Resumo de artigo

"Modelos Generativos e Redes Avançadas"

Felype N. de Souza

Daniel F. de Barros

Emilly Gabrielly

Henrico Bela

Artigo escolhido

"Training Generative Adversarial Networks with Limited Data"

Construído por

Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, Timo Aila

Link para o Artigo

https://arxiv.org/pdf/2006.06676.pdf

https://arxiv.org/pdf/2006.06676.pdf

São Paulo

11/03/2023

## Introdução a problemática - StyleGANs

Uma das principais preocupações no aprendizado de máquina é a geração de dados realistas. A capacidade de gerar novas amostras de dados é uma habilidade muito valiosa, especialmente quando os dados reais são escassos, difíceis de obter, até mesmo corrompidos. Como um modelo generativo, as redes generativas adversárias (GANs) se tornaram populares devido a sua capacidade de gerar dados mais próximos aos dados reais em diversas aplicações, como visão computacional, processamento de linguagem natural e síntese de voz. Entretanto, a GANs ainda possui grandes dificuldades, como por exemplo a instabilidade do treinamento, problemas com a convergência de dados e com o gerenciamento de gradientes soltos.

Para solucionar esses problemas, foram propostas várias possíveis melhorias para as GANs pela comunidade de pesquisa, como a arquitetura de rede mais avançada, melhores técnicas de regularização e novas funções de perda aplicadas. Entre essas propostas, destacam-se a arquitetura de rede “StyleGAN”, pois utiliza técnicas de normalização estilística para gerar imagens de alta qualidade e também, a “DualGAN”, por usar duas redes generativas e discriminativas para preservar a correspondência de domínios em aplicações responsáveis pela tradução de imagem.

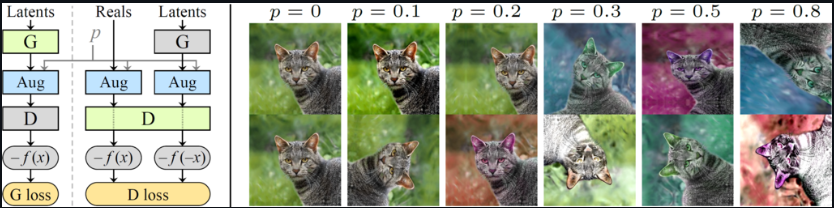
No GAN, a rede é composta por dois modelos: o gerador e o discriminador. O gerador cria as imagens sintéticas enquanto o discriminador avalia se essas imagens são reais ou falsas. O objetivo é treinar o gerador para criar imagens tão realistas que o discriminador não possa mais distingui-las das imagens reais.

Já no StyleGAN, a rede possui um controle mais refinado sobre a aparência das imagens geradas. O modelo permite que o usuário ajuste diversos parâmetros para controlar aspectos específicos da imagem, como cores, texturas, poses e expressões faciais. Dessa forma, o StyleGAN é capaz de gerar imagens com maior nível de detalhe e maior variação do que o GAN tradicional.

Resumindo, enquanto o GAN é um modelo de rede neural que gera imagens a partir de um conjunto de dados, o StyleGAN é uma evolução do GAN, que permite maior controle sobre a aparência das imagens geradas.

## Esboço ou Arquitetura Exemplo do Modelo Abordado

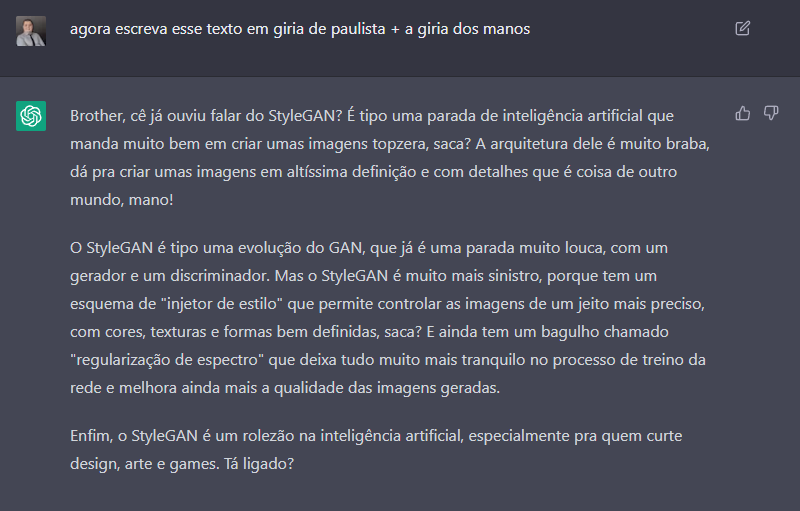
A StyleGAN é uma arquitetura de rede generativa proposta em 2018 que produz imagens de alta qualidade e realistas, com um grande nível de controle sobre as características geradas. Em comparação com outras arquiteturas GAN, a StyleGAN introduz duas técnicas importantes: a normalização estilística e a técnica de mapeamento de vetor latente (latent vector mapping).

*Arquitetura da rede - Fonte: https://github.com/NVlabs/stylegan2-ada*

A normalização estilística permite a normalização de cada canal de ativação da rede em relação ao seu desvio padrão e média para garantir que as diferentes camadas da rede produzam uma saída consistente. Já a técnica de mapeamento de vetor latente transforma um vetor de entrada em um vetor latente no espaço de características, que é então usado para gerar a imagem de saída.

A StyleGAN usa ainda uma técnica de geração progressiva, em que as imagens são geradas em etapas, assim como podemos ver na imagem acima (p = 0… p = 0.8) começando com uma resolução mais baixa e aumentando gradualmente até a resolução desejada. Essa técnica permite a criação de imagens detalhadas com grande controle sobre os recursos gerados.

## O que o ChatGPT acha de Style GANs



ChatGPT com girias!

## Resultados ou Discussão de possíveis resultados que podem advir do uso da técnica

A StyleGAN\* É uma das arquiteturas de rede generativa mais avançadas atualmente, capaz de gerar imagens de alta qualidade e realismo. Sua capacidade de controle sobre as características geradas permite a criação de imagens personalizadas e adaptadas a diferentes aplicações, como a criação de personagens para jogos, o design de roupas e a criação de animais virtuais.

Ao utilizar as StyleGANS, pode-se obter duas vertentes, das quais a primeira convém em utilizar dados limitados, assim como citado no artigo Training Generative Adversarial Networks with Limited Data. No qual enfatiza muito bem as vantagens de se utilizar esse tipo de método, dentre eles: evitar o Overfitting. Porém como toda solução, existem seus problemas, e um deles é o alto custo de uso e processamento com Placas de vídeo, CUDA (NVIDIA), etc.

Outras aplicações possíveis da StyleGAN incluem a geração de imagens médicas para pesquisa médica, a criação de paisagens para simulações e jogos e a geração de dados de treinamento para aprendizado de máquina em tarefas específicas. A capacidade da StyleGAN de produzir imagens realistas.

Os autores deste artigo de pesquisa descrevem a descoberta da eficácia do aumento adaptativo do discriminador (ADA) na melhoria da qualidade dos resultados quando os dados de treinamento são escassos. Eles enfatizam que o aumento deve ser usado para preencher lacunas nos dados reais, e não como substituto para eles. Eles recomendam pesquisas adicionais sobre o conjunto mais efetivo de aumentos e outras técnicas que possam ajudar com dados limitados.

O ADA tem um efeito negligenciável no consumo de energia, tornando-o uma opção economicamente viável para treinar modelos. Os autores observam que a ideia central dos aumentos do discriminador foi descoberta independentemente por três outros grupos de pesquisa, e recomendam a leitura desses artigos. Uma comparação direta dos resultados entre os trabalhos paralelos é difícil devido a diferenças em conjuntos de dados e metodologia, mas parece provável que a combinação de ideias de todos os quatro artigos possa melhorar ainda mais os resultados.

***StyleGAN - É uma técnica de aprendizado de máquina que foi introduzida em 2018 pela equipe de pesquisa do laboratório de pesquisa em inteligência artificial da NVIDIA. É uma abordagem para a geração de imagens fotorrealistas e altamente detalhadas usando redes generativas adversárias (GANs) que são treinadas em grandes conjuntos de dados de imagens.***