

Uma Introdução ao Aprendizado Não-Supervisionado

Modelagem Generativa Visual e Síntese de Imagens

Hugo Oliveira¹

¹Departamento de Informática (DPI)

Universidade Federal de Viçosa (UFV)

31 de Janeiro, 2024





Mais informações: <https://sites.google.com/view/oliveirahugo>

Agenda

- 1 Introdução
- 2 Modelos Autorregressivos
- 3 AutoEncoders
 - Variational AutoEncoder (VAE)
- 4 Generative Adversarial Networks

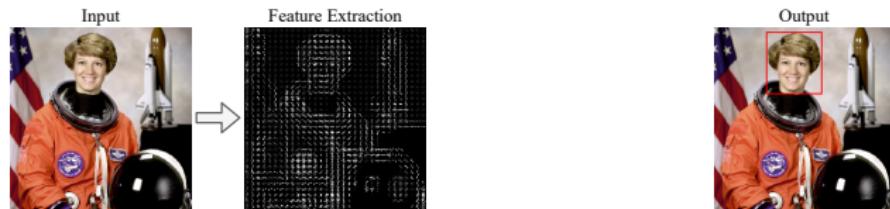
Agenda

- 1 Introdução
- 2 Modelos Autorregressivos
- 3 AutoEncoders
 - Variational AutoEncoder (VAE)
- 4 Generative Adversarial Networks

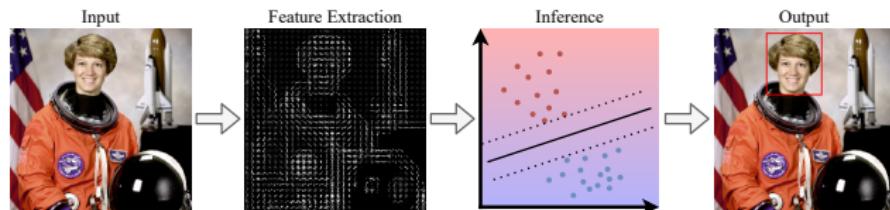
Aprendizado Visual



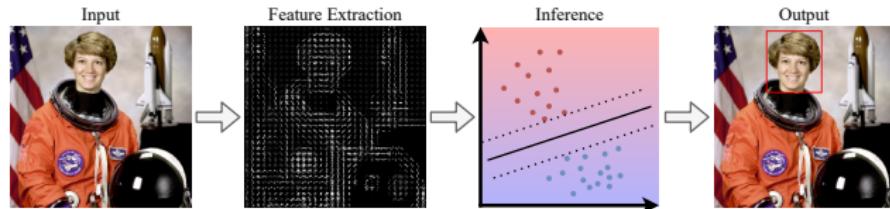
Aprendizado Visual



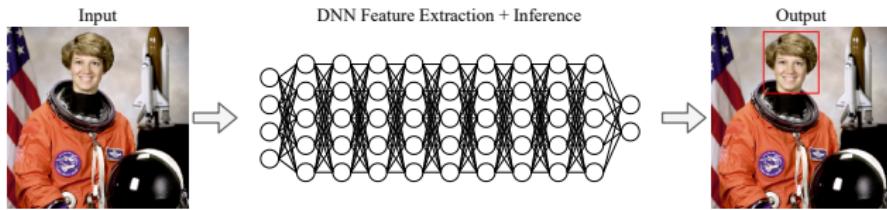
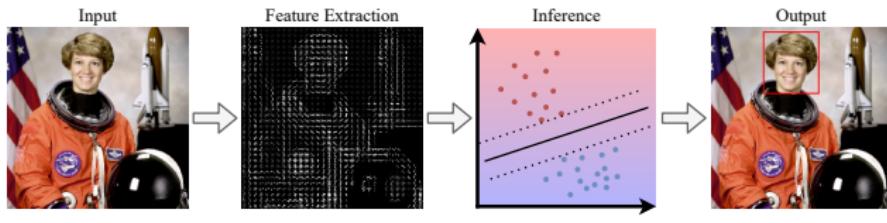
Aprendizado Visual



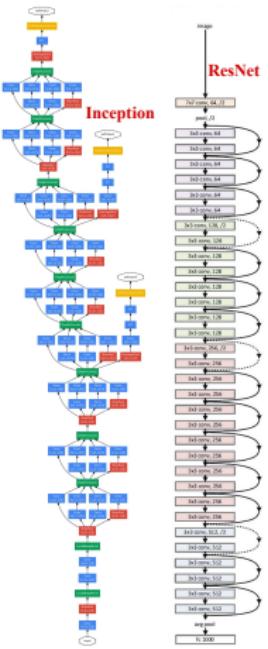
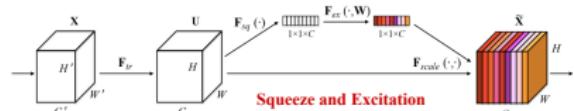
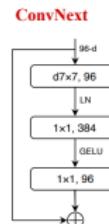
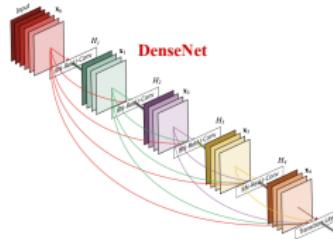
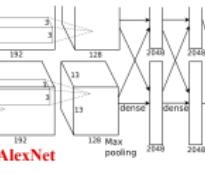
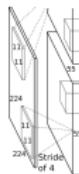
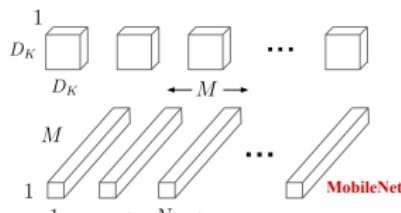
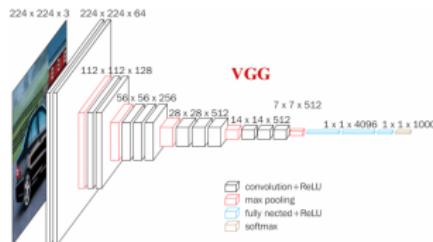
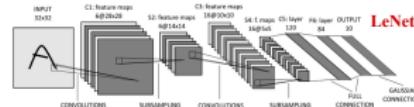
Aprendizado Visual *Shallow* vs. *Deep*



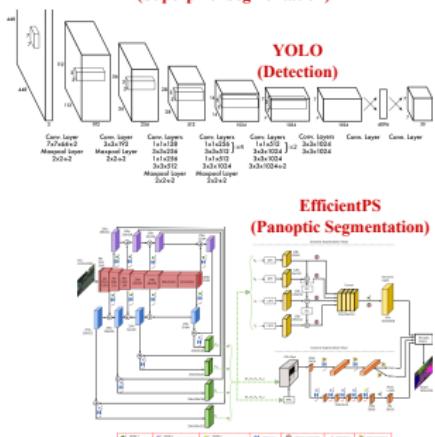
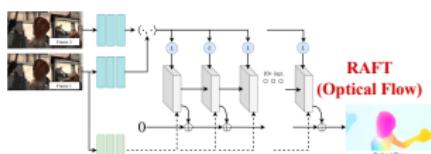
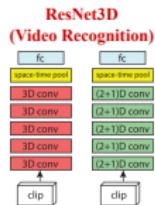
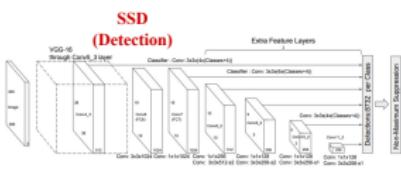
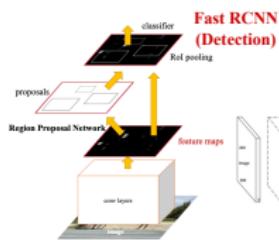
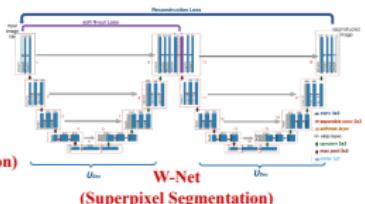
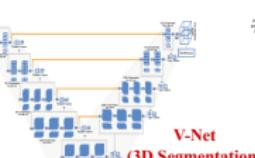
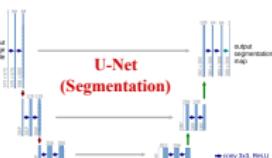
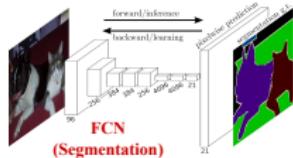
Aprendizado Visual *Shallow* vs. *Deep*



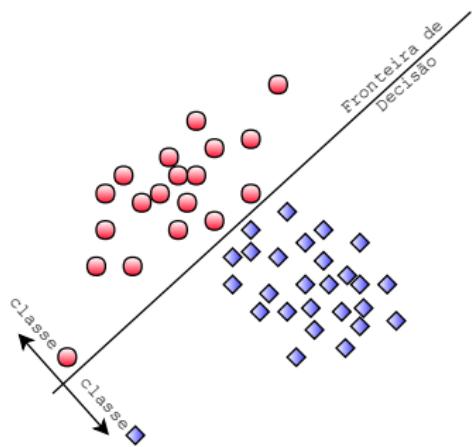
Aprendizado Visual Profundo



Aprendizado Visual Profundo

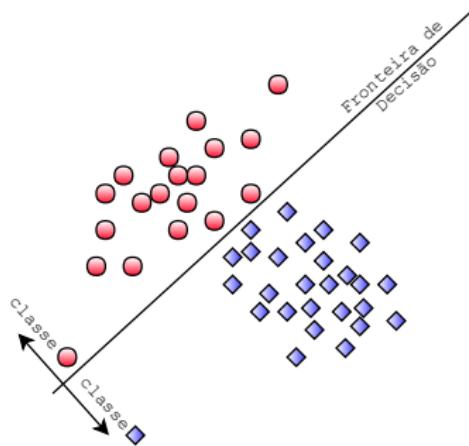


Modelos Discriminativos vs. Modelos Generativos

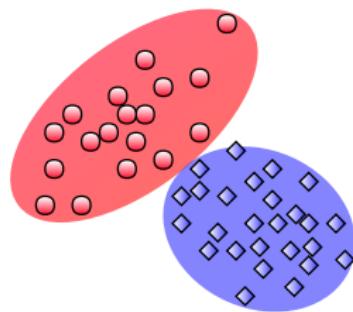


(a) Modelo Discriminativo

Modelos Discriminativos vs. Modelos Generativos



(a) Modelo Discriminativo



(b) Modelo Generativo

Figura: Categorias de modelos em Machine Learning.

Modelos Generativos

Por que estudar modelos generativos?

Modelos Generativos permitem a geração de saídas multimodais. Há tarefas que possuem várias respostas corretas para uma única entrada (i.e. geração de amostras sintéticas, tradução de imagens etc).

Modelos Generativos

Por que estudar modelos generativos?

Modelagem de **tarefas que não possuem uma função de perda bem caracterizada** (i.e. tradução de imagens não-pareadas, geração de imagens fotorrealistas, aprendizado de representações etc).

Modelos Generativos

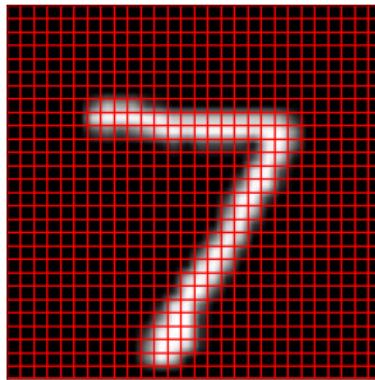
Por que estudar modelos generativos?

Várias tarefas requerem **geração de amostras realísticas** de uma distribuição (i.e. data augmentation, transfer learning etc). Modelos baseados em amostragem via modelos bayesianos de densidade explícita (i.e. MCMC, Gibbs, Metropolis Hastings, etc) possuem performance limitada em aplicações reais.

Modelos Generativos

Por que estudar modelos generativos?

Modelagem de distribuições probabilísticas com um **alto número de dimensões**¹.



¹<https://arxiv.org/abs/2310.11448>

Modelos Generativos

Por que estudar modelos generativos?

Modelagem de distribuições probabilísticas com um **alto número de dimensões**¹.



¹<https://arxiv.org/abs/2310.11448>

Modelos Generativos

Por que estudar modelos generativos?

Reinforcement Learning¹: Um Modelo Generativo treinado para planejamento pode aprender uma distribuição condicional sobre estados futuros do mundo, dado o estado atual e ações hipotéticas do agente.

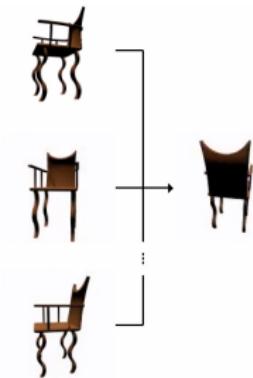
¹<https://skymind.ai/wiki/deep-reinforcement-learning>

²<https://3d-diffusion.github.io/>

Modelos Generativos

Por que estudar modelos generativos?

Reinforcement Learning¹: Um Modelo Generativo treinado para planejamento pode aprender uma distribuição condicional sobre estados futuros do mundo, dado o estado atual e ações hipotéticas do agente.



¹<https://skymind.ai/wiki/deep-reinforcement-learning>

²<https://3d-diffusion.github.io/>

Modelos Generativos

Por que estudar modelos generativos?

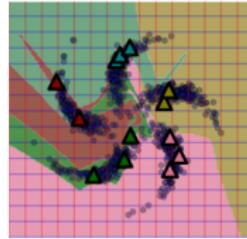
Aprendizado Semi-Supervisionado: Modelos Generativos podem ser treinados com dados incompletos.

¹<http://proceedings.mlr.press/v119/izmailov20a.html>

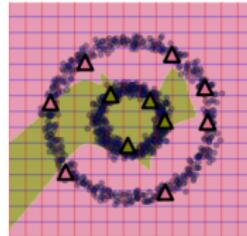
Modelos Generativos

Por que estudar modelos generativos?

Aprendizado Semi-Supervisionado: Modelos Generativos podem ser treinados com dados incompletos.



Supervisionado

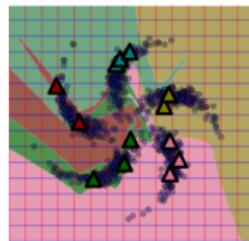


¹<http://proceedings.mlr.press/v119/izmailov20a.html>

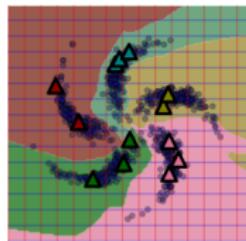
Modelos Generativos

Por que estudar modelos generativos?

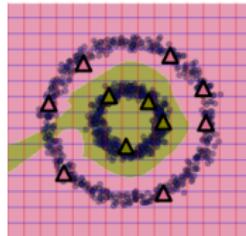
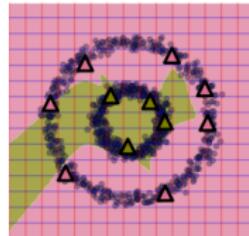
Aprendizado Semi-Supervisionado: Modelos Generativos podem ser treinados com dados incompletos.



Supervisionado



Semi-Supervisionado



¹<http://proceedings.mlr.press/v119/izmailov20a.html>

Modelos Generativos

Por que estudar modelos generativos?



Figura: Superresolução em imagens¹.

¹[https://medium.com/mlarning-ai/diffusion-based-super-resolution-the-sr3-approach-0ef890d5f90f](https://medium.com/mlearning-ai/diffusion-based-super-resolution-the-sr3-approach-0ef890d5f90f)

Modelos Generativos

Por que estudar modelos generativos?

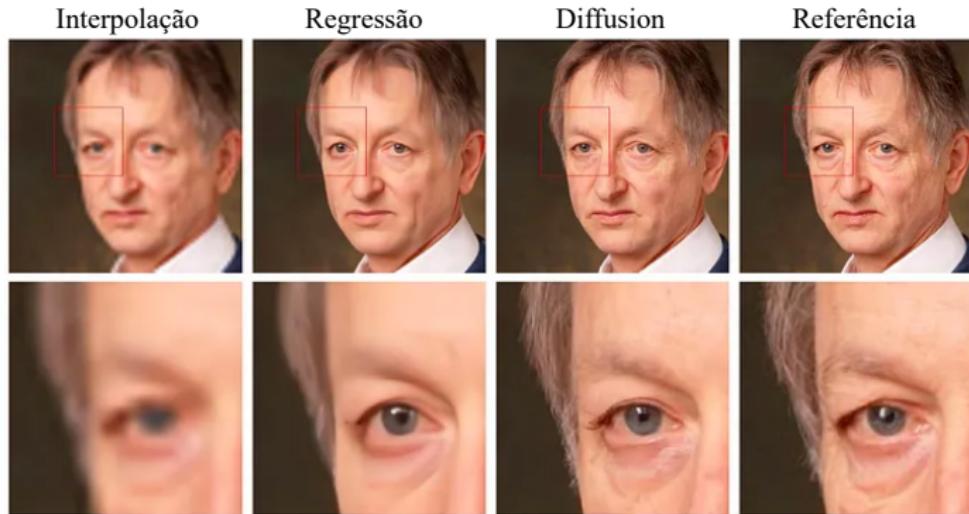


Figura: Superresolução em imagens¹.

¹[https://medium.com/mlarning-ai/diffusion-based-super-resolution-the-sr3-approach-0ef890d5f90f](https://medium.com/mlearning-ai/diffusion-based-super-resolution-the-sr3-approach-0ef890d5f90f)

Modelos Generativos

Por que estudar modelos generativos?



Figura: Síntese de imagens¹.

¹<https://stability.ai/news/stable-diffusion-public-release>

Modelos Generativos

Por que estudar modelos generativos?



Figura: Síntese de imagens¹.

¹<https://stability.ai/news/stable-diffusion-public-release>

Modelos Generativos

Por que estudar modelos generativos?



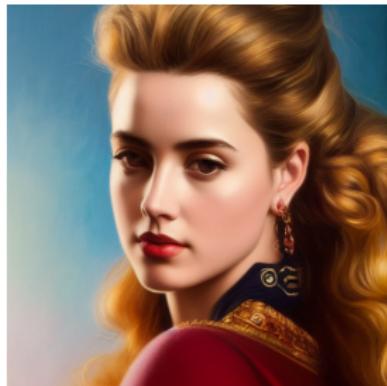
Original

Figura: Transferência de estilo¹.

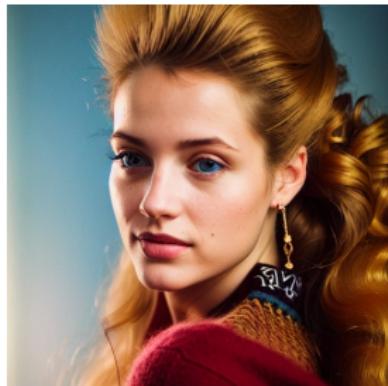
¹<https://stable-diffusion-art.com/stylize-images/>

Modelos Generativos

Por que estudar modelos generativos?



Original



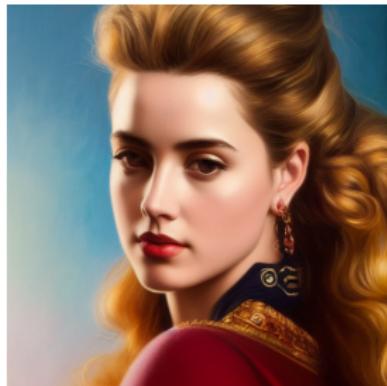
Fotorrealista

Figura: Transferência de estilo¹.

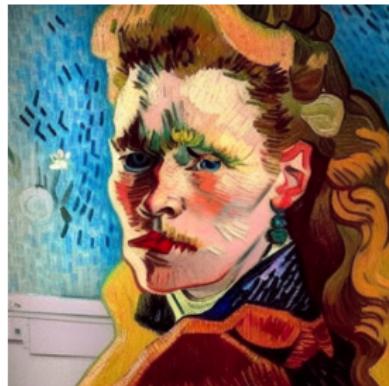
¹<https://stable-diffusion-art.com/stylize-images/>

Modelos Generativos

Por que estudar modelos generativos?



Original



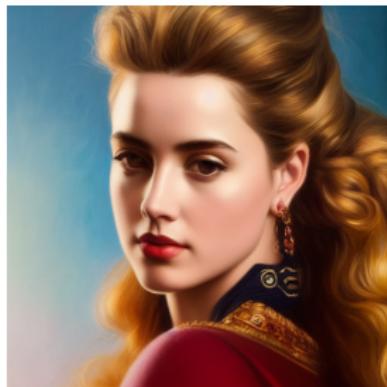
Van Gogh

Figura: Transferência de estilo¹.

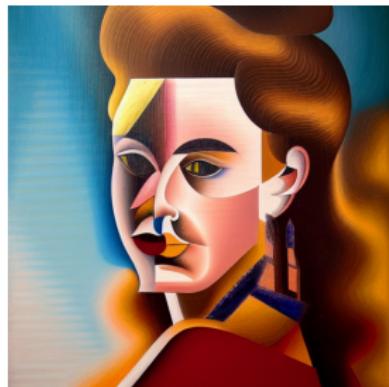
¹<https://stable-diffusion-art.com/stylize-images/>

Modelos Generativos

Por que estudar modelos generativos?



Original



Cubista

Figura: Transferência de estilo¹.

¹<https://stable-diffusion-art.com/stylize-images/>

Modelos Generativos

Por que estudar modelos generativos?



Original



Impressionista

Figura: Transferência de estilo¹.

¹<https://stable-diffusion-art.com/stylize-images/>

Modelos Generativos

Por que estudar modelos generativos?

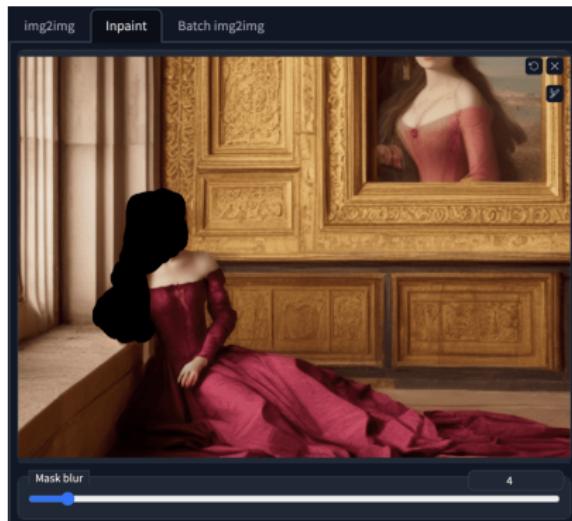


Figura: Inpainting¹ via Diffusion.

Modelos Generativos

Por que estudar modelos generativos?

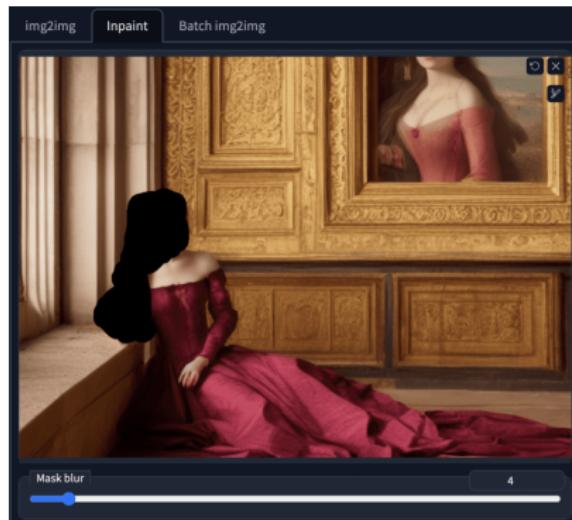


Figura: Inpainting¹ via Diffusion.

Modelos Generativos

Por que estudar modelos generativos?



Figura: Outpainting¹ via Diffusion.

¹<https://stable-diffusion-art.com/outpainting/>

Modelos Generativos

Por que estudar modelos generativos?



Figura: Outpainting¹ via Diffusion.

¹<https://stable-diffusion-art.com/outpainting/>

Conceitos de Modelagem Generativa

Entendendo Modelos Generativos

Como treinar modelos generativos para realizar todas essas tarefas?

Conceitos de Modelagem Generativa

Entendendo Modelos Generativos

Como treinar modelos generativos para realizar todas essas tarefas?

Famílias de Modelos Generativos

Existem múltiplas abordagens para se realizar modelagem generativa, incluindo:

- Amostragem;
- Modelos autorregressivos;
- Reconstrução;
- Inferência variacional;
- Modelos adversariais;
- Diffusion;

Maximum Likelihood Estimation

- Estimativa de Máxima Verossimilhança (MLE) serve de base teórica para modelagens generativas
- Combina evidências (dados) com distribuições *a priori*
- Modelo de estimação não-enviesado com a menor variância possível

Maximum Likelihood Estimation

Maximum Likelihood Estimation

Nem todo modelo gerativo é baseado em MLE (i.e. GANs), mas entender esse método de estimação é imprescindível para compreender a área de Modelos Generativos.

Maximum Likelihood Estimation

- Modelo estatístico: p_{model}
- Parâmetros do modelo: θ
- Amostras de treino: $x^{(i)}, i = 1, 2, \dots, m$
- Modelagem: $\prod_{i=1}^m p_{model}(x^{(i)}; \theta)$

Maximum Likelihood Estimation

- Modelo estatístico: p_{model}
- Parâmetros do modelo: θ
- Amostras de treino: $x^{(i)}, i = 1, 2, \dots, m$
- Modelagem: $\prod_{i=1}^m p_{model}(x^{(i)}; \theta)$

MLE

$$\theta^* = \arg \max_{\theta} \prod_{i=1}^m p_{model}(x^{(i)}; \theta)$$

Maximum Likelihood Estimation

- Modelo estatístico: p_{model}
- Parâmetros do modelo: θ
- Amostras de treino: $x^{(i)}, i = 1, 2, \dots, m$
- Modelagem: $\prod_{i=1}^m p_{model}(x^{(i)}; \theta)$

MLE

$$\theta^* = \arg \max_{\theta} \sum_{i=1}^m \log p_{model}(x^{(i)}; \theta)$$

Maximum Likelihood Estimation

MLE e a Divergência de Kullback-Leibler

É possível entender MLE como uma minimização da **Divergência de Kullback-Leibler** entre a distribuição p_{data} dos dados e a distribuição p_{model} do modelo.

Maximum Likelihood Estimation

Limitações

- A modelagem requer uma representação explícita de p_{model}
- Funciona apenas para funções “comportadas” (i.e. deriváveis, contínuas...)
- Difícil (se não impossível) de calcular para distribuições de probabilidade com muitas dimensões

Maximum Likelihood Estimation

Limitações

- A modelagem requer uma representação explícita de p_{model}
- Funciona apenas para funções “comportadas” (i.e. deriváveis, contínuas...)
- Difícil (se não impossível) de calcular para distribuições de probabilidade com muitas dimensões

O que fazer?

Como estimar essas distribuições mais difíceis de serem modeladas?

Taxonomia de Modelos Generativos

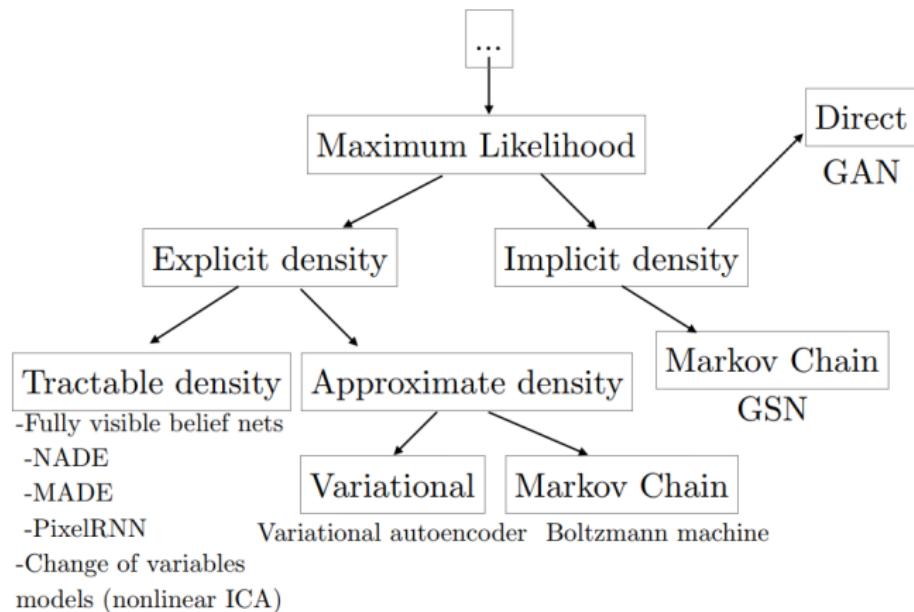


Figura: Uma taxonomia (incompleta) de modelos gerativos¹.

¹<https://arxiv.org/abs/1701.00160>

Agenda

1 Introdução

2 Modelos Autorregressivos

3 AutoEncoders

- Variational AutoEncoder (VAE)

4 Generative Adversarial Networks

Taxonomia de Modelos Generativos

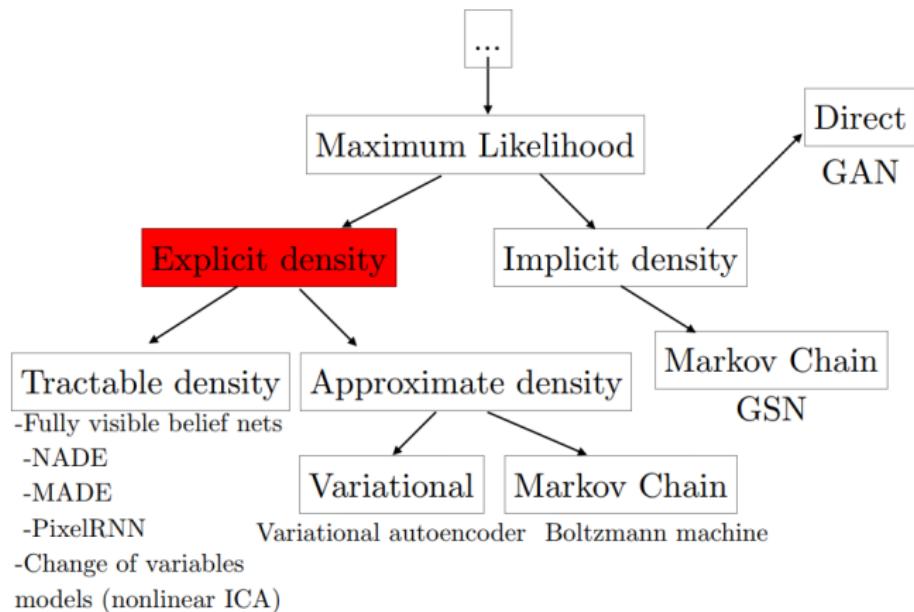


Figura: Uma taxonomia (incompleta) de modelos gerativos¹.

¹<https://arxiv.org/abs/1701.00160>

Taxonomia de Modelos Generativos

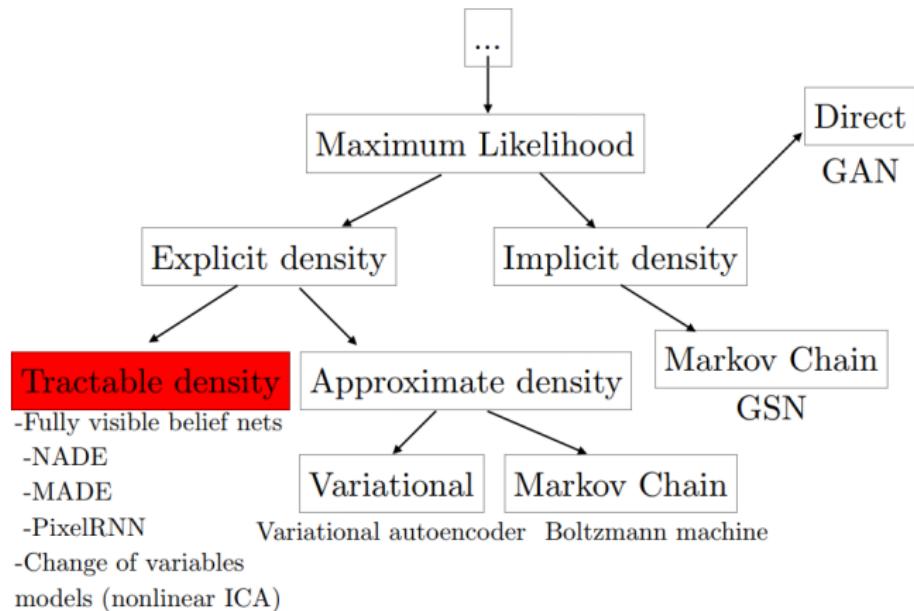


Figura: Uma taxonomia (incompleta) de modelos gerativos¹.

¹<https://arxiv.org/abs/1701.00160>

Modelos Autorregressivos

Como Funcionam?

Possuem uma função de densidade $p_{model}(x; \theta)$ explicitamente descrita no modelo que será aproximada de acordo com os dados.

Modelos Autorregressivos

Como Funcionam?

Possuem uma função de densidade $p_{model}(x; \theta)$ explicitamente descrita no modelo que será aproximada de acordo com os dados.

Dificuldades Associadas

A maior dificuldade é em criar modelos que capturem toda a **complexidade dos dados** que estão sendo modelados deixando os modelos tratáveis computacionalmente.

Modelos Autorregressivos

- PixelRNNs e PixelCNNs¹
 - Construção cuidadosa de modelos que tenham estruturas pensadas para garantir sua tratabilidade

¹<https://arxiv.org/abs/1601.06759>

Modelos Autorregressivos

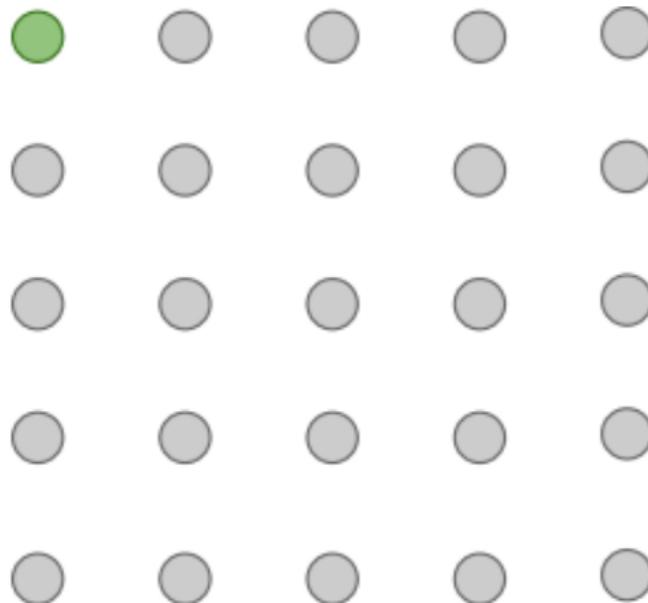


Figura: Geração sequencial de pixels numa PixelRNN.

Modelos Autorregressivos

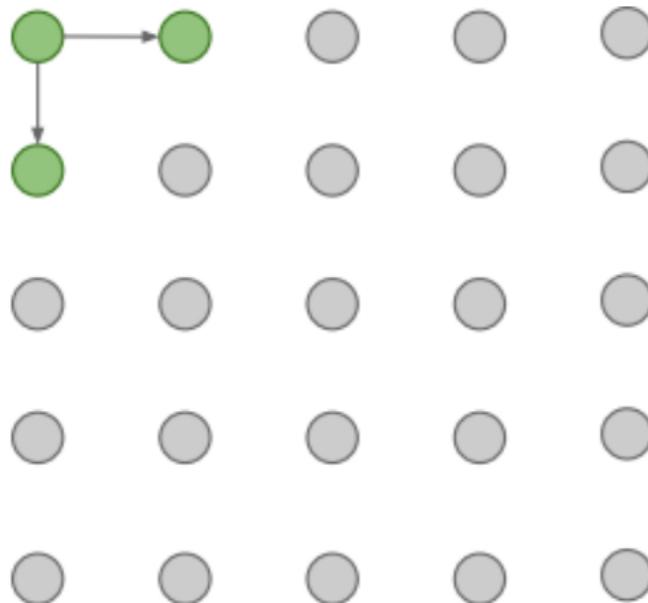


Figura: Geração sequencial de pixels numa PixelRNN.

Modelos Autorregressivos

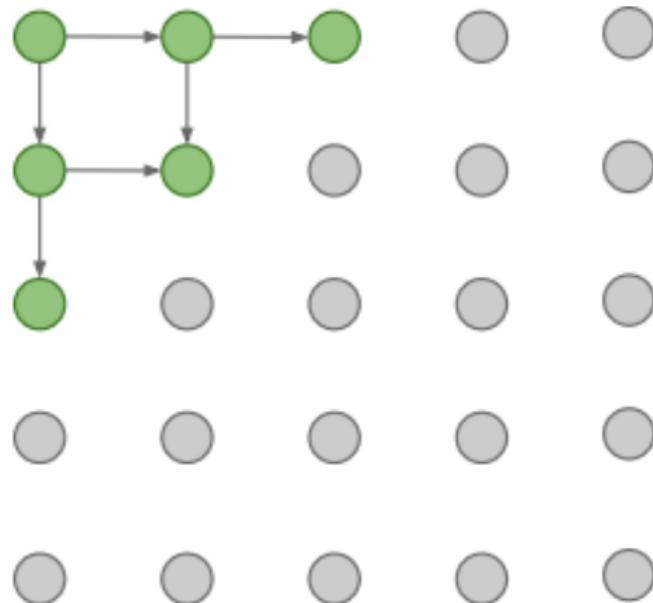


Figura: Geração sequencial de pixels numa PixelRNN.

Modelos Autorregressivos

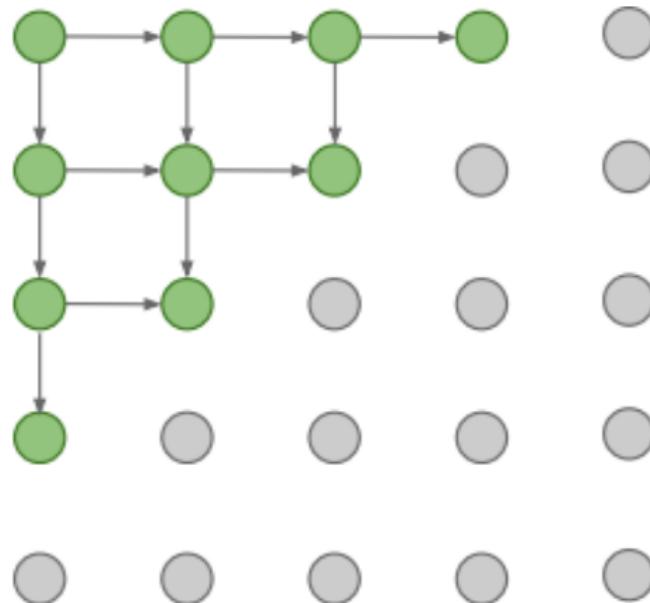


Figura: Geração sequencial de pixels numa PixelRNN.

Modelos Autorregressivos

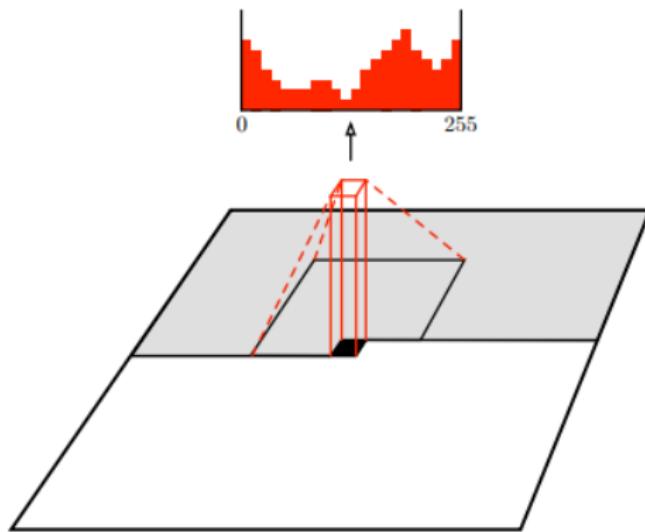


Figura: Geração sequencial de pixels numa PixelCNN.

Modelos Autorregressivos

- Resultados de PixelRNNs e PixelCNNs

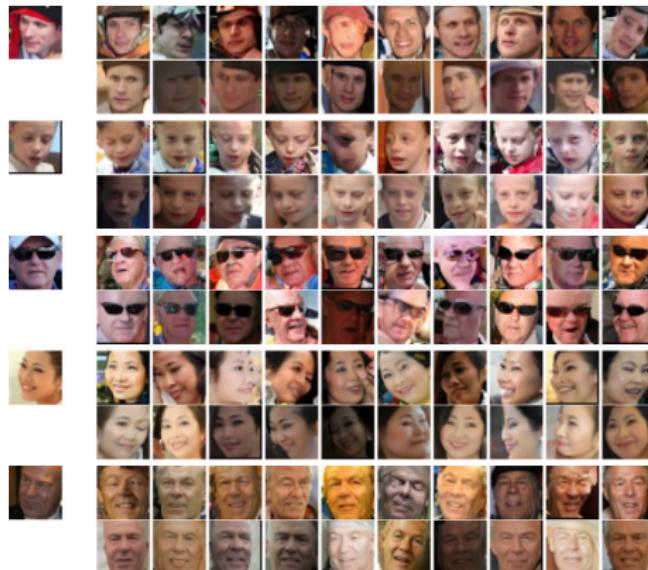


Figura: Samples geradas por meio de PixelCNNs.

Modelos Autorregressivos

Limitações

- A geração pixel-a-pixel de amostras é um **processo muito lento**
- Geração de **amostras não-realistas** devido à falta de contexto global

Agenda

- 1 Introdução
- 2 Modelos Autorregressivos
- 3 AutoEncoders
 - Variational AutoEncoder (VAE)
- 4 Generative Adversarial Networks

AutoEncoders

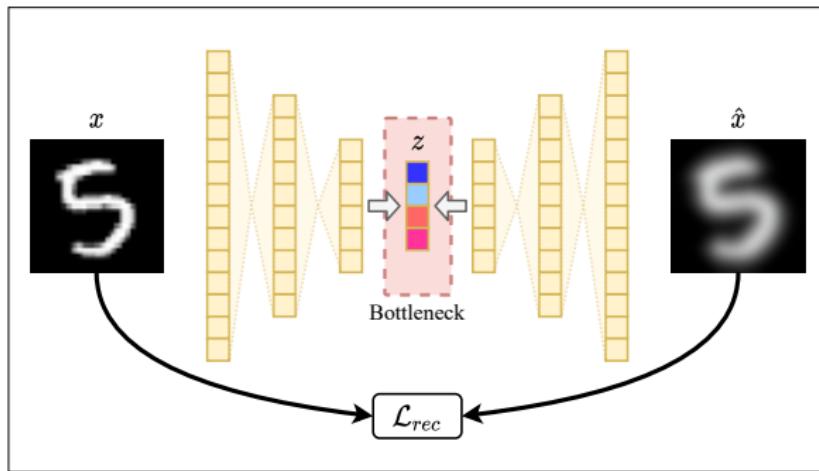


Figura: Arquitetura de um AutoEncoder linear.

AutoEncoders

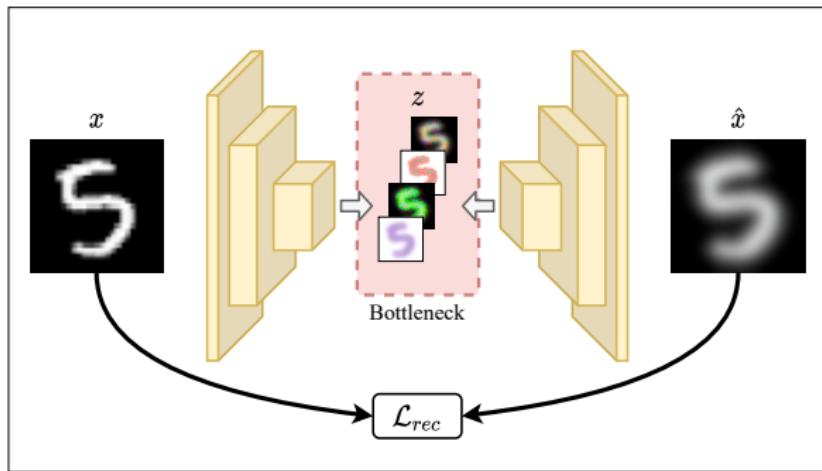


Figura: Arquitetura de um AutoEncoder convolucional.

Taxonomia de Modelos Generativos

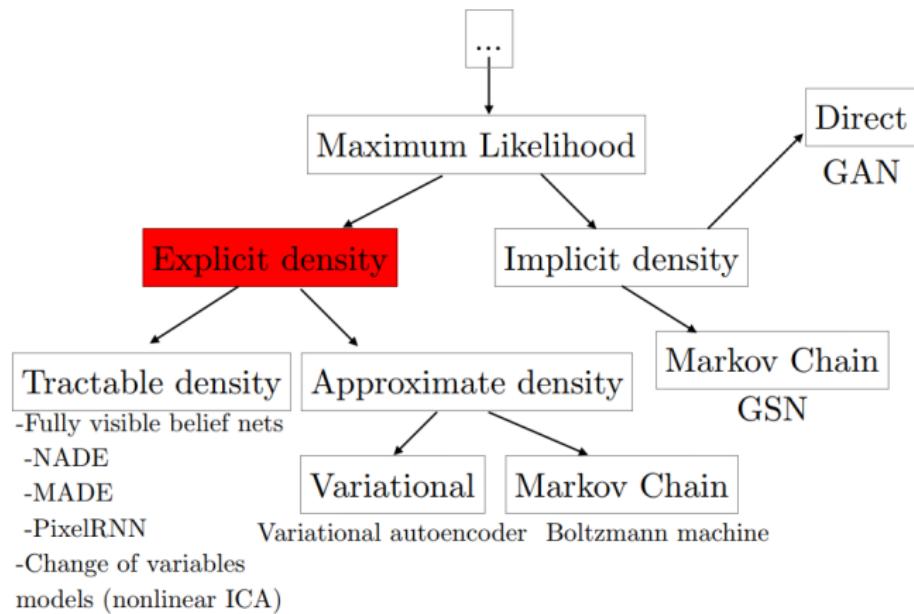


Figura: Uma taxonomia (incompleta) de modelos generativos¹.

¹<https://arxiv.org/abs/1701.00160>

Taxonomia de Modelos Generativos

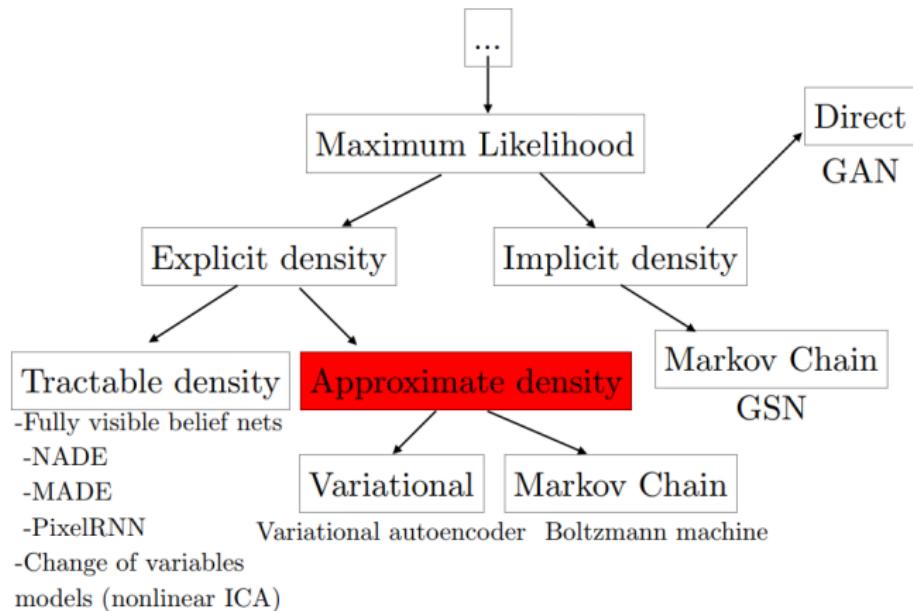


Figura: Uma taxonomia (incompleta) de modelos generativos¹.

¹<https://arxiv.org/abs/1701.00160>

Modelos de Densidade Explícita

Voltando à Síntese de Imagens

Uma outra forma de explicitar a distribuição dos dados num modelo gerativo é admitir **aproximações tratáveis da MLE** e de seus gradientes. Uma espécie de AE chamado Variational AE (VAE)¹ se propõe a modelar dados utilizando essa estratégia.

¹<https://arxiv.org/abs/1312.6114>

Variational AutoEncoder

- Síntese de amostras
- Inferência variacional¹
 - Limite inferior da MLE
- Amostragem a partir de uma distribuição gaussiana
 - Simples de implementar
 - VAE aprende a mapear a gaussiana para variáveis latentes do dado

¹<http://pedrounb.blogspot.com/2017/12/introducao-inferencia-variacional.html>

Variational AutoEncoder

AE para Geração de Amostras

É possível treinar um AE para realizar compressão, remoção de ruído, sparse coding...

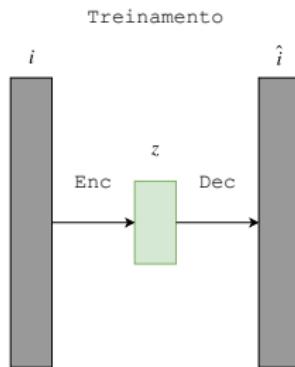


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Será que é possível amostrar valores para o vetor z e gerar novas amostras?

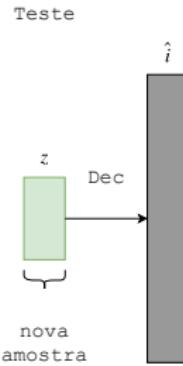


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Será que é possível amostrar valores para o vetor z e gerar novas amostras?

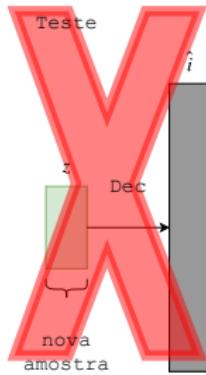


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

AEs comuns oferecem pouco controle sobre como as informações estão sendo representadas em z .

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Variational AEs (VAEs) permitem um maior controle sobre z e, consequentemente, a geração de novas amostras por meio da amostragem de vetores latentes z .

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Para fazer uma amostragem de um valor $z_i \in z$ de acordo com uma Distribuição Gaussiana ($z_i \sim N(\mu_i, \sigma_i)$), é preciso otimizar os parâmetros μ_i e σ_i dessa gaussiana.

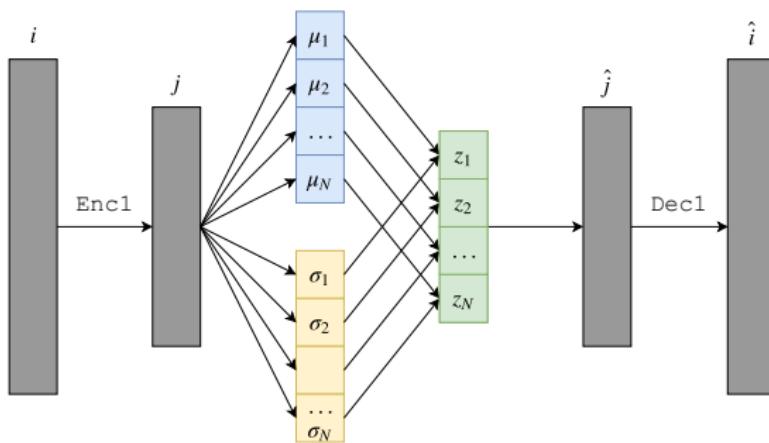


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Como fazer o backpropagation através de valores amostrados de uma distribuição estatística?

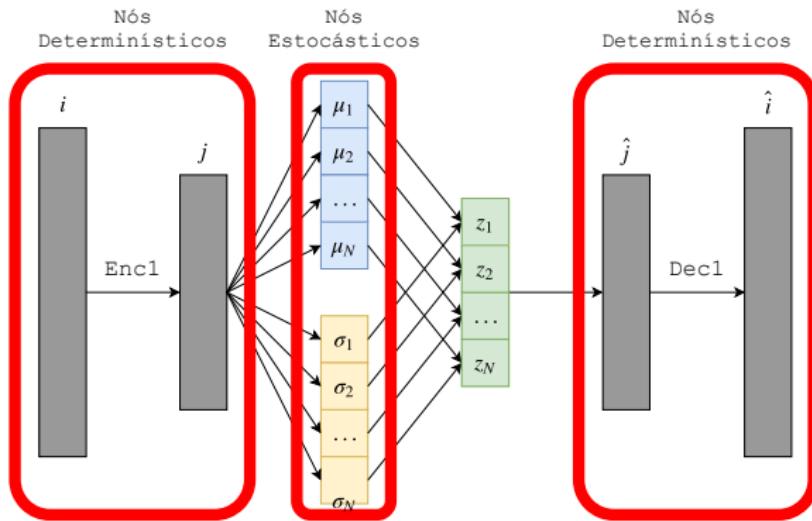


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Nessa arquitetura o backpropagation é impossível.

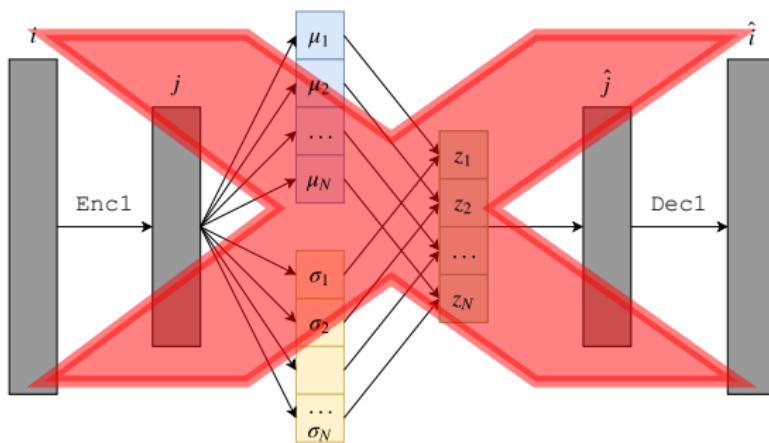


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Para permitir o backpropagation é usado o truque da **reparametrização**.

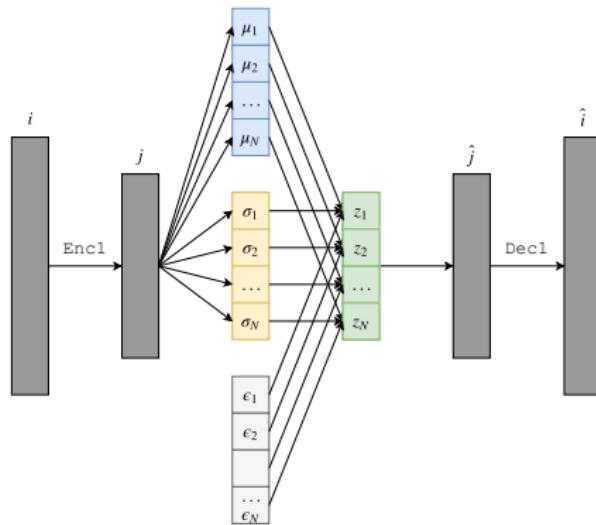


Figura: Geração de amostras em um AE.

Variational AutoEncoder

AE para Geração de Amostras

Para permitir o backpropagation é usado o truque da **reparametrização**.

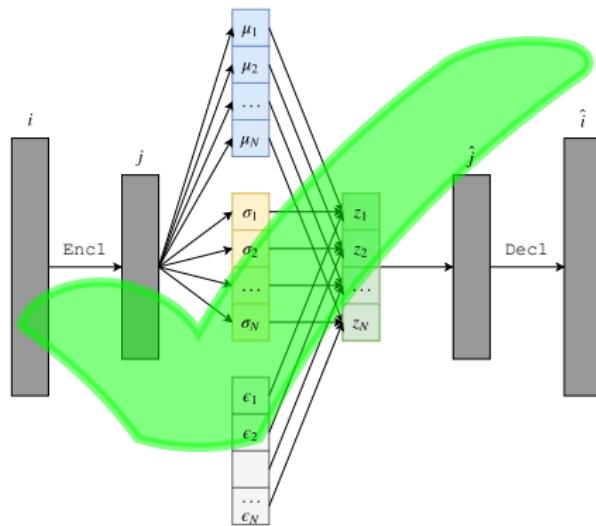


Figura: Geração de amostras em um AE.

Variational AutoEncoder

AE para Geração de Amostras

Para permitir o backpropagation é usado o truque da **reparametrização**.

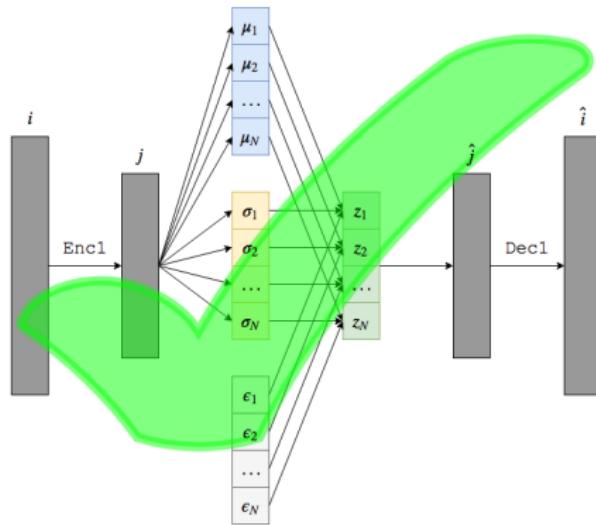


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

AE para Geração de Amostras

Depois de treinada a arquitetura, basta amostrar valores de $\epsilon_i, i = 1, 2, \dots, N$ para gerar amostras novas da distribuição que foi modelada.

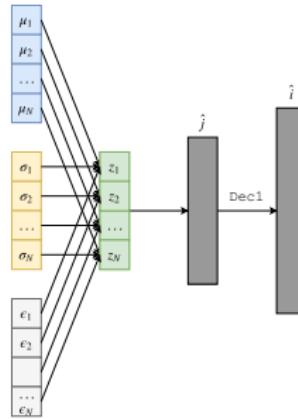


Figura: Geração de amostras em um AE.

Variational AutoEncoder (VAE)

Interpretação do Vetor z

Os valores do vetor latente z tendem a mapear informações de alto nível semântico nos dados.

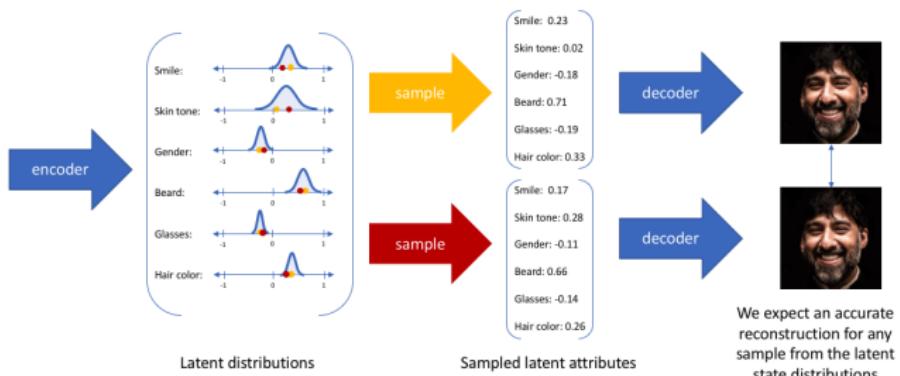


Figura: Gaussianas em um VAE¹.

¹<https://www.jeremyjordan.me/variational-autoencoders/>

Variational AutoEncoder (VAE)



Figura: Variando o espaço de z em um VAE.

Variational AutoEncoder (VAE)

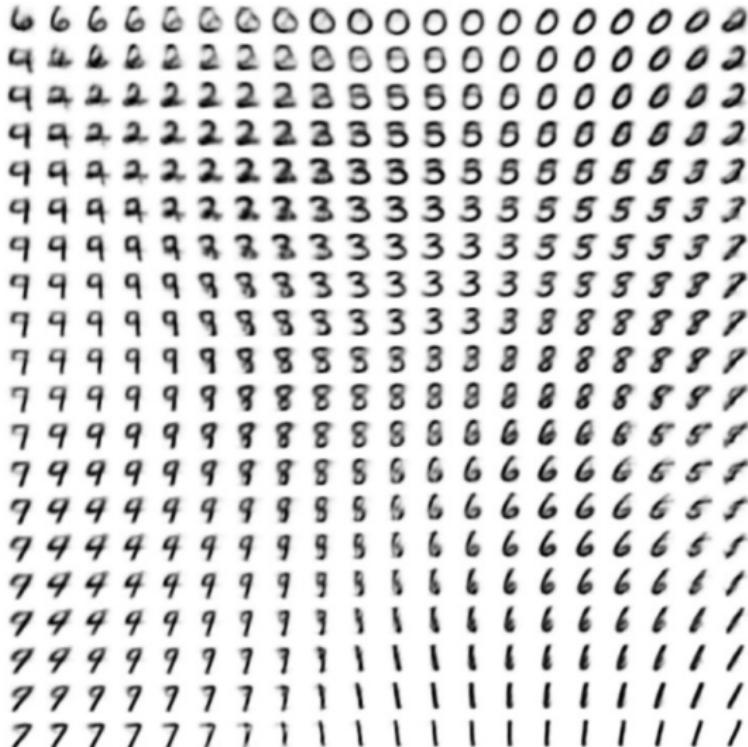


Figura: Variando o espaço de z em um VAE.

VAEs Modernos

- Efficient-VDVAE¹
 - VAEs recentes têm resultados visuais similares ou até superiores ao estado-da-arte de outras metodologias de síntese de imagens deep (i.e. GANs ou até Diffusion)



¹<https://arxiv.org/abs/2203.13751>

VAEs

Demo - VAEs

variational_autoencoder.ipynb

VAEs

Limitações

- Qualidade visual tende a ser inferior a abordagens como GANs e Diffusion
- Funcionam melhor em dados mais comportados (i.e. imagens pouco complexas de rostos, animais, comidas)
- Pouco generalizável para amostras out-of-distribution

Agenda

- 1 Introdução
- 2 Modelos Autorregressivos
- 3 AutoEncoders
 - Variational AutoEncoder (VAE)
- 4 Generative Adversarial Networks

Modelagem Generativa

Abandonando a MLE

Embora modelos autorregressivos (modernos) e principalmente VAEs sejam métodos úteis para modelagem de dados tentando aproximar diretamente a MLE, é possível atingir resultados ainda melhores ao se abandonar essa estratégia. Podemos tentar **modelar a distribuição dos dados implicitamente**, ao invés de calcular a MLE (ou uma aproximação para ela) explicitamente.

Taxonomia de Modelos Generativos

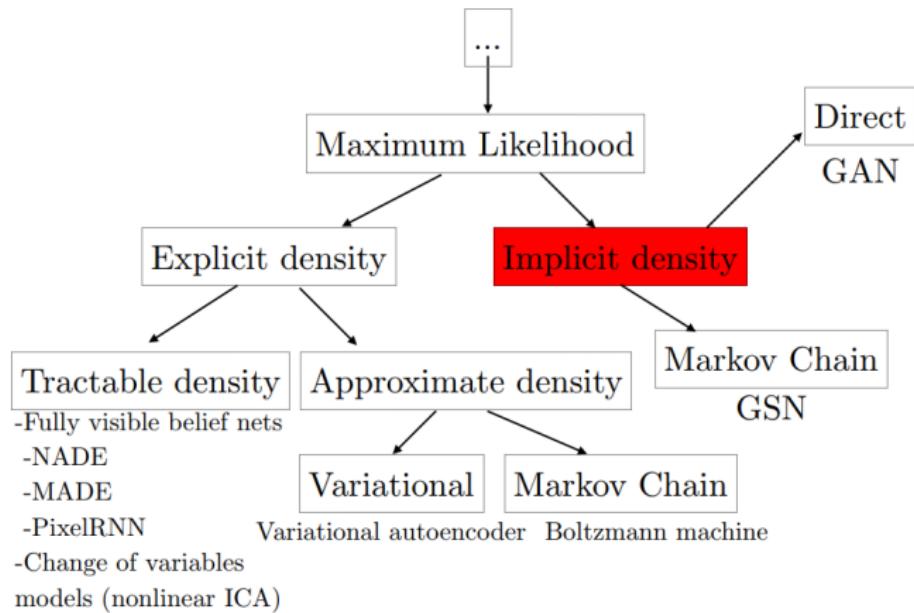


Figura: Uma taxonomia (incompleta) de modelos generativos¹.

¹<https://arxiv.org/abs/1701.00160>

Generative Adversarial Networks

- Aprendizado indireto da distribuição dos dados
- Duas redes com objetivos opostos
- Convergência conjunta
- Background matemático de teoria dos jogos

Generative Adversarial Networks

- Rede Discriminativa (D)
 - Treinada para discernir entre amostras reais do dataset de treino e amostras falsas vindas de G
 - Tem como objetivo discriminar entre amostras reais e falsas

Generative Adversarial Networks

- Rede Discriminativa (D)
 - Treinada para discernir entre amostras reais do dataset de treino e amostras falsas vindas de G
 - Tem como objetivo discriminar entre amostras reais e falsas
- Rede Generativa (G)
 - Mapeia a distribuição de um vetor latente Z para a distribuição das amostras de treino
 - Tem o objetivo de criar amostras cada vez mais verossímeis

Generative Adversarial Networks

- As funções de perda de cada uma das redes devem ser cuidadosamente desenhadas
 - Minimax Loss (MM GAN)
 - Non-Saturating Loss (NS GAN)
 - Wasserstein Distance (WGAN)
 - Mínimos Quadrados (Least Squares – LSGAN)
 - ...
- GAN Zoo
 - <https://github.com/hindupuravinash/the-gan-zoo>

Minimax

- Função de Custo da Rede Discriminativa

- $C^{(D)} = -\frac{1}{2}\mathbb{E}_{x \sim p_{data}} \log D(x) - \frac{1}{2}\mathbb{E}_z \log (1 - D(G(z)))$

- Função de Custo da Rede Generativa

- $C^{(G)} = -C^{(D)}$

Treinamento de uma GAN

Treinamento de uma GAN

As funções de perda \mathcal{L}_G e \mathcal{L}_D de G e D possuem objetivos contrários. Por isso, o “jogo” entre as duas redes é classificado como um **cenário adversarial**.

Treinamento de uma GAN

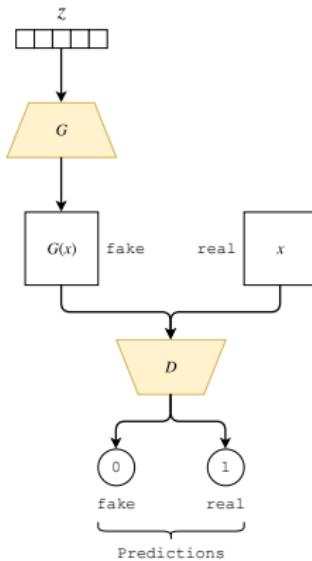


Figura: Arquitetura de uma GAN tradicional.

Treinamento de uma GAN

Treinamento de uma GAN

D quer acertar a classificação entre amostras reais e sintéticas.

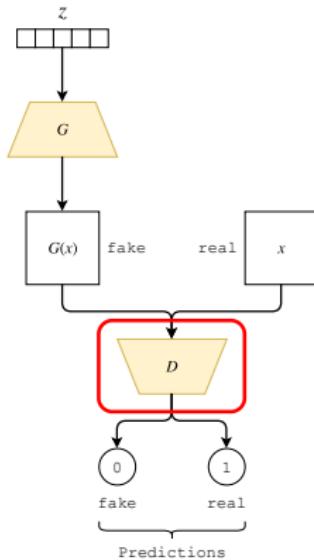


Figura: Arquitetura de uma GAN tradicional.

Treinamento de uma GAN

Treinamento de uma GAN

G quer fazer D errar a classificação entre amostras reais e sintéticas.

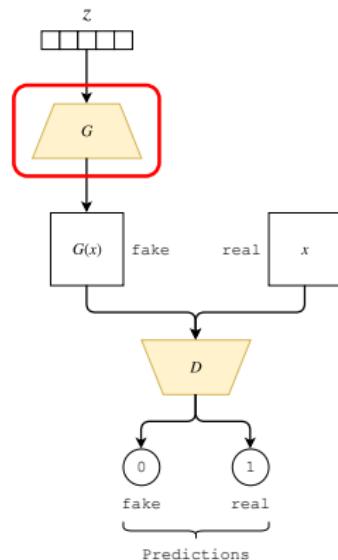


Figura: Arquitetura de uma GAN tradicional.

Treinamento de uma GAN

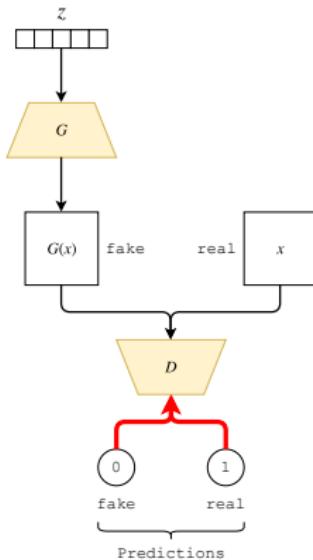


Figura: Backpropagation em uma GAN tradicional.

Treinamento de uma GAN

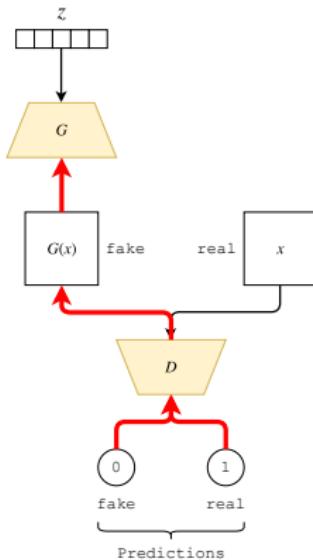


Figura: Backpropagation em uma GAN tradicional.

Treinamento de uma GAN

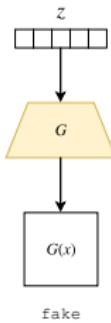


Figura: Geração de imagens em uma GAN tradicional.

Arquitetura de uma GAN

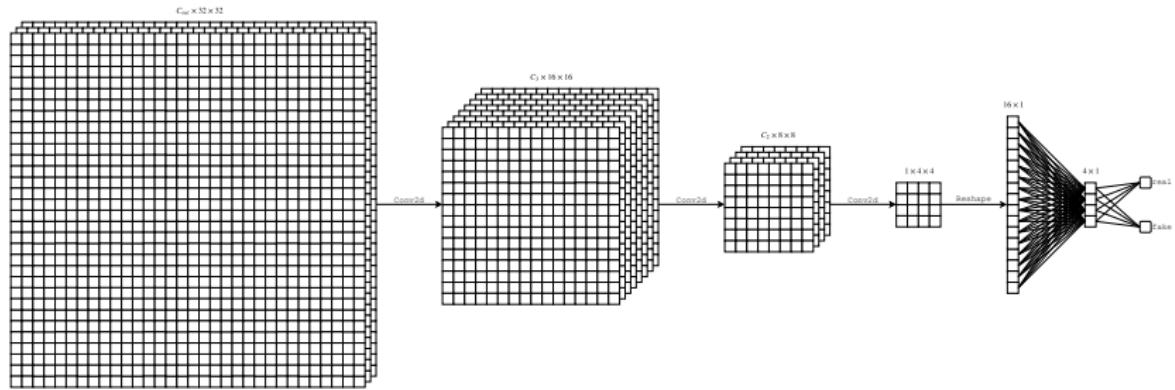


Figura: Arquitetura de D .

Arquitetura de uma GAN

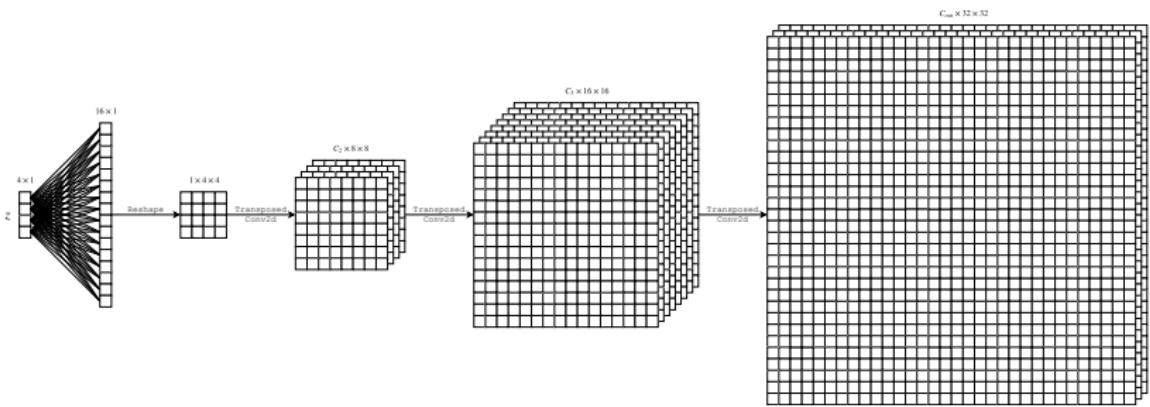


Figura: Arquitetura de G .

Dificuldades de Treinamento

Dificuldades de Treinamento

GANs **não possuem uma única função de loss** como é o caso de arquiteturas tradicionais como MLPs, CNNs, FCNs, R-CNNs, RNNs, Transformers, etc. Ao invés disso, o que é otimizado é o **objetivo composto** do conjunto das losses de G e de D .

Dificuldades de Treinamento

Dificuldades de Treinamento

Como os objetivos das losses é contrário, é possível treinar G e D conjuntamente de forma não supervisionada. Convergência é atingida quando se chega no **Equilíbrio de Nash**^{1,2}.

¹https://medium.com/@jonathan_hui/gan-why-it-is-so-hard-to-train-generative-advisory-networks-819a86b3750b

²<https://ahmedhanibrahim.wordpress.com/2017/01/17/generative-adversarial-networks-when-deep-learning-meets-game-theory/>

GANs

Demo - GANs

convolutional_gan.ipynb

Conditional GANs

Demo (Extra) - Conditional GANs

conditional_gan.ipynb

Generative Adversarial Networks

Limitações

- Treinamento instável
- Funcionam melhor em dados mais comportados (i.e. imagens pouco complexas de rostos, animais, comidas)
- Pouco generalizável para amostras out-of-distribution