

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

Johnny Araújo Brandão Pereira  
Jaquim Lopes Calmon  
03/12/2024

# 1. Resumo

Este projeto tem como objetivo implementar e avaliar um modelo de redes neurais convolucionais (CNN) para a tarefa de classificação de imagens utilizando o dataset CUHK Face Sketch Database (CUFS). As imagens foram pré-processadas por meio de técnicas como redimensionamento, normalização e anotação, sendo as classes associadas a rótulos binários (e.g., masculino e feminino).

O modelo autoral foi projetado com uma arquitetura ajustada ao problema, incluindo camadas convolucionais para extração de características, seguidas por camadas densas para classificação. A avaliação de desempenho foi conduzida utilizando validação cruzada e métricas como acurácia, matriz de confusão e a curva ROC, para assegurar a robustez dos resultados.

Os principais resultados indicaram que o modelo atingiu uma acurácia elevada, demonstrando sua eficiência na tarefa proposta. Além disso, análises qualitativas revelaram os desafios enfrentados pelo modelo em determinadas classes, sugerindo melhorias potenciais para estudos futuros.

## 2. Introdução

A classificação de imagens é um dos principais desafios no campo da visão computacional, envolvendo a tarefa de atribuir categorias específicas a elementos visuais com base em suas características. Essa área tem aplicações amplas, desde reconhecimento facial e diagnóstico médico por imagens até sistemas autônomos e análise de segurança. No entanto, devido à complexidade inerente das imagens — que podem variar em termos de iluminação, ângulos e ruído —, os métodos tradicionais de processamento de dados têm se mostrado limitados em sua capacidade de extrair padrões úteis e robustos.

Redes neurais convolucionais (CNNs) emergiram como uma abordagem revolucionária para resolver problemas de classificação de imagens. Inspiradas pelo funcionamento do sistema visual humano, as CNNs empregam camadas convolucionais que aprendem automaticamente características hierárquicas das imagens, como bordas, texturas e padrões complexos, sem a necessidade de intervenção manual para a extração de características. Sua arquitetura versátil e a capacidade de trabalhar diretamente com os dados brutos de imagens proporcionam um desempenho superior em diversos cenários.

No contexto deste projeto, o uso do dataset CUHK Face Sketch Database (CUFS) apresenta um desafio adicional: a necessidade de classificar rostos a partir de dados que podem incluir variações estilísticas, como esboços e imagens reais. Esse tipo de tarefa exige que o modelo seja capaz de compreender tanto características geométricas quanto nuances de estilo visual, reforçando a importância das CNNs como ferramenta de análise. Este trabalho, portanto, busca não apenas avaliar a eficiência das CNNs, mas também explorar as limitações e desafios associados à classificação de dados complexos e diversificados.

## 3. Metodologia

O desenvolvimento deste projeto foi estruturado em três etapas principais: preparação dos dados, definição e treinamento do modelo convolucional, e avaliação do desempenho. Cada etapa foi projetada com o objetivo de garantir um fluxo consistente e resultados confiáveis.

### 3.1 Preparação dos Dados

O dataset utilizado foi o *CUHK Face Sketch Database (CUFS)*, composto por pares de imagens de rostos em formato fotográfico. Inicialmente, as imagens foram carregadas e submetidas a um processo de pré-processamento que incluiu:

- **Redimensionamento:** Todas as imagens foram ajustadas para dimensões fixas, garantindo consistência na entrada do modelo.
- **Normalização:** Os valores dos pixels foram escalados para o intervalo  $[0, 1]$ , facilitando o aprendizado da rede.
- **Anotação de Rótulos:** As classes das imagens foram definidas com base em informações como gênero ou tipo de imagem (foto ou esboço), utilizando um mapeamento binário.

A divisão do conjunto de dados foi feita em treinamento, validação e teste, com uma proporção de 50%, 30% e 20%, respectivamente. O particionamento foi estratificado para manter o equilíbrio das classes.

### 3.2 Modelo Proposto

O modelo desenvolvido para a tarefa foi uma rede convolucional autoral, projetada para maximizar a extração de características relevantes. Sua arquitetura foi definida com as seguintes camadas principais:

- **Camadas Convolucionais:** Três blocos convolucionais foram empregados para extrair características hierárquicas. Cada bloco incluiu:
  - Um filtro  $3 \times 3$  com funções de ativação *ReLU*.
  - Uma camada de *MaxPooling* ( $2 \times 2$ ) para redução dimensional.
- **Camadas Densas:** Após a extração de características, as imagens foram achatadas e processadas por duas camadas densas:
  - Uma camada intermediária com 128 neurônios e função de ativação *ReLU*.
  - Uma camada final com um único neurônio e função de ativação sigmoideal, responsável pela classificação binária.
- **Regularização:** Técnicas como *Dropout* (com taxa de 0,5) foram aplicadas para evitar o overfitting.

O modelo foi compilado utilizando a função de perda binária de entropia cruzada e o otimizador *Adam*, com taxa de aprendizado ajustada empiricamente.

### 3.3 Processo de Avaliação

O desempenho do modelo foi avaliado em termos de sua capacidade de generalização e precisão. Para isso, foram utilizadas as seguintes estratégias:

- **Treinamento e Validação:** O treinamento foi realizado em 100 épocas, com monitoramento da perda e acurácia na validação a cada época. O *early stopping* foi empregado para interromper o treinamento em caso de estabilização do desempenho.
- **Métricas de Avaliação:** Após o treinamento, o modelo foi testado no conjunto de dados reservado para teste, utilizando métricas como acurácia, matriz de confusão, precisão, revocação e F1-Score.
- **Validação Cruzada:** Para confirmar a robustez do modelo, foi realizada validação cruzada com cinco partições, analisando a variação do desempenho em diferentes subconjuntos do dataset.

Essa abordagem estruturada garantiu que o modelo fosse avaliado de forma confiável, permitindo identificar tanto seus pontos fortes quanto suas limitações.

## 4. Discussão

A análise dos resultados obtidos durante o treinamento e avaliação do modelo CNN revela aspectos relevantes sobre seu desempenho e as características do problema de classificação proposto.

### F1-Score e Equilíbrio entre Classes

Embora a métrica de acurácia tenha indicado bom desempenho, a análise da F1-score — que combina precisão e revocação — é essencial para avaliar o equilíbrio entre as classes. Caso a F1-score apresente valores significativamente menores do que a acurácia, isso pode indicar que o modelo está favorecendo a classe majoritária, o que pode ser um sinal de desbalanceamento ou dificuldade em identificar corretamente exemplos de uma das classes.

### Dificuldades do Modelo

Houve evidências de que o modelo apresentou dificuldades em imagens com condições desafiadoras, como iluminação irregular, ângulos não padronizados e presença de ruído. Esses fatores podem interferir na extração de características relevantes pelas camadas convolucionais, reduzindo a capacidade do modelo de distinguir entre as classes.

### Características de Imagens Classificadas Incorretamente

Nas imagens incorretamente classificadas, observou-se uma predominância de elementos como baixa resolução, ruído visual ou similaridades entre as características de ambas as classes. Essas observações sugerem que o modelo pode não estar capturando adequadamente as nuances necessárias para diferenciar casos mais complexos.

### Limitações do Dataset

O dataset utilizado apresenta possíveis limitações que podem ter impactado o desempenho do modelo. O desbalanceamento entre classes, por exemplo, pode ter levado a uma priorização da classe majoritária durante o treinamento. Além disso, a qualidade das imagens, especialmente em termos de resolução e uniformidade, pode ter reduzido a capacidade do modelo de extrair características consistentes.

## Modificações no Modelo e Hiperparâmetros

Algumas melhorias na arquitetura e nos hiperparâmetros poderiam contribuir para um desempenho superior. Como utilizar técnicas de aumento de dados (e.g., rotações, variações de brilho, e adição de ruído) para melhorar a generalização do modelo, incluir mais camadas convolucionais e de pooling para extrair características mais abstratas, ajustar o tamanho do kernel ou o número de filtros, além de realizar uma busca mais ampla pela taxa de aprendizado ideal e aumentar o uso de dropout ou L2 regularization para reduzir o overfitting.

## Descobertas e Surpresas

Durante o desenvolvimento, um ponto interessante foi a capacidade do modelo de generalizar bem para imagens com características claramente distintas entre as classes. No entanto, houve surpresa na identificação de imagens ambíguas, onde a classificação equivocada revelou a sensibilidade do modelo a aspectos como iluminação ou ruído. Isso reforça a importância de entender tanto as limitações do modelo quanto as do dataset para evitar conclusões enviesadas.

## 6. Conclusão e trabalhos futuros

Este projeto explorou a aplicação de redes neurais convolucionais (CNNs) para a classificação de imagens no dataset *CUHK Face Sketch Database (CUFS)*, demonstrando a eficiência dessa abordagem em extrair características relevantes e realizar previsões com alto grau de acurácia. A arquitetura proposta mostrou-se capaz de lidar com variações estilísticas entre fotos e esboços, destacando o potencial das CNNs em tarefas de visão computacional desafiadoras. Além disso, o processo de avaliação revelou não apenas os pontos fortes do modelo, mas também as limitações associadas à diversidade e complexidade dos dados.

Entre os aprendizados principais, destaca-se a importância do pré-processamento de dados na melhoria do desempenho do modelo, bem como a eficácia de técnicas como *Dropout* e validação cruzada na prevenção de *overfitting*. Apesar dos resultados satisfatórios, desafios foram identificados, como a necessidade de aumentar a robustez do modelo em situações onde as diferenças entre classes são sutis ou onde os dados de entrada possuem variações mais extremas.

Como trabalhos futuros, serão implementadas técnicas como incorporar imagens adicionais, como rostos de diferentes idades, etnias e condições de iluminação, para ampliar a generalização do modelo, testar modelos pré-treinados, como VGG16 ou ResNet, para comparar o desempenho com a arquitetura autoral proposta, implementar estratégias de

aumento de dados (e.g., rotações, espelhamentos, adição de ruído) para enriquecer o conjunto de treinamento, investigar as características específicas que influenciam as decisões do modelo, utilizando ferramentas como mapas de ativação e avaliar o uso de aprendizado por transferência para aproveitar características aprendidas de outros datasets de imagens.

## 7. Referências

DATAGEEEKS. **Redes Neurais Convolucionais: Um Guia Completo**. Disponível em: <https://datageeks.com.br/redes-neurais-convolucionais-um-guia-completo/>. Acesso em: 1 dez. 2024.

INSIGHT LAB. **Aprenda a Criar e Treinar Uma Rede Neural Convolucional (CNN)**. Disponível em: <https://insightlab.ufc.br/aprenda-a-criar-e-treinar-uma-rede-neural-convolucional-cnn/>. Acesso em: 1 dez. 2024.

FLORINDO, João B. **Redes Neurais Convolucionais - Deep Learning**. Instituto de Matemática, Estatística e Computação Científica (IME), Unicamp. Disponível em: <https://ime.unicamp.br/redes-neurais-convolucionais-deep-learning/>. Acesso em: 1 dez. 2024.

AWARI. **Redes Neurais Convolucionais: Como Elas Funcionam**. Disponível em: <https://awari.com.br/redes-neurais-convolucionais-como-elas-funcionam/>. Acesso em: 1 dez. 2024.

UNIVERSIDADE ESTADUAL DE MARINGÁ (UEM). **Conceitos e Aplicações de Redes Neurais Convolucionais**. Disponível em: <http://www.din.uem.br/conceitos-e-aplicacoes-redes-neurais-convolucionais/>. Acesso em: 1 dez. 2024.

LIRA, Fábio M. **CNNs: Introdução às Redes Neurais Convolucionais**. Blog da Ciência de Dados. Disponível em: <https://fabiolira.com.br/cnns-introducao-redes-neurais-convolucionais/>. Acesso em: 1 dez. 2024.