

RESIDÊNCIA
EM SOFTWARE
BAHIA + TECNOLOGIA + EMPREENDEDORISMO

Relatório Técnico: Implementação e Análise de Classificação com Redes Convolucionais e o dataset CUFS

Cauã Tavares Nunes
Vinícius Levi dos Santos

JUAZEIRO - BA
24/11/2024

1. RESUMO

Este projeto visa solucionar o problema de classificação de imagens utilizando Redes Neurais Convolucionais (CNN). O processo envolveu preparação de dados, construção e treinamento de um modelo CNN, seguido de avaliação detalhada dos resultados. Os principais resultados indicam [Inserir Principais Resultados].

2. INTRODUÇÃO

A classificação de imagens é um dos problemas mais desafiadores em visão computacional. Redes Neurais Convolucionais (CNN) têm se destacado como uma abordagem eficiente devido à sua capacidade de extrair automaticamente características relevantes das imagens. Este relatório explora a aplicação de CNNs para resolver um problema específico de classificação de imagens, destacando as etapas e os resultados obtidos.

3. METODOLOGIA

3.1 Preparação dos dados

Coleta e Rotulagem

As imagens foram coletadas de uma pasta chamada "photos". Os arquivos foram rotulados como masculino (0) ou feminino (1) com base no início do nome do arquivo.

Pré-processamento de imagens

As imagens foram redimensionadas para 200 x 250 pixels e normalizadas para o intervalo [0, 1].

Divisão dos dados

O conjunto de dados foi dividido em 50% de treinamento, 30% de validação e 20% de teste. A divisão usou estratificação para manter a proporção de rótulos consistente em cada subset.

3.2 Estrutura do modelo

Arquitetura

Utilizou-se uma rede neural convolucional (CNN) com três camadas convolucionais seguidas por pooling (Max Pooling). Após as camadas convolucionais, há uma etapa de flatten e duas camadas densas, sendo a última a camada de saída com ativação sigmóide para classificação binária.

Camadas convolucionais

A primeira camada tem 32 filtros, a segunda tem 64 e a terceira tem 128, todas com kernel de tamanho (3, 3).

Dropout

Implementado após a primeira camada densa para reduzir o overfitting, com uma taxa de 0,5.

Processo de avaliação

A sequência de passos utilizada proporcionou um entendimento claro do desempenho do modelo e onde ele pode ser melhorado. O código feito estabelece as bases para essa avaliação, e a inclusão de métricas como:

F1-Score: 0.8182

AUC-ROC: 0.9293

Matriz de Confusão

Ajudaram a interpretar melhor os resultados e impactar positivamente em futuros ajustes e aprimoramentos no modelo.

Após treinar e ajustar o modelo, utilizamos o conjunto de teste para uma avaliação final. O que nos deu uma estimativa do desempenho real do modelo.

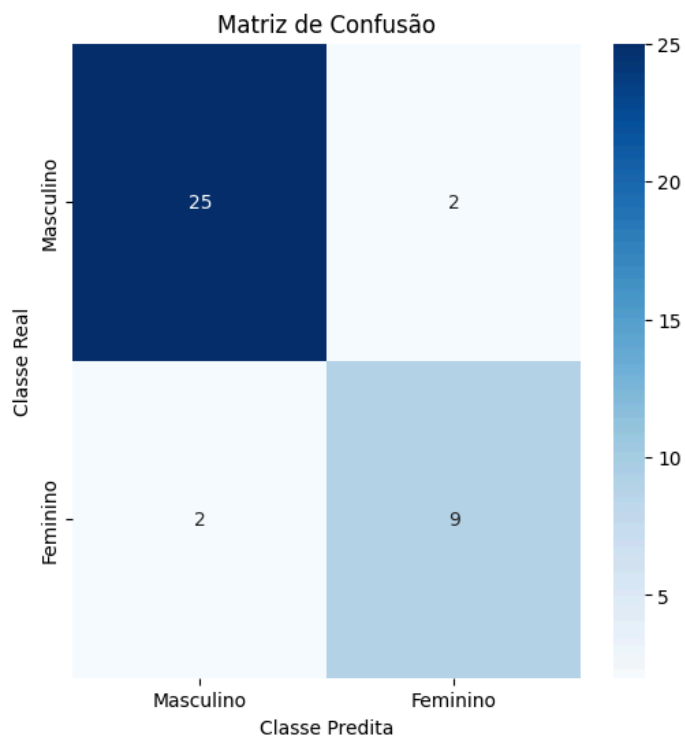


Figura 1 - Matriz de confusão

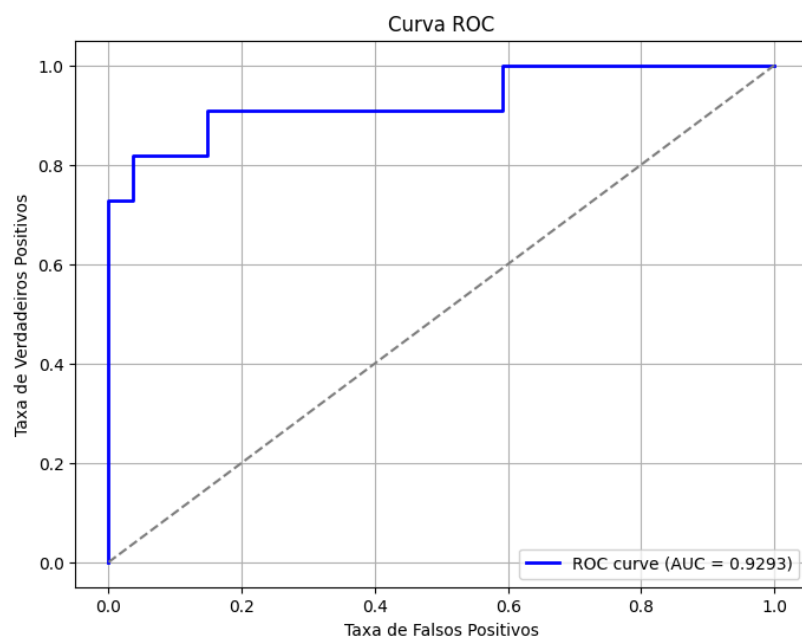


Figura 2 - Curva ROC

4. DISCUSSÃO

O F1-score mostrou bom equilíbrio entre as classes?

O F1-score é a média harmônica da precisão e do recall. Um valor de 0.8182 indica que o modelo teve um bom equilíbrio entre as classes, embora haja espaço para melhorias, este score sugere que o modelo está efetivamente distinguindo entre as classes masculino e feminino, mas ainda pode haver alguns falsos positivos ou falsos negativos.

Dificuldades do modelo

Algumas imagens problemáticas observadas foram aquelas que apresentam características que comumente são associadas a gênero, como o tamanho do corte de cabelo. Que foi a principal causa de confusão no modelo, pois, na maioria dos nossos dados, mulheres possuem cabelos longos e homens têm cabelos curtos e alguns samples que continham imagens com riscos pretos.

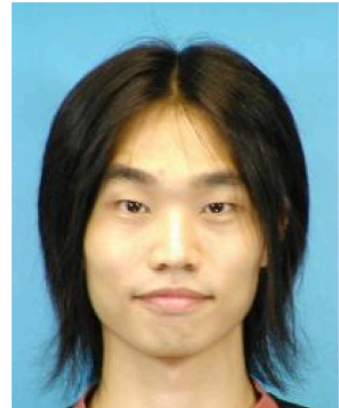
Real: Masculino, Pred: Feminino



Real: Feminino, Pred: Masculino



Real: Masculino, Pred: Feminino



Real: Feminino, Pred: Masculino



Figura 3 - Exemplos de imagens classificadas incorretamente

Alguma limitação do dataset pode ter influenciado os resultados (e.g., desbalanceamento entre classes, qualidade das imagens)?

Uma das principais limitações do dataset é a representatividade étnica. O modelo foi treinado com dados de pessoas de uma única etnia, o que limita sua aplicabilidade a populações diversificadas. Para criar um modelo mais geral, seria necessário expandir o dataset para incluir pessoas de diferentes países, cores de cabelo, tipos de cabelo, entre outros aspectos. Outra característica que faltou foi a variedade de ângulos de tomada de foto, o que pode afetar a precisão do modelo.

Há indícios de que o modelo tenha tido dificuldade com certos tipos de imagens (e.g., iluminação, ângulos, ruído)? Quais características são comuns nas imagens classificadas incorretamente?

Imagens com uma iluminação mais elevada, imagens que possuíam algum tipo de ruído e tamanho do cabelo.

Quais modificações na arquitetura do modelo ou nos hiperparâmetros poderiam melhorar o desempenho?

Expansão do número de filtros e adição de camadas convolucionais são passos indispensáveis para capturar detalhes complexos. Técnicas como Batch Normalization estabilizam o aprendizado, enquanto funções de ativação alternativas, como Leaky ReLU, aumentam a eficiência. Ajustar a taxa de aprendizado e aplicar técnicas de aumento de dados, como rotação e inversão, aprimoram a robustez e generalização do modelo. Essas estratégias, combinadas, podem maximizar a precisão e a capacidade de generalização do modelo.

Descobertas

A precisão do modelo ao caracterizar corretamente as fotos foi bastante impressionante, com erros apenas em alguns casos onde até mesmo um humano poderia ter dificuldade em determinar o gênero da pessoa apresentada. Isso indica que o modelo conseguiu se adaptar bem às características visuais e, na maioria das situações, forneceu resultados corretos e confiáveis.

5. CONCLUSÃO E TRABALHOS FUTUROS

Este projeto reforçou a importância de uma análise de dados minuciosa e da escolha de técnicas de modelagem apropriadas. Trabalhos futuros poderiam incluir a integração de dados temporais para prever tendências futuras e o uso de técnicas de aprendizado profundo para potencialmente melhorar o desempenho do modelo.

6. REFERÊNCIAS

Scikit-learn: Scikit-learn developers. Scikit-learn: Machine Learning in Python. Disponível em: <https://scikit-learn.org/>. Acesso em: 20/11/2024

Pandas: The Pandas development team. Pandas: Powerful data analysis tools for Python. Disponível em: <https://pandas.pydata.org/>. Acesso em: 20/11/2024.

Matplotlib: Hunter, J. D. Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90-95, 2007.

Seaborn: Waskom, M. L. seaborn: Statistical data visualization. Journal of Open Source Software, 1(0.010), 218, 2017.

AMARAL, Fernando. Formação Cientista de Dados: O Curso Completo. Udemy. Disponível em: <https://www.udemy.com/course/cientista-de-dados/>.