

Relatório Técnico – Classificador de Frutas (Maçã x Banana)

Este documento descreve as decisões técnicas tomadas no desenvolvimento do modelo de classificação de imagens capaz de distinguir entre maçãs e bananas utilizando uma Rede Neural Convolutiva (CNN).

1. Por que o dataset é pequeno?

O dataset utilizado é propositalmente pequeno, contendo poucas dezenas de imagens por classe. O objetivo principal não é alcançar desempenho industrial, mas compreender profundamente o funcionamento das CNNs, desde a criação manual do dataset até o ajuste fino da arquitetura e dos hiperparâmetros. Além disso, um dataset reduzido facilita a observação de overfitting, underfitting e o impacto de técnicas como data augmentation e dropout.

2. Data Augmentation

Para compensar o baixo número de imagens, foi aplicado data augmentation durante o treinamento. As transformações incluem rotações, espelhamento horizontal e ajustes de brilho, contraste e saturação. Isso permite que o modelo veja variações das mesmas imagens, aumentando sua capacidade de generalização.

3. Arquitetura da Rede (3 Convoluções)

A rede utiliza três camadas convolucionais com aumento progressivo do número de filtros (4 → 8 → 16). Essa escolha permite que o modelo aprenda características em diferentes níveis: bordas simples na primeira camada, padrões intermediários na segunda e estruturas mais complexas na terceira. Mais camadas seriam desnecessárias para imagens pequenas (32x32) e poderiam causar overfitting.

4. Dropout

Foi aplicado Dropout de 12% na camada totalmente conectada. Mesmo com apenas 64 neurônios, o dropout ajuda a evitar que a rede memorize o conjunto de treino, forçando-a a aprender representações mais robustas.

5. Acurácia do Modelo

O modelo alcançou aproximadamente 80% de acurácia no conjunto de validação. Considerando o tamanho reduzido do dataset e o fato de as imagens de teste serem totalmente inéditas, esse resultado é consistente e demonstra que a rede aprendeu padrões relevantes, não apenas memorizou os dados de treino.

6. Objetivo do Projeto

Este projeto tem caráter educacional e experimental. Ele demonstra todo o pipeline de um modelo de visão computacional: criação do dataset, definição da arquitetura, treinamento, validação, salvamento do modelo e inferência em imagens nunca vistas anteriormente.