

Uma investigação comparativa entre dados sintéticos vs aumento de dados tradicional em problemas de classificação de imagens

João Vítor Correia Pessoa

Universidade Federal do Rio de Janeiro (UFRJ)

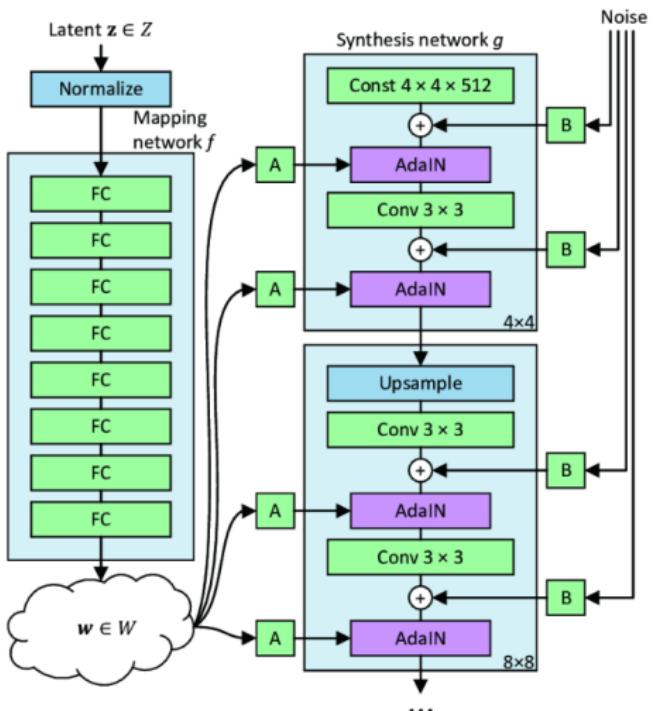
December 10, 2025

Motivação

- Muitos domínios possuem poucas amostras rotuladas
- Técnicas convencionais de aumentos de dados podem ser insuficientes
- GANs modernos, como StyleGAN2-ADA, permitem gerar imagens sintéticas realistas
- Objetivo: avaliar impacto do uso de dados sintéticos vs. aumento tradicional no treinamento de classificadores

- Modelo generativo baseado em StyleGAN2 com foco em datasets pequenos
- ADA (Adaptive Discriminator Augmentation): aplica aumento adaptativo para evitar overfitting
- Permite treinar com poucas centenas de imagens mantendo qualidade
- Usado aqui para gerar imagens sintéticas que expandem o conjunto de dados real

Arquitetura geral da StyleGAN2-ADA



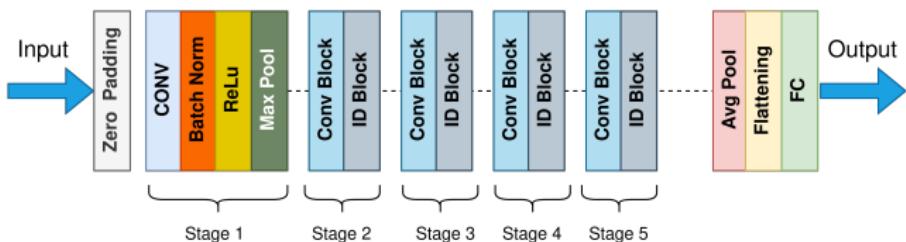
Arquitetura geral do modelo StyleGAN2-ADA



Tipos de aumento de imagens realizados pelo algoritmo ADA

Introduz **blocos residuais** com conexões de atalho (*skip connections*) que permitem:

- Treinamento de redes muito profundas
- Redução da degradação de performance à medida que profundidade aumenta

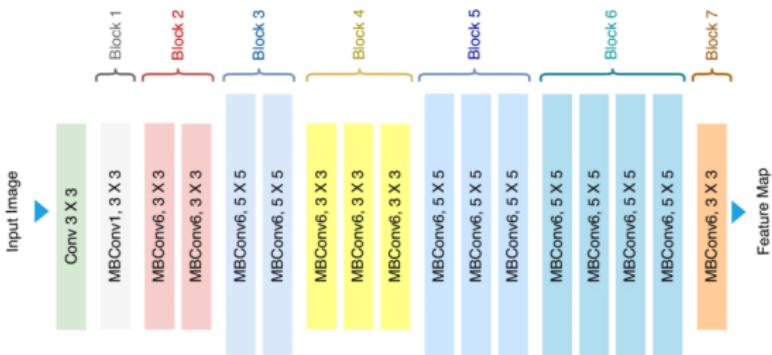


Arquitetura geral do modelo ResNet50

EfficientNet-B0

Baseada em blocos **MBConv** (Mobile Inverted Bottleneck), eficientes e leves, possui com características principais:

- Alto desempenho com baixo custo computacional
- Ótima escolha para datasets pequenos e treinamento rápido
- Favorece generalização devido ao design eficiente



Arquitetura geral do modelo EfficientNet-B0

Parâmetros importantes

Table 1: Configurações Gerais do Experimento

Parâmetro	Valor
Dataset	AFHQ
Número de classes	3
Cenários	real_only, generated_only, real_plus_gan, real_aug_standard
Device	CUDA

Table 2: Modelos

Modelo	Configuração
StyleGAN2-ADA	Pesos AFHQ
ResNet50	Pesos ImageNet (IMAGENET1K_V2), camada FC substituída
EfficientNet-B0	Pesos ImageNet (IMAGENET1K_V1), classifier[1] substituído

Parâmetros importantes

Table 3: Hiperparâmetros de Treinamento da StyleGAN2-ADA

Parâmetro	Valor
Resolução da Imagem	512×512
Épocas	10
Minibatch size	64
Minibatch stdev	8
Dataset x-flips	True
Feature maps	$1 \times$
Learning rate $\eta \times 10^{-3}$	2.5
R_1 regularization λ	0.5
G moving average	20k
Mixed-precision	True
Mapping net depth	8
Style mixing reg.	True
Path length reg.	True

Table 4: Hiperparâmetros de Treinamento dos classificadores

Parâmetro	Valor
Épocas	5
Batch size	32
Otimizador	AdamW
Learning rate	1×10^{-4}
Loss function	CrossEntropyLoss
Num. workers (DataLoader)	4
Normalização	mean = (0.5,0.5,0.5), std = (0.5,0.5,0.5)

Parâmetros importantes

Table 5: Transformações aplicadas

Resize	256 × 256
RandomHorizontalFlip	50%
RandomRotation	15°
ColorJitter	brightness=0.2, contrast=0.2, saturation=0.1, hue=0.05
Normalize	mean=0.5, std=0.5

Table 6: Cenários Experimentais

Cenário	Descrição
real_only	Apenas imagens reais
generated_only	Apenas imagens sintéticas (StyleGAN2-ADA)
real_plus_gan	Reais + sintéticas
real_aug_standard	Apenas reais + data augmentation tradicional

Geração das imagens



Imagens reais



Imagens geradas

Resumo dos experimentos

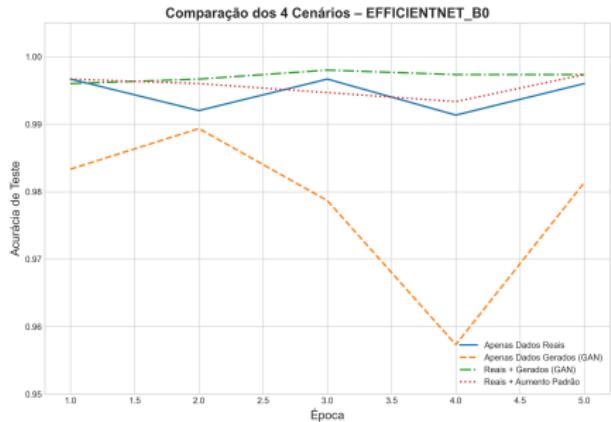
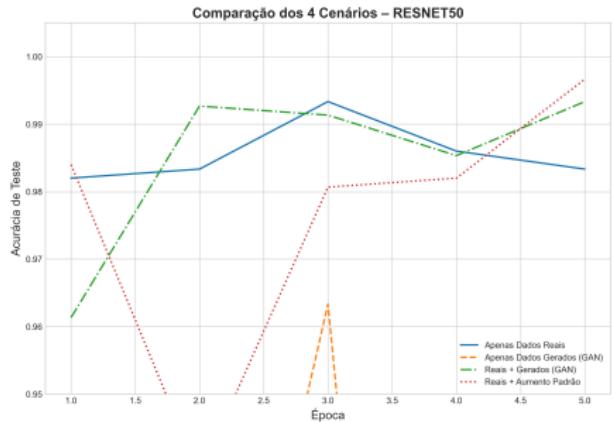
Table 7: Métricas do treinamento da StyleGAN2-DAN

Categoria	FID 50k	Tempo total(min/s)
cat	4.292	19m 46s
dog	8.230	19m 45s
wild	3.251	19m 47s

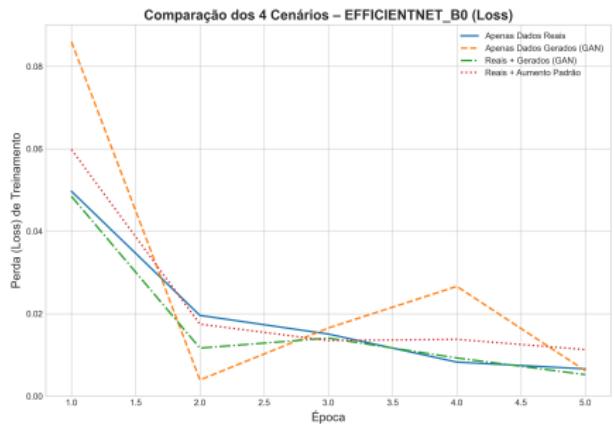
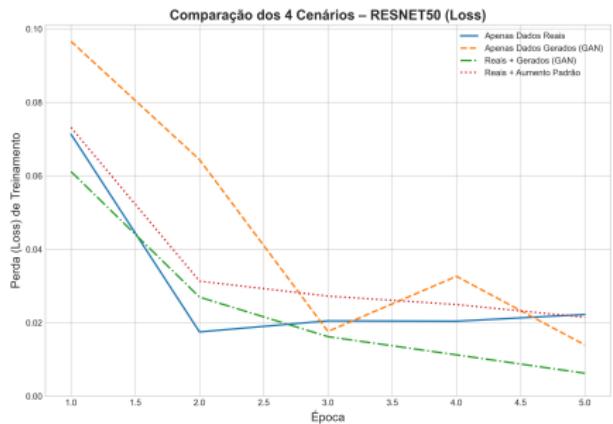
Table 8: Resultados dos Experimentos

Experimento	Train Loss	Train Acc	Test Acc	F1	Precision	Recall
resnet50_generated_only	0.014	0.995	0.943	0.942	0.949	0.943
resnet50_real_aug_standard	0.021	0.993	0.997	0.997	0.997	0.997
resnet50_real_only	0.021	0.993	0.997	0.997	0.997	0.997
resnet50_real_plus_gan	0.022	0.992	0.983	0.983	0.983	0.983
efficientnet_b0_generated_only	0.006	0.999	0.981	0.981	0.982	0.981
efficientnet_b0_real_aug_standard	0.011	0.996	0.997	0.997	0.997	0.997
efficientnet_b0_real_only	0.007	0.998	0.996	0.996	0.996	0.996
efficientnet_b0_real_plus_gan	0.005	0.998	0.997	0.997	0.997	0.997

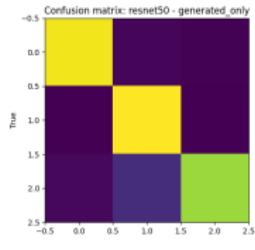
Curvas de Acurácia



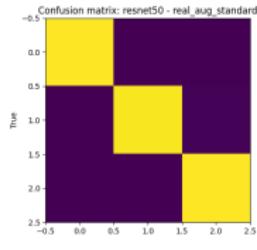
Curvas de Perda



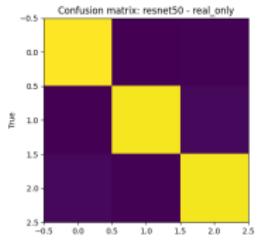
Matrizes de confusão



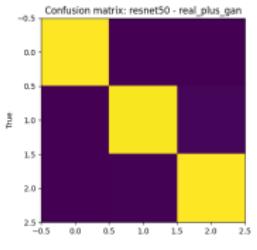
Confusion matrix: efficientnet_b0 - generated_only



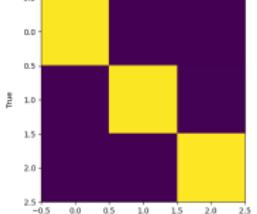
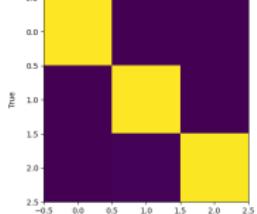
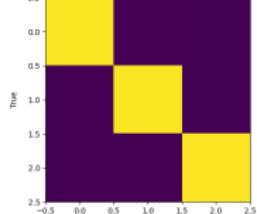
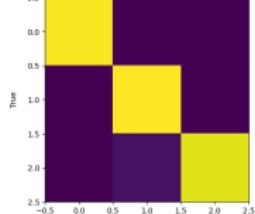
Confusion matrix: efficientnet_b0 - real_aug_standard



Confusion matrix: efficientnet_b0 - real_only



Confusion matrix: efficientnet_b0 - real_plus_gan



- Os dados sintéticos contribuíram para reduzir overfitting nos modelos profundos
- O treinamento apenas com dados sintéticos apresentou melhor desempenho na rede EfficientNet-B0
- Para o conjunto de dados analisado, o aumento de dados tradicional continua relevante

Considerações finais

- Modelos GAN avançados como StyleGAN2-ADA se mostram eficazes em alguns cenários de poucos dados
- Dados sintéticos podem complementar a augmentation tradicional
- As melhorias dependem da qualidade das imagens geradas
- Trabalhos futuros:
 - Avaliar datasets em que augmentation tradicional não gera impacto significativo
 - Avaliar a relação entre quantidade máxima de dados sintéticos no treinamento com a qualidade do gerador