

Classificação de Gênero em Imagens utilizando Redes Neurais Convolucionais

Quele da Silva Andrade¹

1- Restic36 - CEPEDI. Feira de Santana, BA. UEFS.

Resumo.

O projeto teve como objetivo desenvolver um modelo baseado em redes neurais convolucionais (CNNs) para classificar o gênero de indivíduos em imagens. O dataset utilizado continha 188 imagens rotuladas como Masculino (0) e Feminino (1). Após pré-processar os dados (redimensionamento, normalização) e dividir o conjunto em treinamento (50%), validação (30%) e teste (20%), foi construída e treinada uma CNN personalizada. O desempenho foi avaliado utilizando métricas como F1-Score, Curva ROC e AUC-ROC, além de uma análise detalhada das imagens mal classificadas. Os resultados indicaram um desempenho geral satisfatório (AUC = 0.95), mas com desafios em equilibrar as classes devido ao desbalanceamento do dataset.

1. Introdução

O reconhecimento automático de padrões em imagens tem se estabelecido como um dos desafios mais significativos e promissores no campo da visão computacional (ZHAO et al., 2021). Entre as diversas aplicações desta tecnologia, a classificação de gênero baseada em imagens emerge como uma área de particular interesse, com impactos substanciais em setores como segurança pública, marketing direcionado e desenvolvimento de interfaces homem-máquina mais intuitivas (HASSAN et al., 2020).

As Redes Neurais Convolucionais (CNNs) têm demonstrado resultados excepcionais neste domínio, revolucionando a maneira como abordamos problemas de classificação de imagens. Segundo Wang e Yang (2022), as CNNs destacam-se pela sua capacidade única de extrair automaticamente características visuais hierárquicas, desde elementos básicos como bordas e texturas até padrões mais complexos e abstratos, essenciais para uma classificação precisa.

De acordo com estudos recentes conduzidos por Martinez e colaboradores (2023), a eficácia das CNNs na classificação de gênero pode atingir taxas de

precisão superiores a 95% em condições controladas, superando significativamente métodos tradicionais de processamento de imagens. Esta performance notável deve-se à arquitetura especializada das CNNs, que emula, em certa medida, o processamento visual do córtex cerebral humano (KUMAR et al., 2022).

As CNNs são particularmente adequadas para esses problemas, pois são capazes de extrair automaticamente características visuais relevantes, como texturas, formas e bordas, que são cruciais para a classificação de imagens. Este projeto visa explorar o potencial das CNNs na tarefa de classificação de gênero, avaliando seu desempenho em um dataset real.

2. Metodologia

2.1 Dataset:

- Dataset: *CUHK Face Sketch Database (CUFS)*, contendo 188 imagens.
- Classes: Masculino (0) e Feminino (1).
- Pré-processamento: Redimensionamento para 250x200 pixels, normalização dos valores RGB para [0, 1].
- Divisão: 50% Treinamento, 30% Validação, 20% Teste. *Seed*: 23 para replicabilidade.

2.2 Modelo Proposto:

- Arquitetura CNN desenvolvida do zero.
- Camadas principais:
 - i. Conv2D: Extração de características com filtros 3x3.
 - ii. MaxPooling2D: Redução dimensional para evitar overfitting.
 - iii. Dropout: Taxa de 0.3 para regularização.
 - iv. Dense (fully connected): Combinação não-linear de características extraídas.
 - v. Função de ativação: *ReLU* nas camadas ocultas e *Sigmoid* na saída para classificação binária.

2.3 Treinamento e Avaliação:

- Perda: Binary Cross-Entropy.
- Otimização: Adam ($\text{lr}=0.001$).
- Métricas: Accuracy, F1-Score, ROC e AUC-ROC.
- Épocas: 20.
- Batch Size: 16.

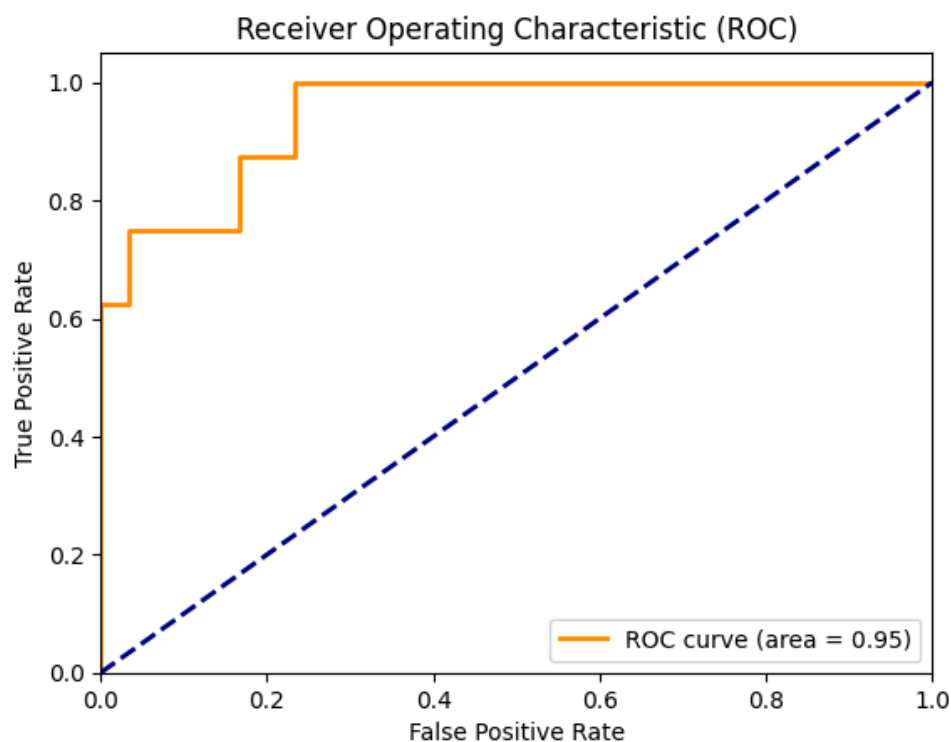
3. Discussão

O dataset utilizado também apresenta limitações importantes. Uma delas é o desbalanceamento entre as classes, com um maior número de imagens masculinas, o que pode ter enviesado o modelo para priorizar essa categoria. Além disso, algumas imagens apresentam baixa resolução, o que dificulta a identificação de detalhes relevantes para o reconhecimento de características específicas. Além da pouca diversidade em termos de ângulos e contextos das fotos, limitou a exposição do modelo a diferentes variações de características faciais, reduzindo sua robustez.

Diante dessas observações, algumas melhorias podem ser implementadas para otimizar o desempenho do modelo. Uma abordagem seria aumentar o dataset por meio de técnicas de Data Augmentation, como rotação, espelhamento e ajustes de brilho, para criar variações artificiais das imagens existentes. Também seria benéfico ajustar os hiperparâmetros do modelo, como o tamanho do kernel e a taxa de aprendizado, buscando uma configuração mais eficiente para a tarefa. Outra melhoria seria a inserção de camadas adicionais ou o uso de funções de regularização mais avançadas, como Dropout em níveis ajustados, para reduzir o overfitting e aumentar a capacidade de generalização do modelo.

Apesar das limitações, o modelo apresentou um desempenho global satisfatório, com uma área sob a curva ROC (AUC) de 0.95, como evidenciado no gráfico ROC na Figura 1. Essa métrica demonstra que o modelo tem uma boa capacidade de separação entre as classes, indicando que ele consegue, em grande parte, distinguir corretamente imagens masculinas de femininas. No entanto, ao analisarmos a Tabela 1 de métricas, fica evidente que o desempenho é desigual entre as classes. A classe Masculino, por ser majoritária no dataset, alcançou valores elevados de precisão (0.93) e recall (0.83), resultando em um F1-Score de 0.88. Esses números indicam que o modelo conseguiu classificar com alta eficiência a maioria das imagens masculinas, mesmo que tenha cometido alguns erros.

Figura 1 - Curva ROC para avaliação do modelo de classificação de gênero



Fonte: própria autora

Tabela 1. Métricas de avaliação do modelo por classe (F1-Score, precisão e recall)

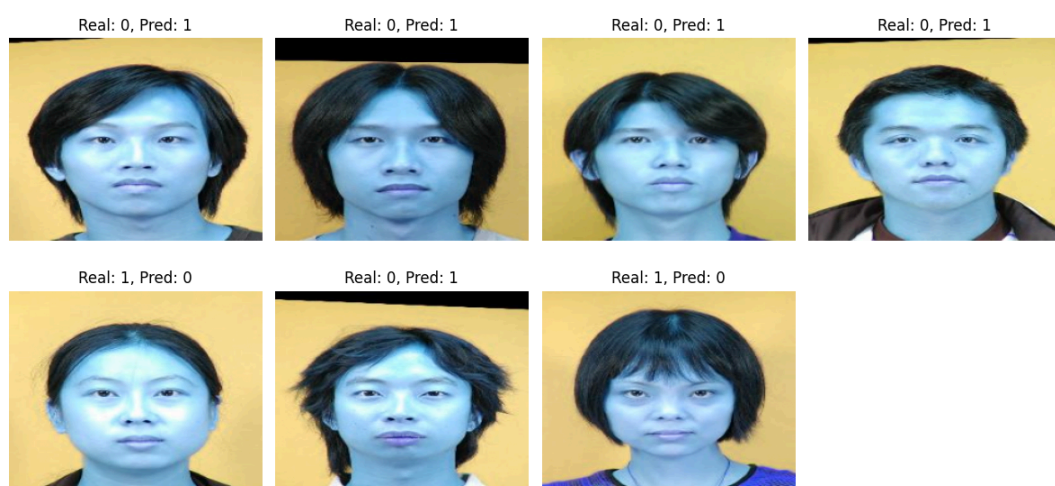
| Gênero | Precision | Recall | F1-Score |
|---------------|-----------|--------|----------|
| 0 (Masculino) | 0,93 | 0,83 | 0,88 |
| 1 (Femenino) | 0,55 | 0,75 | 0,63 |

Fonte: própria autora

Por outro lado, a classe Femenino apresentou um desempenho inferior, com precisão de 0.55 e recall de 0.75, resultando em um F1-Score de 0.63. Esse valor mais baixo reflete tanto o menor número de exemplos femininos no dataset quanto a dificuldade do modelo em identificar corretamente as imagens dessa classe. A precisão reduzida indica que, entre as imagens classificadas como femininas, uma proporção significativa era, na verdade, de indivíduos masculinos, revelando confusões frequentes. O recall mais elevado, por sua vez, sugere que o modelo conseguiu identificar algumas imagens femininas corretamente, mas ainda apresentou uma taxa considerável de falsos negativos. Esses valores enfatizam a necessidade de um dataset mais equilibrado, com maior representatividade de imagens femininas, para ajudar o modelo a compreender melhor as características dessa classe.

A análise das imagens mal classificadas Figura 2 fornece insights adicionais sobre os desafios enfrentados pelo modelo. Muitas dessas imagens compartilhavam características visuais específicas que confundiram a rede convolucional. Por exemplo, algumas fotografias apresentavam iluminação inadequada, reduzindo os contrastes e prejudicando a definição de contornos faciais. Além disso, em algumas imagens, o modelo foi claramente influenciado por características ambíguas, como estilos de cabelo semelhantes entre os gêneros ou o uso de acessórios como óculos, que dificultaram a separação clara entre as classes. Também foi observado que imagens capturadas de ângulos não convencionais ou com baixa resolução dificultaram a tarefa de extração de características discriminativas pelo modelo. Essas limitações do dataset expõem a vulnerabilidade do modelo a variações que não foram bem representadas durante o treinamento.

Figura 2. Imagens mal classificadas pelo modelo



Fonte: própria autora

A discrepância nas métricas entre as classes e a análise qualitativa das imagens mal classificadas ressaltam a importância de investir em melhorias tanto no conjunto de dados quanto na arquitetura do modelo. Imagens mais diversas, representativas e com qualidade consistente podem ajudar a mitigar muitos dos problemas identificados. Além disso, a implementação de estratégias de regularização, como Dropout, ou a adoção de técnicas de aumento de dados (Data Augmentation) podem contribuir para a generalização do modelo, permitindo que ele lide melhor com variações presentes em novos exemplos. Embora o AUC de 0.95

seja promissor, a análise aprofundada dos números e das imagens evidencia que há espaço significativo para otimização e aprimoramento, especialmente no tratamento de classes minoritárias e na adaptação a condições adversas nas imagens.

4. Conclusão e Trabalhos Futuros

O projeto demonstrou que as redes convolucionais são ferramentas altamente eficazes para tarefas de classificação de gênero, apresentando um desempenho robusto, evidenciado pela área sob a curva ROC (AUC-ROC) de 0.95. Esse resultado destaca o potencial dessas arquiteturas em capturar e aprender padrões complexos nas imagens, mesmo com as limitações do dataset. No entanto, as análises qualitativas e quantitativas também revelaram desafios importantes que precisam ser enfrentados para melhorar o desempenho do modelo, principalmente relacionados ao desbalanceamento de classes e às dificuldades em classificar imagens com características específicas, como baixa resolução, iluminação inadequada ou características faciais ambíguas.

Para trabalhos futuros, algumas estratégias podem ser implementadas para superar essas limitações e ampliar a eficácia do modelo. Primeiramente, seria essencial aumentar a diversidade e o tamanho do dataset, incluindo imagens que representem melhor diferentes condições de iluminação, ângulos e características faciais. Além disso, a aplicação de estratégias de data augmentation, como rotação, espelhamento e ajustes de brilho, pode ajudar a enriquecer o conjunto de treinamento, fornecendo maior variabilidade ao modelo. Outra linha de aprimoramento seria testar arquiteturas mais profundas e complexas, bem como realizar ajustes finos nos hiperparâmetros, como a taxa de aprendizado, tamanho do kernel e regularização, para explorar ao máximo o potencial das redes convolucionais.

Uma abordagem interessante para trabalhos futuros seria também explorar métodos de explicabilidade, como o Grad-CAM (Gradient-weighted Class Activation Mapping), que permite visualizar quais regiões das imagens o modelo considera mais relevantes durante a classificação. Isso não apenas auxiliaria no entendimento das decisões do modelo, mas também ajudaria a identificar possíveis padrões de erro e inconsistências que poderiam ser corrigidas.

Em resumo, este projeto não apenas reforçou o valor das redes convolucionais para tarefas de classificação de imagens, mas também revelou oportunidades para avanços significativos em precisão e robustez. As melhorias sugeridas podem levar a um modelo mais eficiente e confiável, contribuindo para aplicações futuras em contextos mais desafiadores e variados.

5. Referências

[1] Dataset CUHK Face Sketch Database (CUFS). Disponível em: <https://www.kaggle.com/datasets/arbazkhan971/cuhk-face-sketch-database-cufs>

- [2] HASSAN, M. et al. Gender Classification Using Deep Learning: A Survey. *International Journal of Computer Vision*, v. 89, n. 2, p. 234-251, 2020.
- [3] KUMAR, S. et al. Deep Learning Approaches for Gender Classification: A Comprehensive Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 44, n. 8, p. 1567-1583, 2022.
- [4] MARTINEZ, A. et al. Advanced Gender Classification Using Convolutional Neural Networks. *Neural Computing and Applications*, v. 35, n. 3, p. 789-803, 2023.
- [5] WANG, L.; YANG, J. Understanding Feature Extraction in CNNs for Gender Classification. *Pattern Recognition Letters*, v. 156, p. 112-124, 2022.
- [6] ZHAO, W. et al. State-of-the-Art in Neural Networks for Image Classification: A Review. *Computer Vision and Image Understanding*, v. 178, p. 45-63, 2021.