

## 6. Geração de Faces com base em redes adversárias

O artigo "Progressive Growing of GANs for Improved Quality, Stability, and Variation" apresenta uma metodologia de treinamento para Redes Adversárias Generativas (GANs), visando aprimorar a qualidade, estabilidade e diversidade das imagens geradas.

### Introdução às GANs

As Redes Adversárias Generativas (GANs), introduzidas por Ian Goodfellow em 2014, são uma arquitetura de aprendizado profundo usada para gerar dados sintéticos realistas. A estrutura básica das GANs consiste em dois componentes principais:

- **Gerador (Generator):** Gera amostras falsas (dados sintéticos) a partir de um vetor de entrada aleatório (geralmente ruído). O objetivo do gerador é enganar o discriminador criando amostras que se assemelhem às amostras reais do conjunto de treinamento.
- **Discriminador (Discriminator):** Recebe tanto amostras reais quanto amostras geradas pelo gerador. O objetivo do discriminador é distinguir entre amostras reais e geradas, classificando-as corretamente.

O treinamento das GANs é formulado como um jogo min-max onde o gerador tenta minimizar a função de perda do discriminador, enquanto o discriminador tenta maximizá-la. Isso cria uma competição que, idealmente, converge para um equilíbrio onde o gerador produz dados indistinguíveis dos dados reais.

### Desafios iniciais nas GANs

Primeiramente referente a instabilidade no treinamento, pois as redes muitas vezes sofrem de oscilação ou colapso de modo (quando o gerador foca em produzir apenas algumas variações dos dados reais). Também no que se refere a qualidade dos dados gerados as GANs básicas frequentemente lutam para produzir imagens de alta resolução e detalhes consistentes.

### Crescimento Progressivo

O gerador e o discriminador começam a treinar em baixa resolução e, gradualmente, novas camadas são adicionadas para aumentar a resolução e detalhamento das imagens. O crescimento progressivo permite a aceleração do treinamento e maior estabilidade, permitindo a geração de imagens de alta qualidade, como imagens CelebA em resolução 1024×1024.

### Aumento da Variedade das Imagens

Proposta de um método simples para incrementar a diversidade das imagens geradas, alcançando um recorde de pontuação de inception de 8,80 no conjunto de dados CIFAR-10 não supervisionado.

Figura 1 - Resultados

UNSUPERVISED			LABEL CONDITIONED		
Method		Inception score	Method		Inception score
ALI	(Dumoulin et al., 2016)	$5.34 \pm 0.05$	DCGAN	(Radford et al., 2015)	6.58
GMAN	(Durugkar et al., 2016)	$6.00 \pm 0.19$	Improved GAN	(Salimans et al., 2016)	$8.09 \pm 0.07$
Improved GAN	(Salimans et al., 2016)	$6.86 \pm 0.06$	AC-GAN	(Odena et al., 2017)	$8.25 \pm 0.07$
CEGAN-Ent-VI	(Dai et al., 2017)	$7.07 \pm 0.07$	SGAN	(Huang et al., 2016)	$8.59 \pm 0.12$
LR-AGN	(Yang et al., 2017)	$7.17 \pm 0.17$	WGAN-GP	(Gulrajani et al., 2017)	$8.67 \pm 0.14$
DFM	(Warde-Farley & Bengio, 2017)	$7.72 \pm 0.13$	Splitting GAN	(Grinblat et al., 2017)	$8.87 \pm 0.09$
WGAN-GP	(Gulrajani et al., 2017)	$7.86 \pm 0.07$			
Splitting GAN	(Grinblat et al., 2017)	$7.90 \pm 0.09$			
Our (best run)		<b><math>8.80 \pm 0.05</math></b>			
Our (computed from 10 runs)		<b><math>8.56 \pm 0.06</math></b>			

Fonte: Karras, T.; Aila, T.; Laine, S.; Lehtinen, J. *Progressive Growing of GANs for Improved Quality, Stability, and Variation* (2017).

### Detalhes de Implementação

- **Estratégia de Crescimento Progressivo:**
  - O treinamento começa com redes simples que geram imagens de baixa resolução (como 4×4 pixels).
  - À medida que o treinamento avança, novas camadas são adicionadas gradualmente ao gerador e ao discriminador para lidar com resoluções mais altas, como 8×8, 16×16, até 1024×1024.
  - Durante a transição para uma resolução maior, as redes interpolam suavemente entre camadas antigas e novas, evitando saltos abruptos que possam desestabilizar o treinamento.
- **Métodos de Suavização:**
  - O processo de adição de camadas utiliza interpolação linear, permitindo que as camadas recém-adicionadas sejam introduzidas com suavidade, em vez de serem ativadas diretamente. Isso reduz oscilações durante o treinamento.
- **Normalização e Equalização de Peso:**
  - A normalização de pixel e a equalização de pesos foram implementadas para estabilizar o treinamento.
    - **Normalização de Pixel:** Força os pixels gerados a terem estatísticas específicas (como média e variância) para melhorar a estabilidade.
    - **Equalização de Peso:** Automatiza o escalonamento dos pesos, eliminando a necessidade de inicializações complexas.

- **Aprimoramento do Discriminador:**
  - O discriminador é projetado para focar em detalhes de alta frequência (como texturas) em imagens de alta resolução, enquanto mantém sua capacidade de distinguir estruturas globais em resoluções menores.
- **Dados de Treinamento:**
  - As imagens de alta qualidade são usadas desde o início, mas o gerador e o discriminador começam operando em versões reescaladas dessas imagens para baixa resolução. Isso garante que informações críticas sejam capturadas em níveis iniciais antes que detalhes finos sejam aprendidos.
- **Evitar Colapso de Modo:**
  - Para mitigar o problema do colapso de modo (onde o gerador ignora a diversidade do conjunto de dados), o artigo implementa técnicas como minibatch discrimination e o uso de um discriminador que avalia conjuntos de amostras ao invés de amostras isoladas.

Esses avanços na implementação tornam o modelo mais estável, eficiente e capaz de produzir imagens que alcançam níveis sem precedentes de qualidade e realismo.

## **Conclusão**

Sugestão de uma nova métrica para avaliar os resultados das GANs, considerando tanto a qualidade quanto a variedade das imagens geradas. Desenvolvimento de uma versão de maior qualidade do conjunto de dados CelebA, utilizada para treinar e avaliar o modelo proposto.

## **REFERENCIAS**

KARRAS, T.; AILA, T.; LAINE, S.; LEHTINEN, J. *Progressive Growing of GANs for Improved Quality, Stability, and Variation*. arXiv preprint arXiv:1710.10196, 2017. Disponível em: <https://arxiv.org/pdf/1710.10196v3>. Acesso em: 17 dez. 2024.