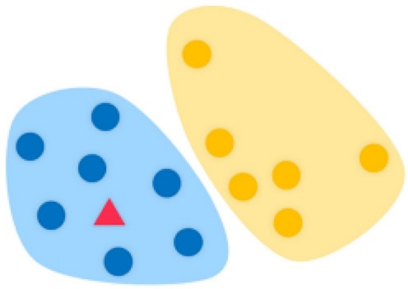
  Definem uma fronteira de decisão

  Discriminam o problema em classes

- **Generativos:**

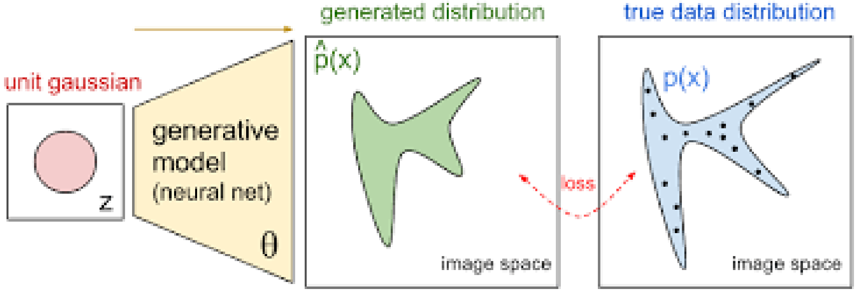
  Aprendem a distribuição dos dados

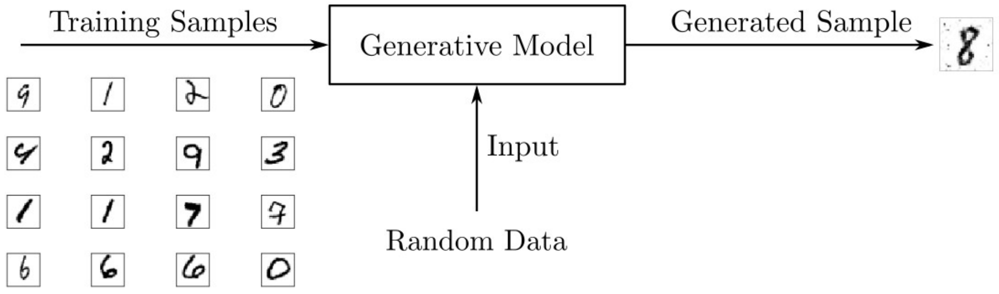
  São capazes de gerar novos dados



Modelos Discriminativos
Arquitetura

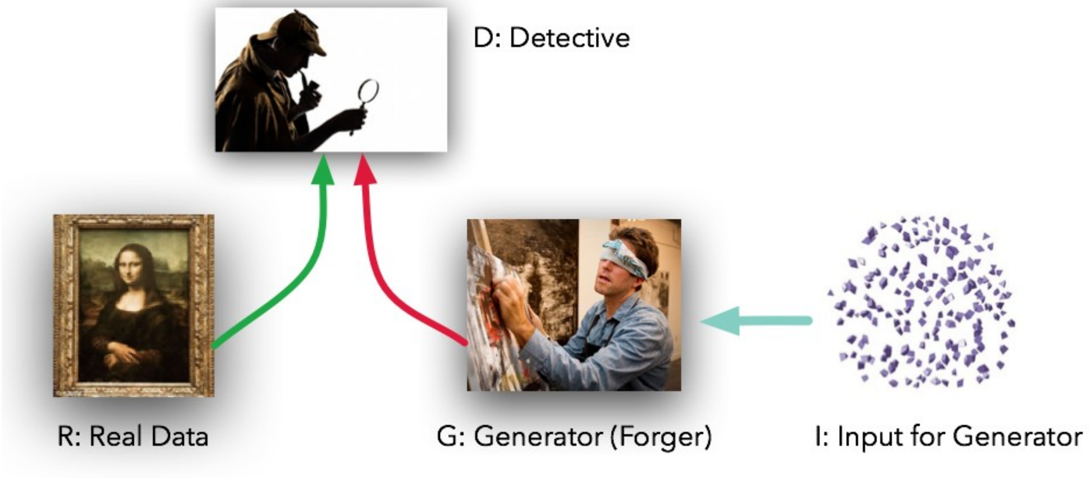
- O treinamento consiste em aprender a distribuição a partir de um ruído
- Gerar dados aleatórios (sintéticos) dentro da distribuição real





Redes Adversárias Generativas

- Generative Adversarial Nets (GANs) – Ian GoodFellow - 2015
- Criar dados sintéticos realistas (*fakes)*
- Objetivo da rede generativa é confundir o discriminador
- Vasta aplicação em Visão Computacional



Generative Adversarial Nets

Ian J. Goodfellow, Jean Pouget-Abadie,* Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair,† Aaron Courville, Yoshua Bengio‡
Département d'informatique et de recherche opérationnelle
Université de Montréal
Montréal, QC H3C 3J7

Abstract

We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions G and D, a unique solution exists, with G recovering the training data distribution and D equal to $\frac{1}{2}$ everywhere. In the case where G and D are defined by multilayer perceptrons, the entire system can be trained with backpropagation. There is no need for any Markov chains or unrolled approximate inference networks during either training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.

1 Introduction

The promise of deep learning is to discover rich, hierarchical models [2] that represent probability distributions over the kinds of data encountered in artificial intelligence applications, such as natural images, audio waveforms containing speech, and symbols in natural language corpora. So far, the most striking successes in deep learning have involved discriminative models, usually those that map a high-dimensional, rich sensory input to a class label [14, 22]. These striking successes have primarily been based on the backpropagation and dropout algorithms, using piecewise linear units [19, 9, 10] which have a particularly well-behaved gradient . Deep *generative* models have had less of an impact, due to the difficulty of approximating many intractable probabilistic computations that arise in maximum likelihood estimation and related strategies, and due to difficulty of leveraging the benefits of piecewise linear units in the generative context. We propose a new generative model estimation procedure that sidesteps these difficulties. [1]

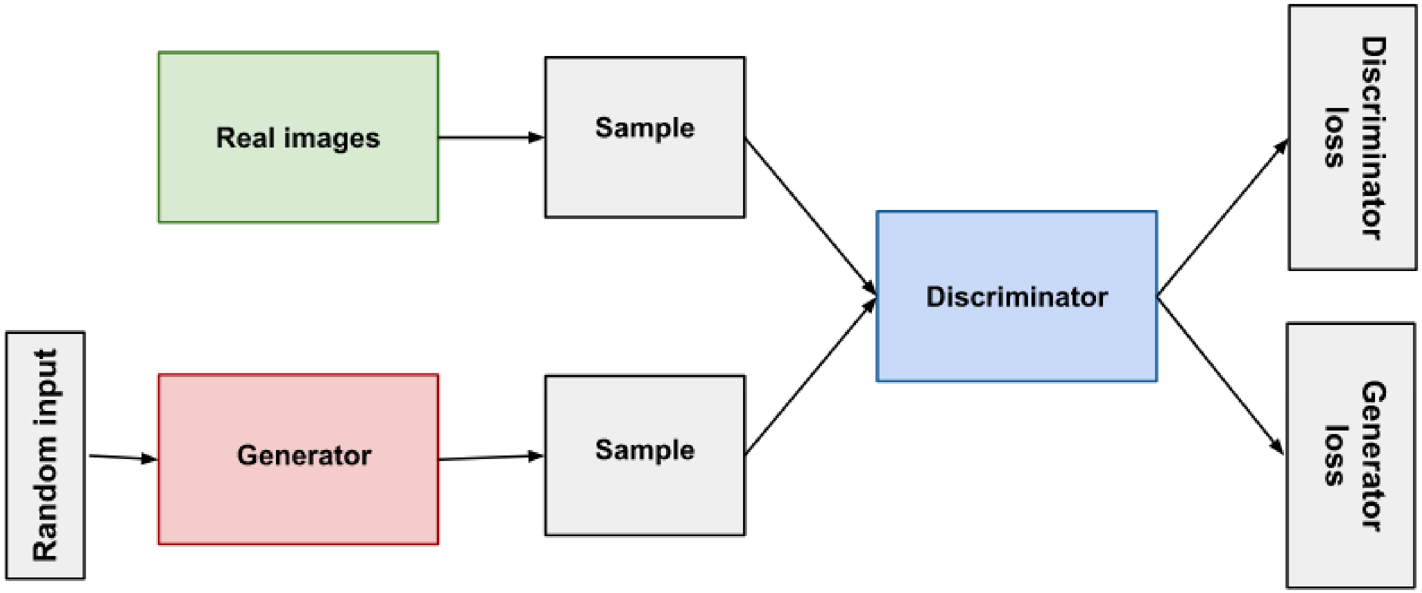
In the proposed *adversarial nets* framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistiguishable from the genuine articles.

*Jean Pouget-Abadie is visiting Université de Montréal from Ecole Polytechnique.
†Sherjil Ozair is visiting Université de Montréal from Indian Institute of Technology Delhi
‡Yoshua Bengio is a CIFAR Senior Fellow.
[1] All code and hyperparameters available at http://www.github.com/goodfeli/adversarial

1

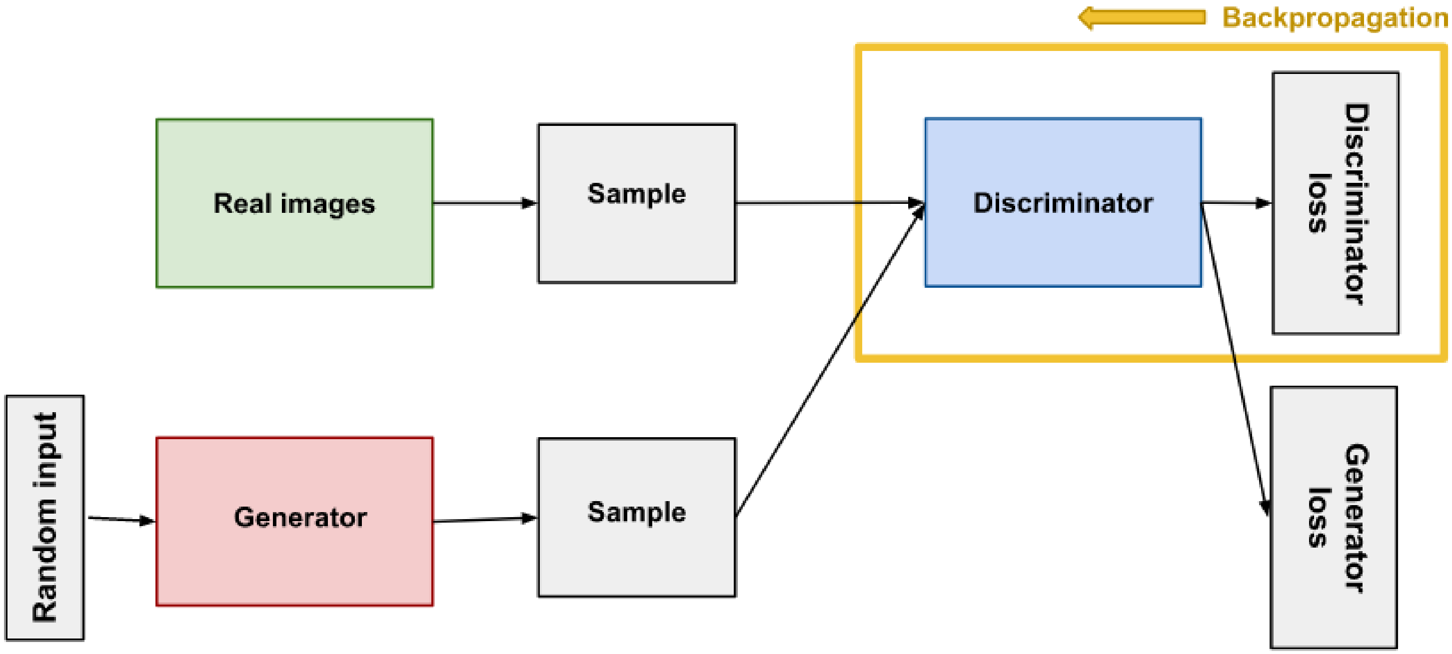
Redes Adversárias Generativas Arquitetura

- Consiste em uma rede discriminativa (D) e outra generativa (G)
- D = {Real | Falso}
- G = Produz imagens sintéticas
- A rede generativa é treinada até confundir a rede discriminativa (False→True)



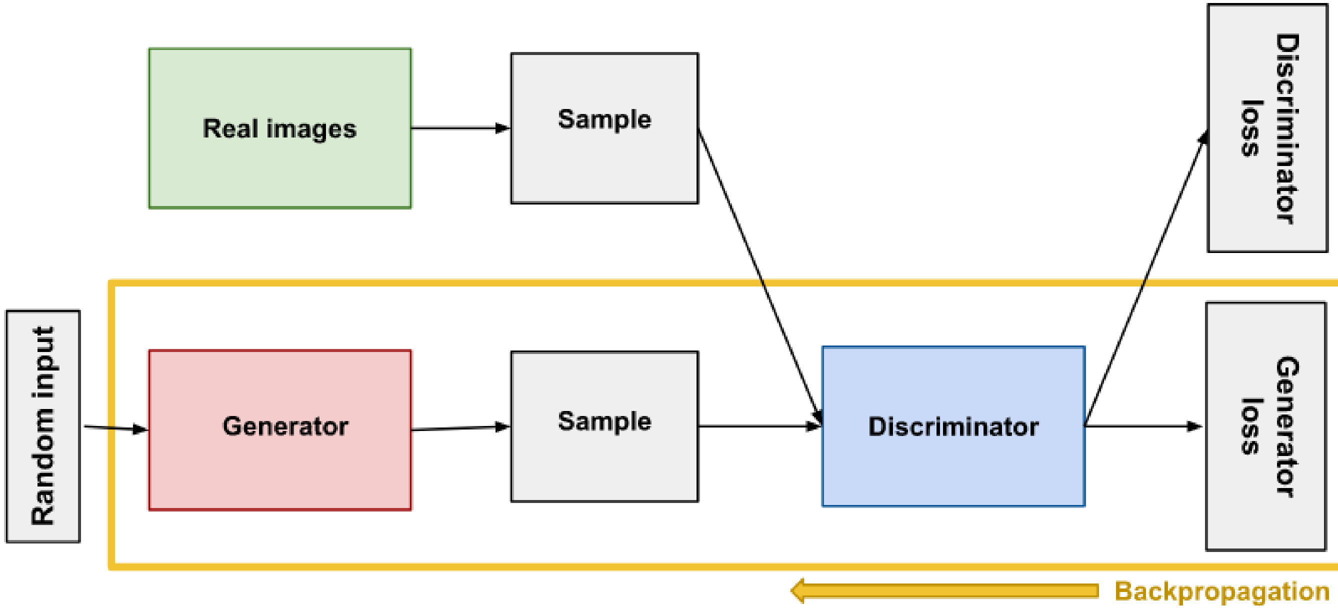
Redes Adversárias Generativas Treinamento

- Modelo Discriminativo
  - 2 Classes (Real | Falso)
  - Modelo Generativo provê imagens falsas



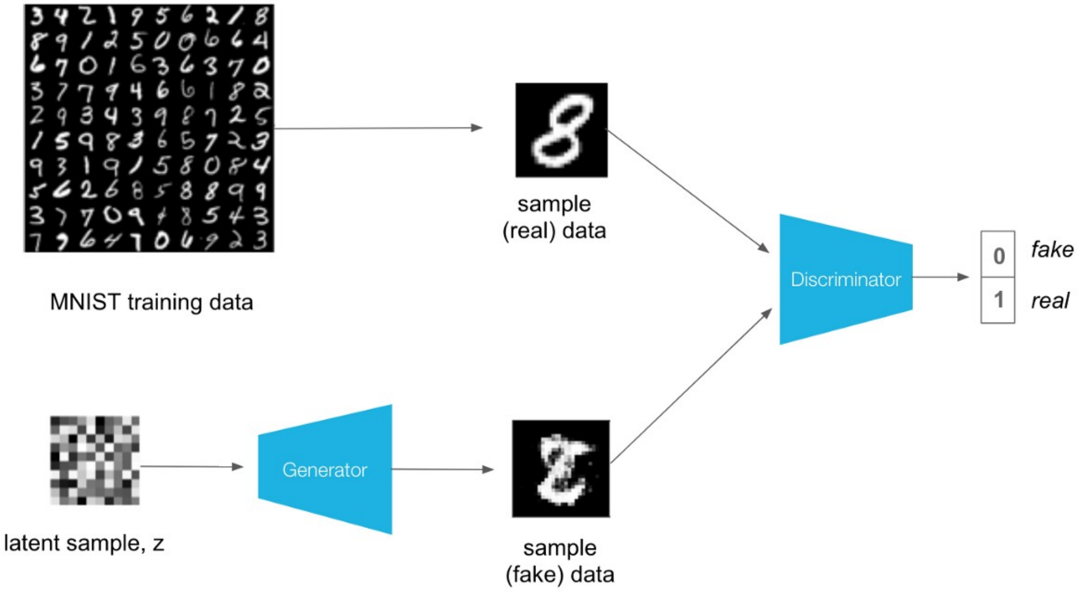
Redes Adversárias Generativas Treinamento

- Modelo Generativo
  - Ruído Z → Imagem NxM
  - Objetivo: Confundir o discriminador (1-loss discriminator)



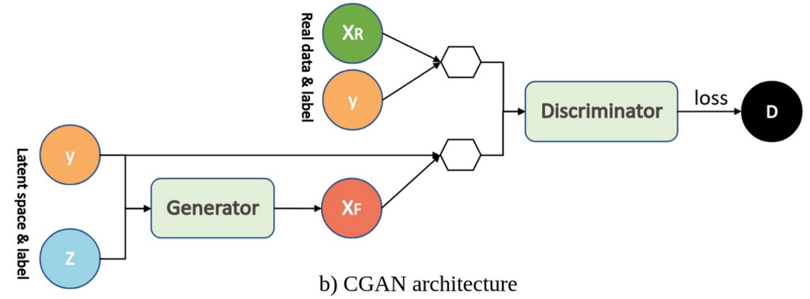
Redes Adversárias Generativas

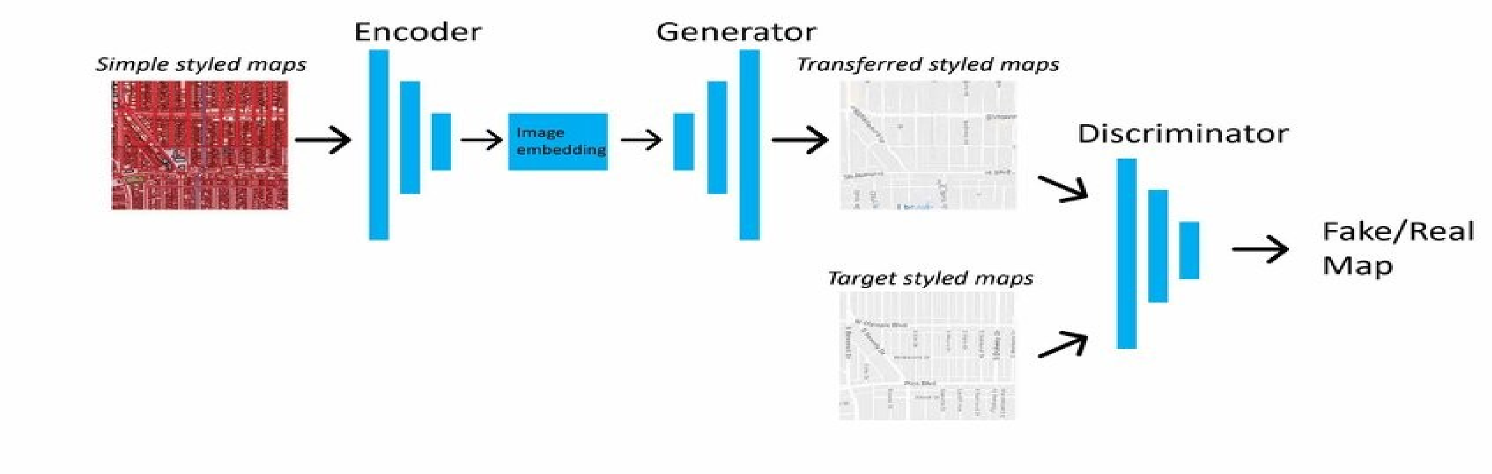
- Tutorial Fake Digits!
  - https://tinyurl.com/37z4xw6x



Arquiteturas e Aplicações
C-GANs

- GANs Condicionais
  - Transferência de contexto
  - Treinamento é feito por um par de dados (Imagem, Label)

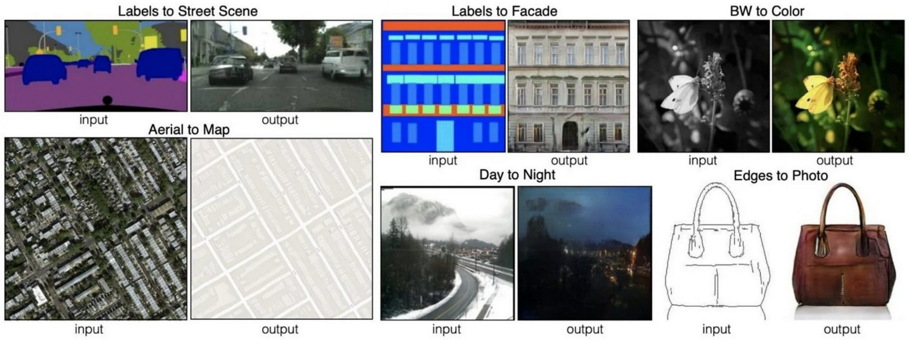


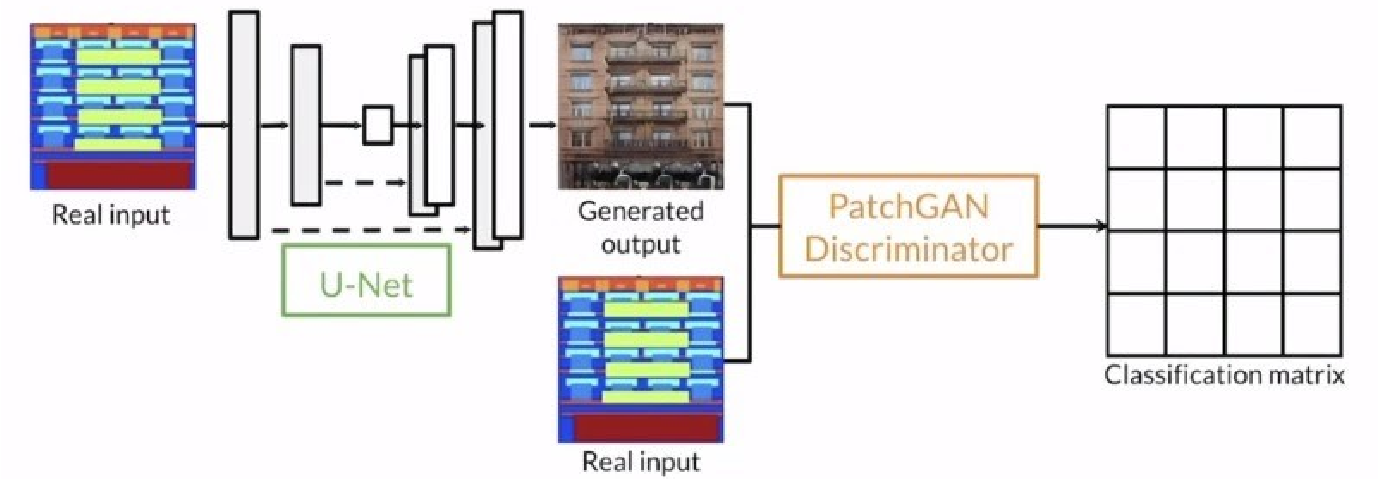


Arquiteturas e Aplicações
C-GANs

- Pix2pix Network

[Phillip Sola - 2017]





Arquiteturas e Aplicações
C-GANs

- Tutorial Pix2Pix – Esboço2Fachadas

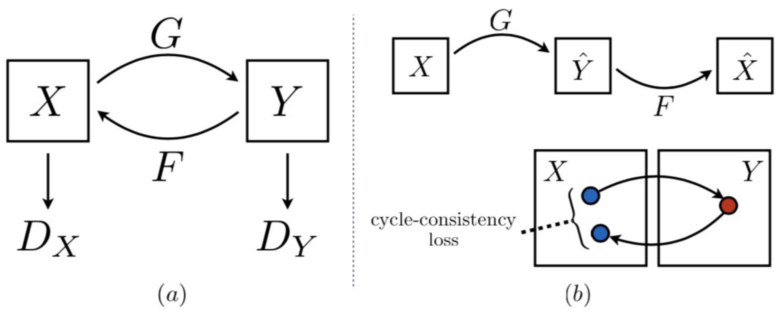
  https://tinyurl.com/yt6dkfw4

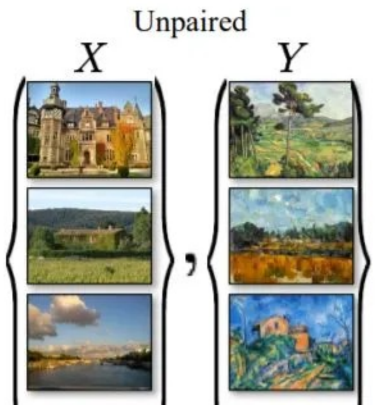
- WebPage Pix2Pix

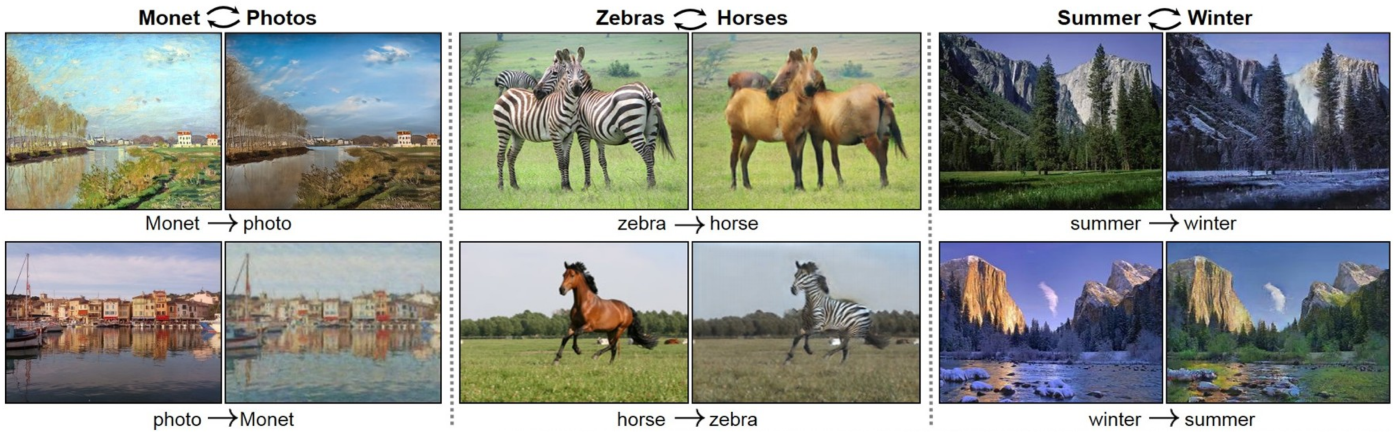
  https://affinelayer.com/pixsrv/

Arquiteturas e Aplicações
Cycles-GANs

- GANs Ciclicas [Jun-Yan Zhu, Taesung Park, **Phillip Isola**, Alexei A. Efros - 2017]
- Transferência de contexto A→B e B→A
- Dataset não é pareado







Arquiteturas e Aplicações
Cycles-GANs

- Tutorial CycleGans – Cavalos e Zebras

  https://tinyurl.com/rmf4e9m9

Arquiteturas e Aplicações

- Tutorial CycleGans – Cavalos e Zebras

  https://tinyurl.com/rmf4e9m9