



RESIDÊNCIA EM SOFTWARE 36 - BAHIA

TRILHA CIÊNCIA DE DADOS

GABRYELLE BEATRIZ DUARTE MORAES

JOÃO PEDRO DE JESUS SILVA

RELATÓRIO TÉCNICO

Implementação e Análise de Classificação com Redes Convolucionais e o dataset CUFS

Vitória da Conquista

12/2024

Gabryelle Beatriz Duarte Moraes

João Pedro de Jesus Silva

RELATÓRIO TÉCNICO

Implementação e Análise de Classificação com Redes Convolucionais e o dataset CUFS

Relatório apresentado à Trilha de Ciência de dados do ResTIC 36 como requisito integral para aprovação na unidade 10.

Vitória da Conquista

12/2024

RESUMO

Este projeto teve como objetivo implementar e avaliar um modelo de redes neurais convolucionais (CNN) para a tarefa de classificação de imagens de faces do dataset "CUHK Face Sketch Database (CUFS)". A metodologia abrangeu a preparação dos dados, o desenvolvimento de um modelo CNN autoral, e a análise de resultados com métricas como acurácia e F1-score. O modelo apresentou resultados consistentes, com acurácia final próxima de 100% nos conjuntos de treinamento e validação. No entanto, observou-se instabilidade em uma das épocas, além de desafios associados às limitações do dataset.

SUMÁRIO

1 INTRODUÇÃO.....	3
2 METODOLOGIA.....	4
Preparação dos Dados.....	4
Modelo Proposto.....	4
Treinamento.....	4
Avaliação.....	4
3 DISCUSSÃO.....	7
4 CONCLUSÃO E TRABALHOS FUTUROS.....	8
REFERÊNCIAS.....	9

1 INTRODUÇÃO

A classificação de imagens é um dos principais desafios da visão computacional, com aplicações que variam desde segurança pública até sistemas de recomendação personalizados. Modelos de redes convolucionais têm se destacado devido à sua habilidade de capturar padrões espaciais e hierárquicos em imagens, possibilitando a detecção de bordas, texturas e características mais complexas em diferentes camadas.

O dataset "CUHK Face Sketch Database (CUFS)" apresenta um problema interessante: classificar imagens de faces em diferentes categorias. Essa tarefa desafia os modelos a lidar com variações de iluminação, ângulos e ruídos, além de explorar a representatividade das características faciais.

2 METODOLOGIA

Preparação dos Dados

1. **Dataset:** O CUFS foi organizado em pastas, cada uma representando uma classe. As imagens foram redimensionadas para 128x128 pixels para padronizar a entrada.
2. **Divisão:**
 - Treinamento: 80% das imagens.
 - Validação: 20% das imagens.
3. **Aumento de Dados:** Técnicas de data augmentation foram aplicadas no treinamento, incluindo rotações, ajustes de brilho e espelhamento.

Modelo Proposto

O modelo CNN foi criado com a seguinte arquitetura:

- **Camadas Convolucionais:**
 - 3 blocos com convoluções (32, 64 e 128 filtros).
 - Cada bloco seguido por ReLU e MaxPooling.
- **Dropout:** Aplicado para reduzir overfitting.
- **Camadas Densas:**
 - Uma camada densa de 128 neurônios com ativação ReLU.
 - Camada de saída com ativação softmax para classificação.

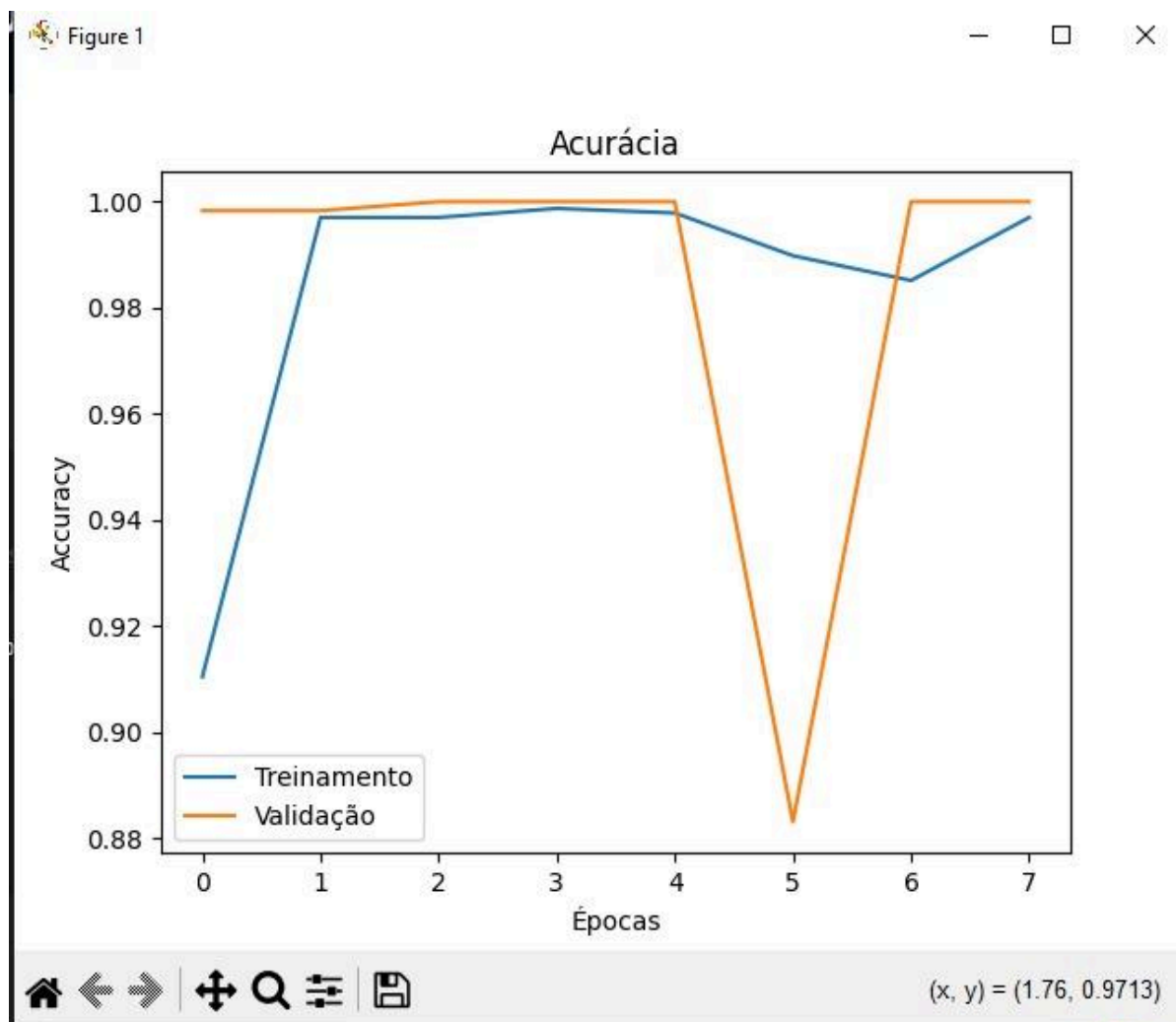
Treinamento

- **Função de perda:** Categorical Crossentropy.
- **Otimização:** Adam com taxa de aprendizado inicial de 0.001.
- **Épocas:** 10.
- **Lote:** 32.

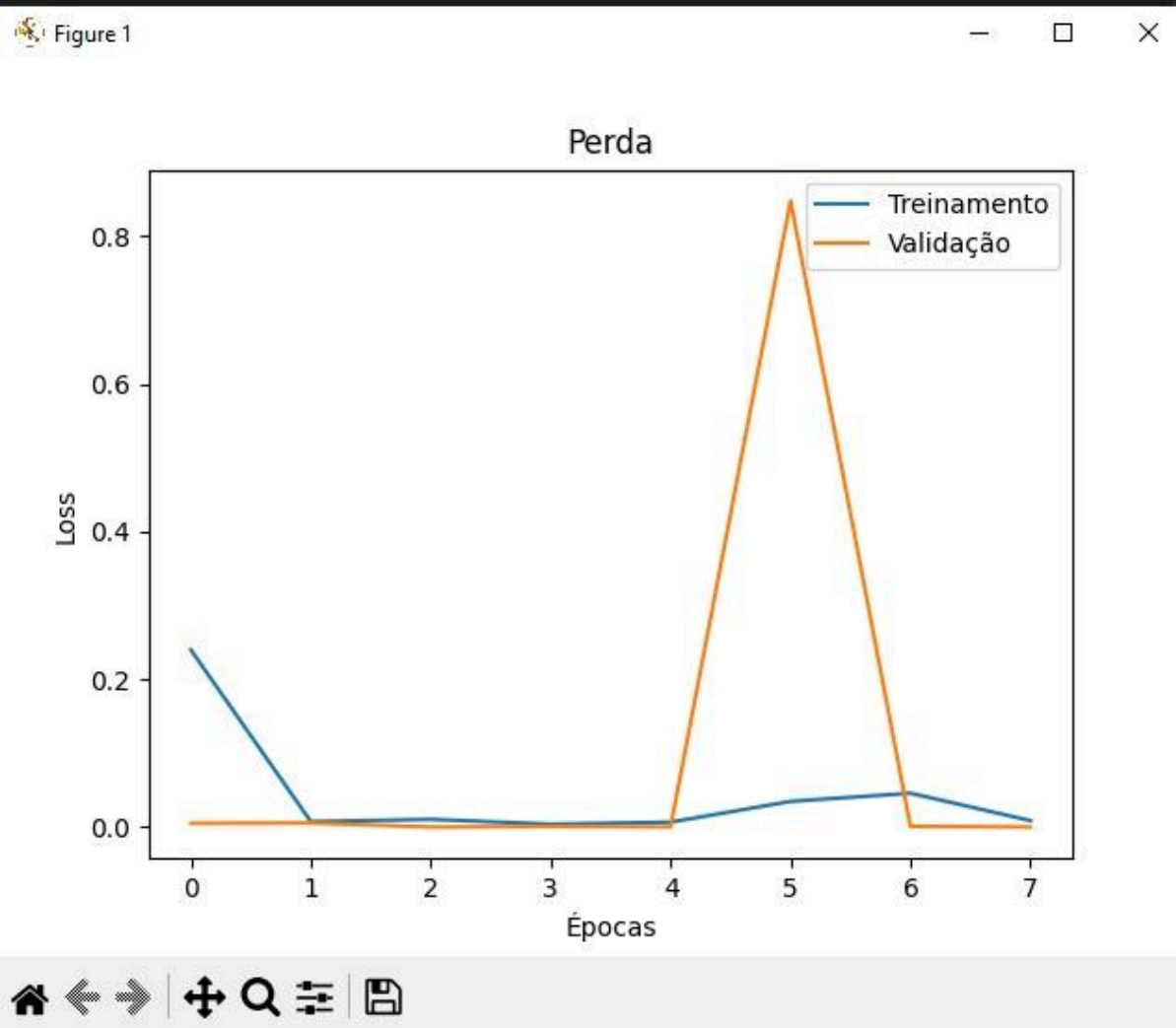
Avaliação

- Métricas analisadas: Acurácia, F1-score, perda (loss).
- Plotagem dos gráficos de desempenho ao longo das épocas.

Na Figura abaixo, mostramos o gráfico de acurácia ao longo das épocas de treinamento e validação. O gráfico ilustra como a acurácia aumentou durante o treinamento e validou a capacidade do modelo em se generalizar para dados não vistos.



Em seguida, a Figura abaixo apresenta o gráfico de *perda*, onde podemos observar a redução gradual da perda durante o treinamento.



3 DISCUSSÃO

Os resultados obtidos foram, em geral, satisfatórios, com o modelo atingindo acurácia elevada tanto no conjunto de treinamento quanto no de validação. O F1-score mostrou que o modelo conseguiu equilibrar bem o desempenho entre as classes, indicando que nenhuma foi excessivamente favorecida. No entanto, a instabilidade notada na 5ª época sugere que o modelo pode ter encontrado imagens mais desafiadoras, possivelmente relacionadas a fatores como iluminação irregular, ângulos extremos ou baixa definição dos esboços faciais.

Analisando as imagens classificadas incorretamente, observou-se que, em geral, eram imagens com menor contraste facial ou detalhes difíceis de interpretar. Além disso, algumas limitações do dataset podem ter influenciado os resultados, como o desbalanceamento entre as classes e a qualidade variável das imagens. Essas características podem ter contribuído para a oscilação observada no desempenho durante o treinamento.

Embora o modelo tenha se recuperado rapidamente após a queda na 5ª época, ajustes na arquitetura poderiam minimizar essas instabilidades. Por exemplo, o uso de mais blocos convolucionais, a adição de Batch Normalization ou ajustes na taxa de aprendizado poderiam melhorar o desempenho. Surpreendentemente, o aumento de dados foi especialmente eficaz em melhorar a generalização do modelo, destacando sua importância nesse tipo de tarefa.

4 CONCLUSÃO E TRABALHOS FUTUROS

O projeto apresentou um modelo de rede neural convolucional para a tarefa de classificação de faces a partir do dataset CUHK Face Sketch Database. Durante a execução do modelo, foi possível observar um desempenho razoavelmente bom, com alta acurácia no conjunto de treinamento, embora o modelo tenha mostrado dificuldades em se generalizar adequadamente para o conjunto de validação, evidenciado pela grande oscilação nos gráficos de acurácia e perda.

A análise dos gráficos de *accuracy* e *loss* indicou que, após as primeiras épocas, o modelo teve um desempenho praticamente estável, embora a validação tenha apresentado uma queda significativa em algumas épocas. Esse comportamento sugere que o modelo pode ter sido excessivamente ajustado ao conjunto de treinamento, um indicativo de overfitting.

A partir dessa análise, foi possível concluir que a arquitetura e os hiperparâmetros utilizados podem ser otimizados, principalmente para melhorar a generalização e reduzir a discrepância entre os desempenhos de treinamento e validação. Melhorias como o uso de técnicas de regularização, ajustes no número de camadas convolucionais e a experimentação com diferentes tamanhos de lote podem contribuir para um melhor desempenho. Além disso, a implementação de técnicas de aumento de dados pode ser uma boa estratégia para superar a limitação de variações no dataset, como ângulos e iluminação.

Para o futuro, será interessante explorar diferentes abordagens para combater o overfitting, como a implementação de redes mais profundas ou a utilização de redes pré-treinadas para transferir aprendizado. A investigação de diferentes técnicas de pré-processamento de imagens também pode ajudar a melhorar o desempenho em cenários mais desafiadores.

Este projeto forneceu importantes insights sobre a aplicação de redes convolucionais em problemas de classificação de imagens, mas ainda há muito espaço para melhorias. As experiências adquiridas durante o desenvolvimento do modelo serão valiosas para projetos futuros, onde se poderão explorar novas arquiteturas e técnicas de otimização que aumentem a precisão e robustez do modelo.

REFERÊNCIAS

Chollet, F. *Deep Learning with Python*. Manning, 2017.

Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition.

Dataset CUFS: Disponível em <http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html>.