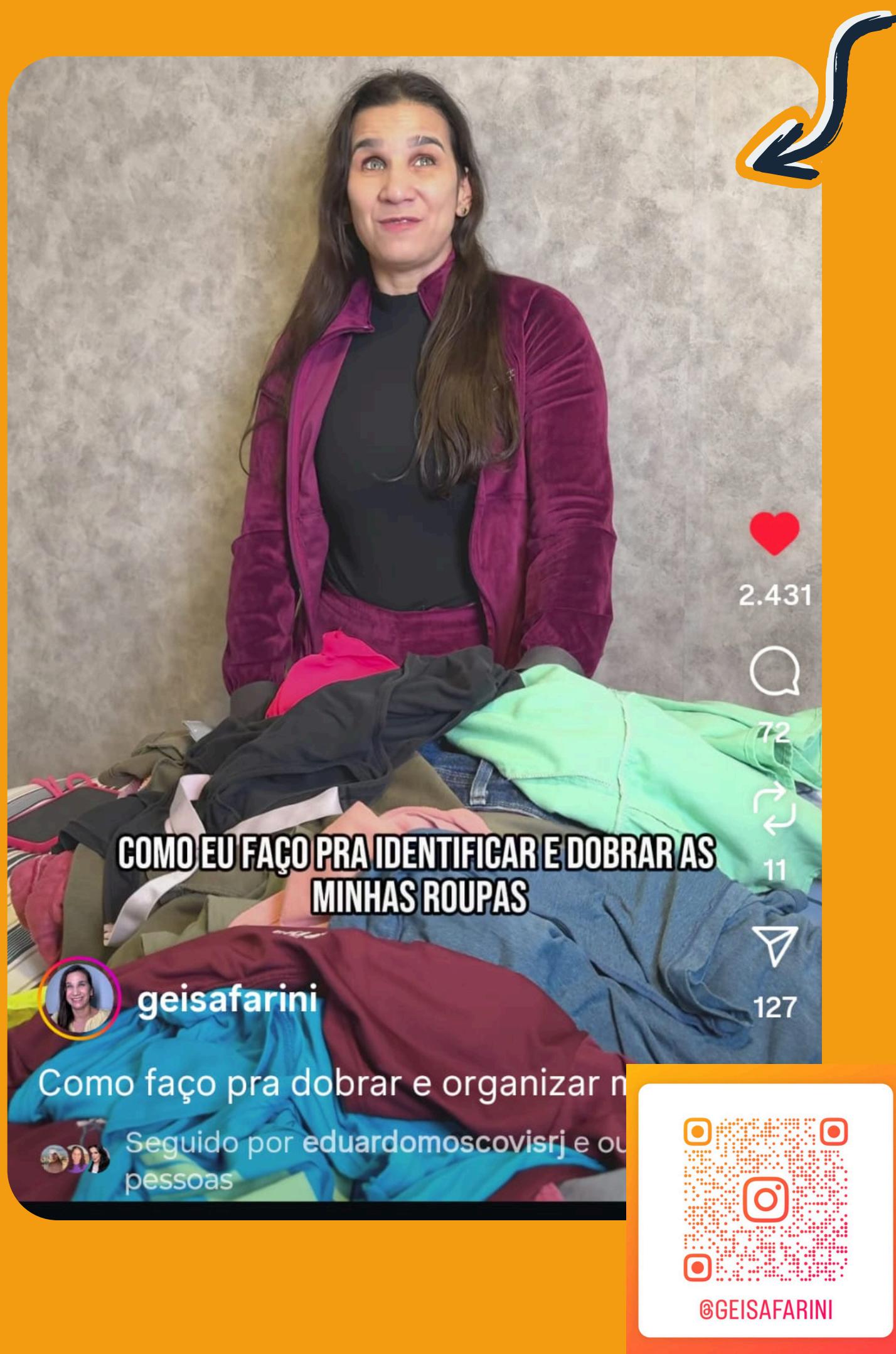




Visão Computacional Aplicada à Acessibilidade:

Classificação de Vestuário com Redes Neurais
Convolucionais para Pessoas com Deficiência Visual

Docentes: Filipe Rolim & Valmir Macário.
Discente: Jacqueline Navarro da Silva.



Motivação Pessoal e Relevância Social

- Curso Recursos Educacionais Acessíveis (IBC/MEC, 2023)
- Inspiração em Geisa Farini
(Engenheira de Software | seleção brasileira de futebol de cegas | Palestrante)
- Soluções acessíveis para pessoas com deficiência visual
- Tecnologia como ferramenta de equidade

Desafios para Pessoas com Deficiência Visual

- Dificuldade em identificar roupas
- Impacto na autonomia e inclusão social



Instituto Benjamin Constant (IBC)

Objetivos do Projeto

- Comparar o desempenho de quatro modelos de aprendizado profundo, incluindo o de referência.
- Identificar o modelo mais eficiente e aplicar otimizações.
- Avaliar os ganhos obtidos em relação ao baseline após a otimização.

Pessoas cegas: Classificação de categorias de roupas e detecção de manchas usando aprendizado por transferência.



Artigo base de referência (2023)



Article

Blind People: Clothing Category Classification and Stain Detection Using Transfer Learning

Daniel Rocha ^{1,2,3}, Filomena Soares ^{1,*}, Eva Oliveira ² and Vitor Carvalho ^{1,2,*}

¹ Algoritmi Research Center/LASI, University of Minho, 4800-058 Guimarães, Portugal

² 2AI, School of Technology, Polytechnic Institute of Coimbra and Ave, 4750-810 Barcelos, Portugal

³ INL—International Nanotechnology Laboratory, 4715-330 Braga, Portugal

* Correspondence: soares@dei.uminho.pt (ES); vcarvalho@pt.pt (VC.)

Abstract: The ways in which people dress, as well as the styles that they prefer for different contexts and occasions, are part of their identity. Every day, blind people face limitations in identifying and inspecting their garments, and dressing can be a difficult and stressful task. Taking advantage of the great technological advancements, it becomes of the utmost importance to minimize, as much as possible, the limitations of a blind person when choosing garments. Hence, this work aimed at categorizing and detecting the presence of stains on garments, using artificial intelligence algorithms. In our approach, transfer learning was used for category classification, where a benchmark was performed between convolutional neural networks (CNNs), with the best model achieving an F1 score of 91%. Stain detection was performed through the fine tuning of a deep learning object detector, i.e., the mask R (region-based)-CNN. This approach is also analyzed and discussed, as it allowed us to achieve better results than those available in the literature.

Keywords: blind people; clothing recognition; stain detection; transfer learning; deep learning

Table 5. Test performance results.		
Network	Train Accuracy	Validation Accuracy
ResNet-50	0.946	0.903
ResNet-18	0.930	0.904
MobileNet V2	0.917	0.898
GoogLeNet	0.940	0.907
MobileNet V3 (small)	0.917	0.883
EfficientNet-B0	0.927	0.900

GoogLeNet

Topics: Clothing Category Classification and Stain Detection Using Transfer Learning, Appl. Sci. 2023, 13, 1925. <https://doi.org/10.3390/app13011925>
Academic Editor: Chao-Xia Zhang
Received: 22 December 2022
Review: 19 January 2023
Accepted: 21 January 2023
Published: 2 February 2023



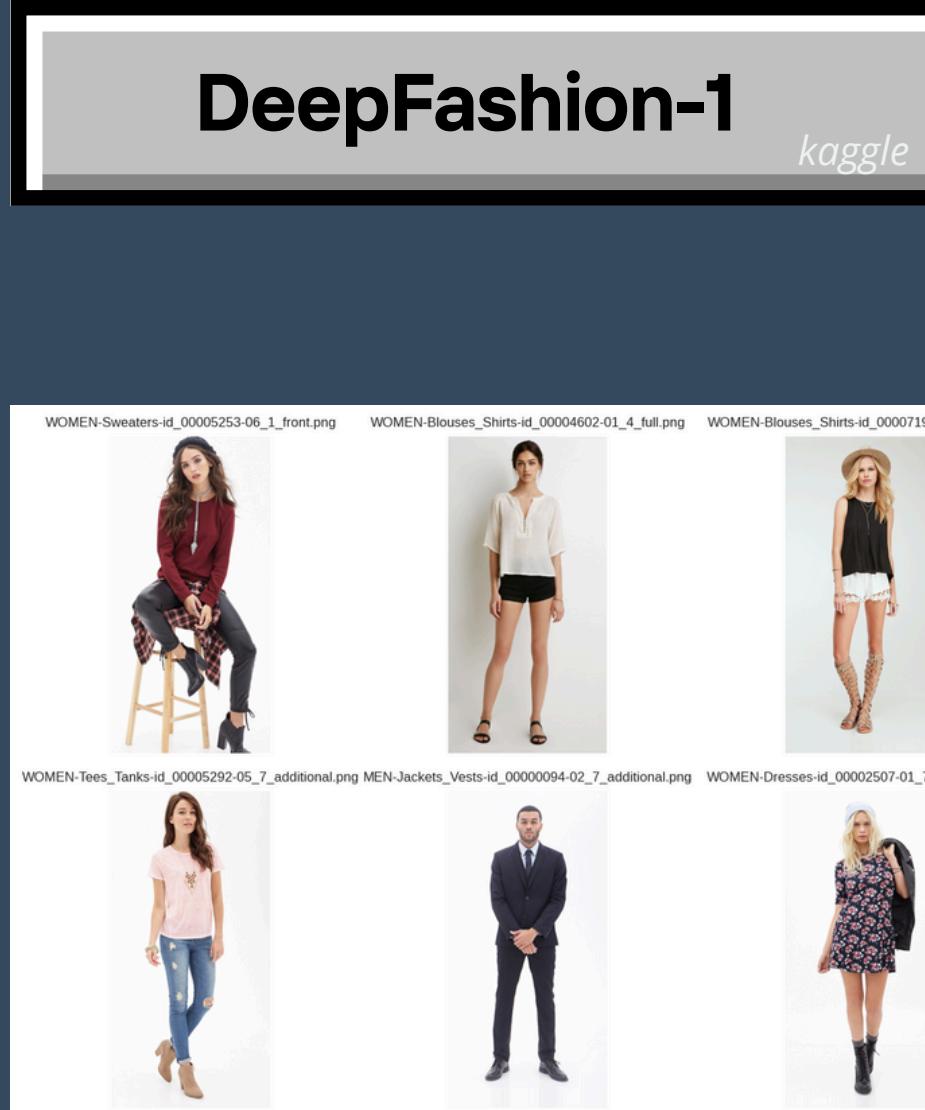
Copyright © 2023 by the authors. License MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Appl. Sci. 2023, 13, 1925. <https://doi.org/10.3390/app13011925>

<https://www.mdpi.com/journal/applsci>

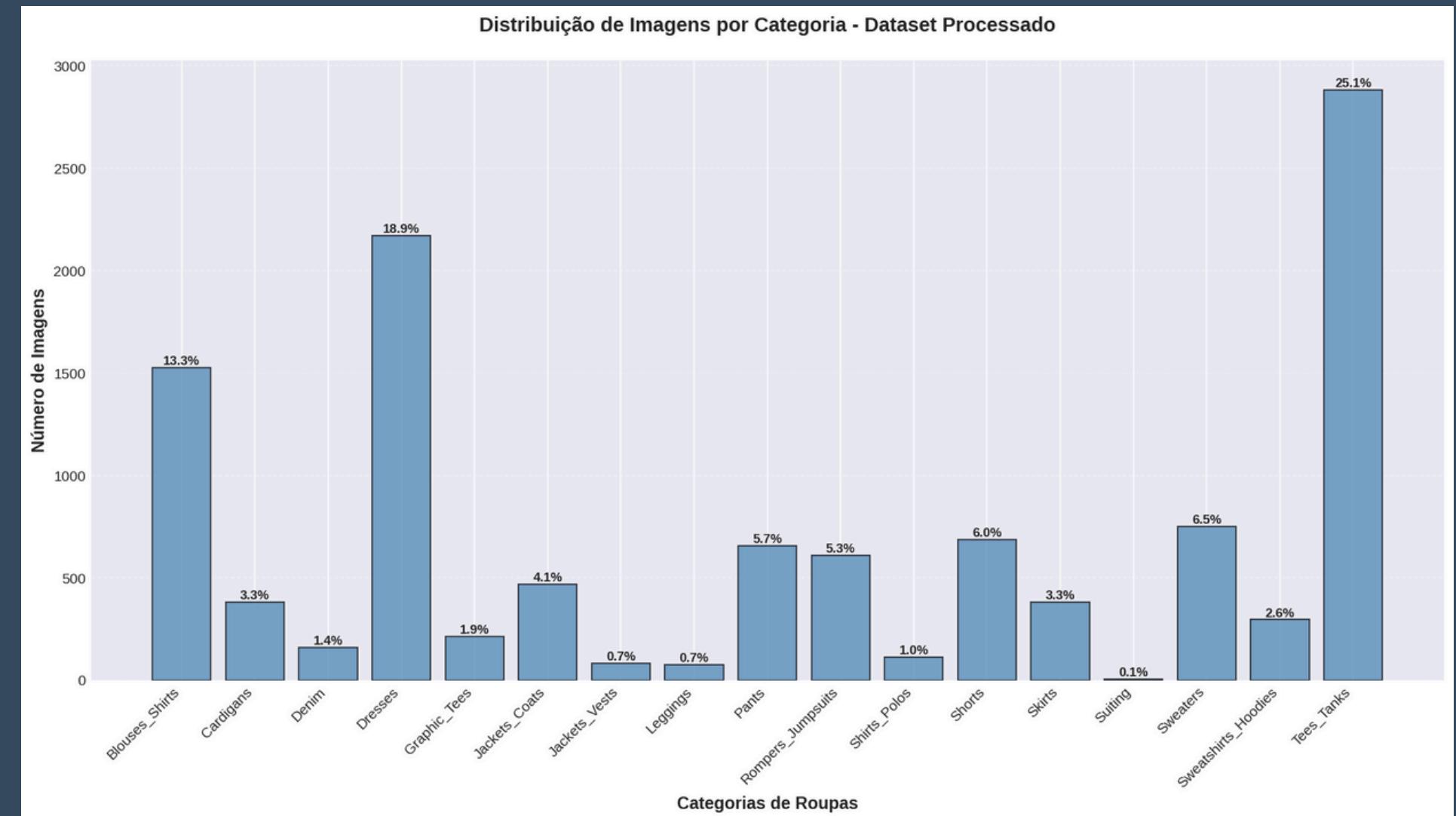
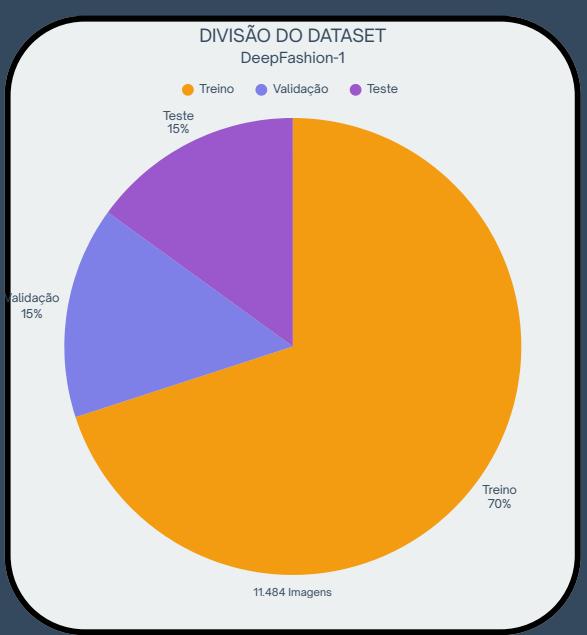
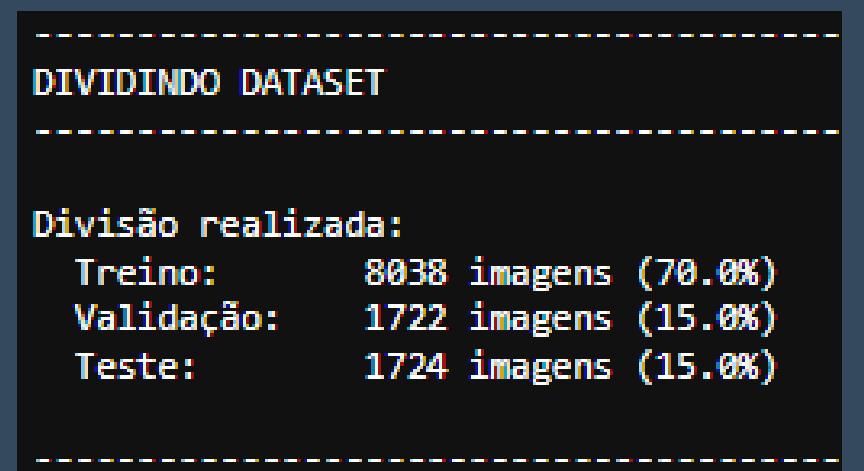
Entre os modelos testados no artigo de referência, a **GoogLeNet** apresentou o melhor desempenho (acurácia de validação, de **90,7%**).

Metodologia: Dataset



DeepFashion-1

kaggle



Mosaico das primeiras 10 imagens do dataset DeepFashion-1.

O dataset DeepFashion-1, originado do Kaggle, contém 11.484 imagens distribuídas em 17 categorias de vestuário, apresentando um desafio crítico devido ao desbalanceamento extremo entre classes.



Pré-processamento e Aumento de Dados

- Redimensionamento: 224×224 pixels
- Normalização: parâmetros de média e desvio-padrão do ImageNet
- Data Augmentation:
 - RandomHorizontalFlip
 - RandomRotation (até 15°)
 - ColorJitter (brilho, contraste, saturação)
- Desafio central: severo desbalanceamento de classes
 - Classe majoritária (Tees_Tanks): 25.12%
 - Classe minoritária (Suiting): 0.07%

Resultados dos Modelos Comparados

- **GoogLeNet ***
- **ResNet-50**
- **MobileNet-v2**
- **EfficientNet-B0**

Comparação de Performance dos Modelos

EfficientNet-B0

melhor desempenho

Modelo	Train Acc (%)	Val Acc (%)	Test Acc (%)	Precision (%)	Recall (%)	F1-Score (%)	Épocas	Tempo (s)
GOOGLENET	74.05	62.66	63.07	63.67	63.07	62.73	10	955.0
RESNET50	74.51	64.17	66.84	68.71	66.84	65.85	10	1060.0
MOBILENET_V2	80.17	66.20	67.48	68.59	67.48	67.22	10	921.6
EFFICIENTNET_B0	89.53	70.85	70.79	70.45	70.79	70.26	10	925.2

*GoogLeNet - Melhor modelo no artigo de referência.

Otimização do Melhor Modelo

EfficientNet-B0

Class Weighting

A técnica de **class weighting** ajusta a importância de cada classe durante o treinamento, aumentando o peso das classes raras para melhorar a acurácia em dados desbalanceados.

Fine-Tuning

O **fine-tuning conservador** aplica ajustes finos nos parâmetros do modelo pré-treinado, permitindo uma adaptação eficaz aos dados específicos do projeto sem perder o que foi aprendido.

Regularização

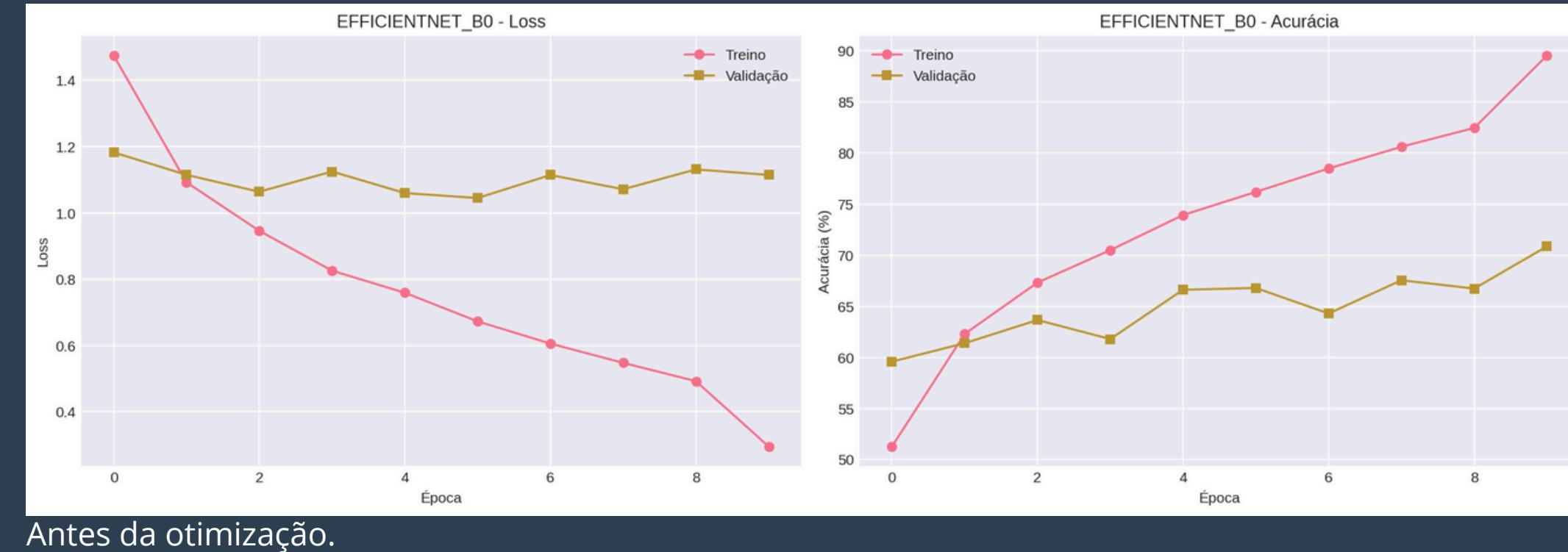
Regularização envolve técnicas como weight decay e dropout, que ajudam a evitar o overfitting, promovendo uma melhor generalização do modelo em dados não vistos.

Otimização do Melhor Modelo

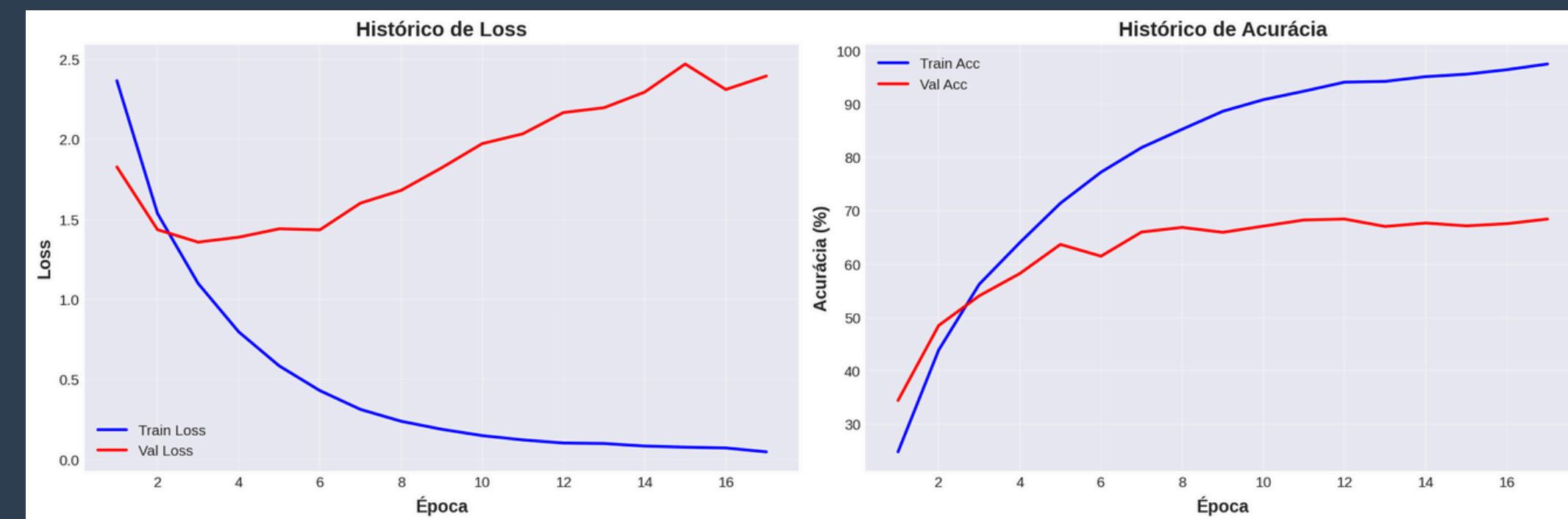
Técnicas aplicadas

- Class Weighting
- Fine-Tuning
- Dropout
- LR Scheduler
- Early Stopping
- TTA

Os resultados da otimização foram inversos ao esperado. Ocorreu o aumento de overfitting e queda no F1-Score.

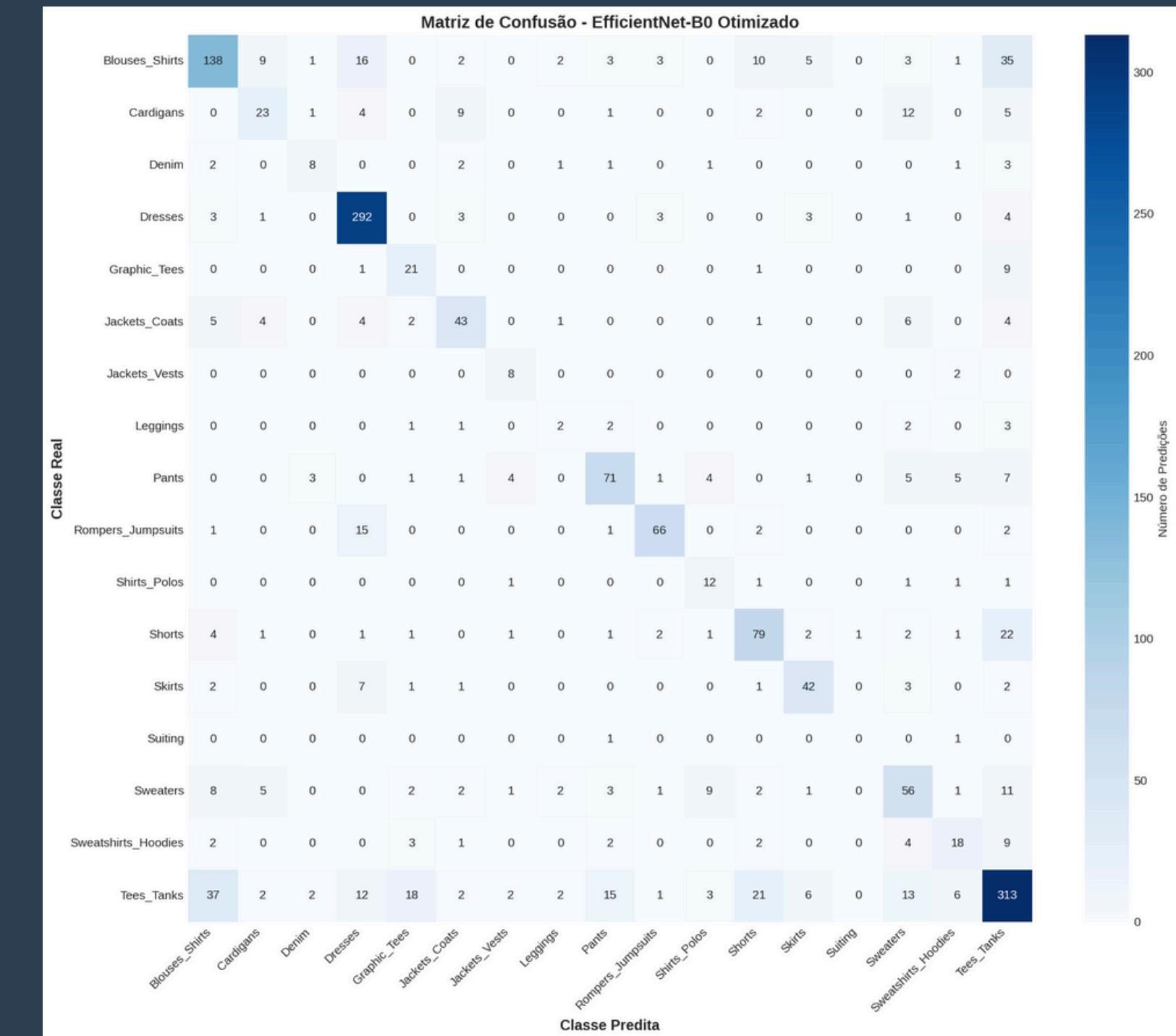
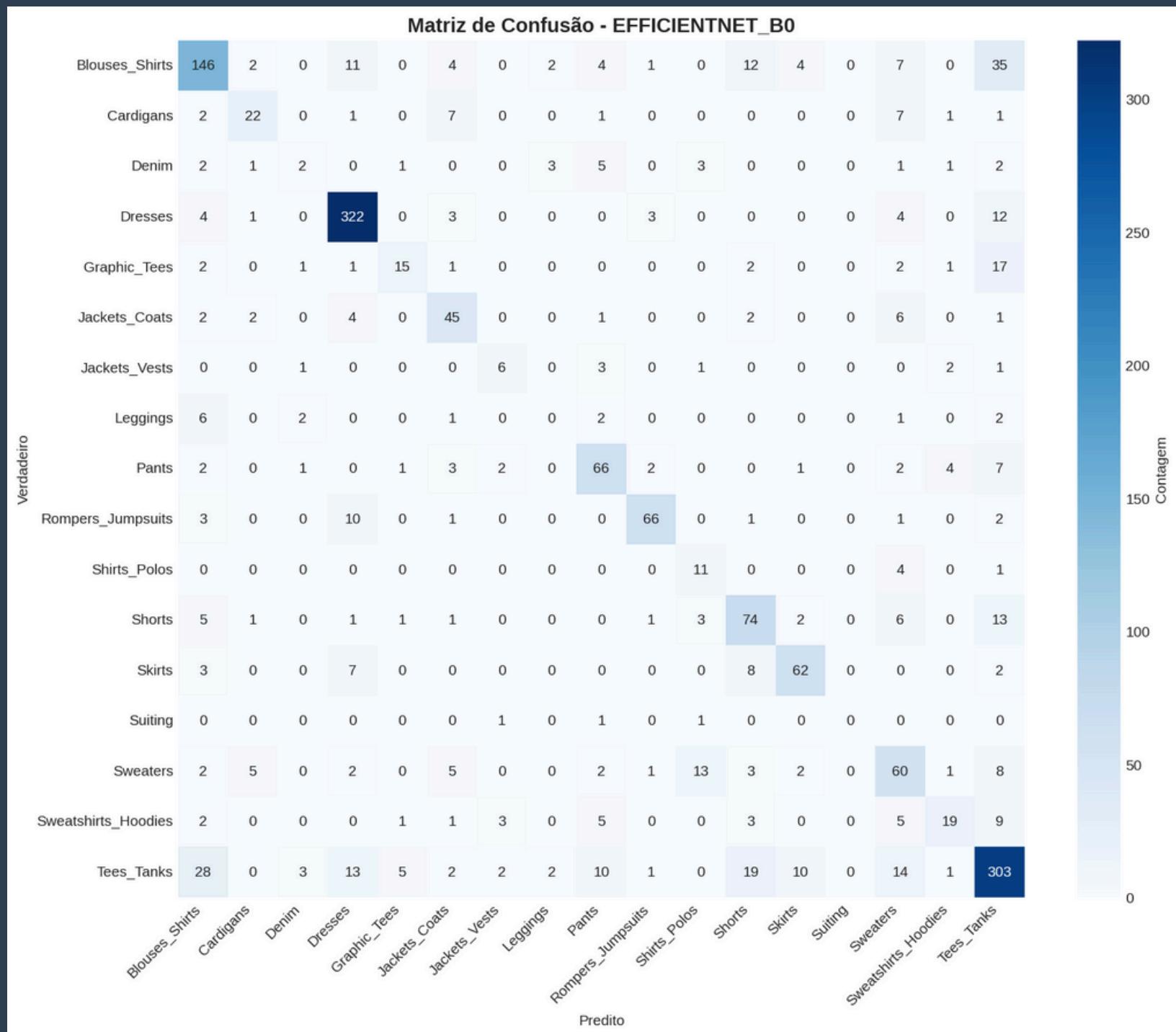


Antes da otimização.



Depois da otimização: Histórico de treinamento do modelo EfficientNet-B0 Otimizado. À esquerda, a curva de Loss mostra a divergência entre a perda de treino (azul) e a perda de validação (vermelha) a partir da Época 4. À direita, a curva de Acurácia mostra a estagnação da validação, confirmando o aumento do overfitting.

Matriz de Confusão



Antes da otimização

Depois da otimização

Comparação do EFFICIENTNET-B0 (Original) com o Modelo Otimizado e Modelo Otimizado + TTA

Comparação: EfficientNet-B0 Original vs Otimizado

Modelo	Train Acc (%)	Val Acc (%)	Test Acc (%)	F1-Score (%)
EfficientNet-B0 (Original)	89.5273631840796	70.84785133565622	70.7897793263647	70.25654893044872
EfficientNet-B0 (Otimizado)	99.88803184871858	68.46689895470384	69.19953596287704	56.13110818199595
EfficientNet-B0 (Otimizado + TTA)	99.88803184871858	68.46689895470384	69.1415313225058	55.522760130985915

- Overfitting* aumentou de 18,68% para 31,42%, reduzindo a capacidade de generalização.
- O uso de Class Weighting e Learning Rate baixo em dataset desbalanceado levou à memorização das classes minoritárias, reduzindo significativamente o F1-Score.

*Overfitting = Train Acc - Val Acc = 99.89% - 68.47% = 31.42%

Considerações Finais

O projeto atingiu o objetivo de comparar diferentes arquiteturas de Deep Learning na classificação de roupas, validando a viabilidade de sistemas assistivos para pessoas com deficiência visual. O EfficientNet-BO destacou-se como o modelo mais promissor, alcançando 70,79% de acurácia e 70,26% de F1-Score na baseline. Contudo, a etapa de otimização revelou limitações: o uso de Class Weighting e Learning Rate baixo aumentou o overfitting (de 18,68% para 31,42%) e reduziu o desempenho, evidenciando que o desbalanceamento severo exige estratégias mais robustas. Em síntese, o estudo demonstrou a eficácia do Transfer Learning e ressaltou a importância de calibrar técnicas de balanceamento em cenários multiclasse desbalanceados.

Sugestões para Trabalhos Futuros

- Classificação Multi-Label: categoria + cor da roupa em predição simultânea.
- Balanceamento de Dados: uso de oversampling/SMOTE e curadoria de classes minoritárias.
- Aplicação Acessível: protótipo móvel/web com feedback de áudio em tempo real.

Obrigada pela atenção!

Jacqueline Navarro da Silva
jacqueline.navarro@ufrpe.br

