

Relatório – CNN e GANs no MNIST

1. Introdução

Redes Generativas Adversariais (GANs) têm se mostrado extremamente eficazes na geração de imagens sintéticas realistas. Em particular, modelos condicionais, como a *Conditional GAN* (*cGAN*), permitem controlar a classe das amostras geradas, enquanto a *InfoGAN* introduz códigos latentes interpretáveis, favorecendo diversidade e controle semântico.

Uma aplicação prática importante dessas arquiteturas é o data augmentation generativo, no qual dados sintéticos são utilizados para aumentar o conjunto de treinamento de classificadores supervisionados. Neste trabalho, investigamos o impacto do uso de *cGANs* e *InfoGANs*, em versões MLP e convolucionais, no desempenho de uma CNN classificadora treinada sobre o conjunto de dados MNIST.

2. Conjunto de Dados

Foi utilizado o conjunto de dados MNIST, composto por:

- 70.000 imagens de dígitos manuscritos (0–9)
- Resolução: 28×28 pixels
- Tons de cinza
- Divisão padrão:
 - 60.000 imagens para treinamento
 - 10.000 imagens para teste

O MNIST é adequado para este estudo por possuir baixa complexidade visual, permitindo comparações claras entre arquiteturas MLP e convolucionais.

3. Metodologia

3.1 CNN Classificadora Base

Inicialmente, foi treinada uma CNN convolucional como baseline, utilizando apenas os dados reais do MNIST. Essa rede serviu como referência para avaliar o impacto da inclusão de dados sintéticos.

3.2 Modelos Generativos

Foram implementados e treinados quatro modelos generativos:

1. cGAN MLP

- Gerador e discriminador baseados em camadas totalmente conectadas
- Entrada: ruído + rótulo one-hot
- Saída: imagem flatten (784 dimensões)

2. cGAN Convolucional

- Gerador baseado em convoluções transpostas
- Discriminador totalmente convolucional
- Geração de imagens diretamente no formato 28×28

3. InfoGAN MLP

- Arquitetura MLP
- Uso de código latente categórico e contínuo
- Treinada com perda de informação mútua (InfoLoss)

4. InfoGAN Convolucional

- Arquitetura convolucional
- Maior capacidade de modelar estrutura espacial
- Código latente interpretável

3.3 Avaliação dos Modelos Generativos

A qualidade dos modelos geradores foi avaliada utilizando:

- **FID (Fréchet Inception Distance)**
- **KID (Kernel Inception Distance)**
- **Precision / Recall para modelos generativos**

As features foram extraídas por uma CNN treinada no MNIST, garantindo coerência entre domínio real e sintético.

3.4 Data Augmentation Generativo

Para cada gerador:

- Foram criadas 2.000 imagens sintéticas por classe
- As imagens foram combinadas com o conjunto de treino real

- Uma nova CNN foi treinada em cada cenário
- A avaliação foi feita sempre no conjunto de teste real

4. Resultados

4.1 Qualidade das Imagens Geradas

Modelo	FID ↓	KID ↓	Precision	Recall
cGAN MLP	0.387	0.0035	1.000	0.000
cGAN Conv	0.013	0.00007	0.400	0.443
InfoGAN MLP	0.368	0.0032	0.716	0.000
InfoGAN Conv	0.033	0.00012	0.255	0.357

Análise:

- Modelos convolucionais superaram amplamente os MLPs
- cGAN Conv apresentou o melhor equilíbrio entre fidelidade e diversidade
- Modelos MLP sofreram **colapso de modo**, evidenciado por recall próximo de zero

4.2 Impacto no Classificador

Tabela 2 – Acurácia da CNN

Conjunto de Treinamento	Acurácia
MNIST (baseline)	0.9854
MNIST + cGAN MLP	0.9829
MNIST + cGAN Conv	0.9860
MNIST + InfoGAN MLP	0.9873
MNIST + InfoGAN Conv	0.9566

Análise:

- cGAN Conv resultou em leve melhora no desempenho da CNN
- InfoGAN MLP apresentou ganho inesperado, possivelmente devido a efeito de regularização

- InfoGAN Conv prejudicou o classificador, sugerindo excesso de diversidade ou ruído
-

5. Discussão

Os resultados indicam que a qualidade do gerador é crucial para que dados sintéticos sejam úteis no treinamento de classificadores. Modelos convolucionais, por respeitarem a estrutura espacial das imagens, produzem amostras mais realistas e consistentes.

Embora o InfoGAN introduza diversidade por meio de códigos latentes interpretáveis, essa diversidade nem sempre se traduz em ganho de desempenho, especialmente quando o gerador produz amostras ambíguas.

O caso do InfoGAN MLP sugere que dados sintéticos de baixa qualidade podem, em alguns cenários, atuar como ruído estruturado, promovendo regularização da CNN.

6. Conclusão

Neste trabalho, foi demonstrado que:

- Arquiteturas convolucionais são superiores às MLPs para geração de imagens MNIST
- O cGAN convolucional foi o modelo mais eficaz para data augmentation
- A inclusão de códigos latentes (InfoGAN) aumenta diversidade, mas não garante melhoria no classificador
- O impacto positivo do data augmentation depende diretamente da qualidade do gerador

Os resultados reforçam que GANs podem melhorar classificadores tradicionais, desde que o modelo gerativo seja estável e produza imagens realistas.

7. Trabalhos Futuros

- Avaliação em bases mais complexas (Fashion-MNIST, CIFAR-10)
- Ajuste da proporção entre dados reais e sintéticos
- Análise de matrizes de confusão
- Estudo do impacto de códigos contínuos do InfoGAN