

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

Referências



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.

Table 5. Test performance results.

Network	Train Accuracy	Validation Accuracy
ResNet-50	0.946	0.903
ResNet-18	0.939	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

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

applied
sciences

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 Vítor Carvalho ^{1,2,*}

¹ Algorithm Research Centre/LASI, University of Minho, 4800-058 Guimarães, Portugal

² 2A4, School of Technology, Polytechnic Institute of Castelo and Ave, 4750-810 Barcelos, Portugal

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

* Correspondence: danielr@lasi.uminho.pt (D.R.); vsoares@lasi.uminho.pt (F.S.); vcarvalho@lasi.uminho.pt (V.C.)

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

1. Introduction

Vision is one of the senses that dominates our lives. It allows us to form perceptions of the surrounding world and give meaning to objects, concepts, ideas, and tastes [1]. Therefore, any type of visual loss can have a great impact on our daily routines, significantly affecting even the simplest tasks in our day-to-day habits. Vision loss can be sudden and severe, or the result of a gradual deterioration, where objects at great distances become increasingly difficult to see. Therefore, the wording “vision impairment” encompasses all conditions in which vision deficiency exists [2]. The individual who is born with the sense of sight and later loses it stores visual memories and can remember images, lights, and colors. This particularity is of the highest importance for re-adaptation. On the other hand, those who are born without the capacity of seeing can never form or possess visual memories. For both cases, clothing represents a demanding challenge.

Recently, there has been an increasing focus on assistive technology for people with visual impairments and blindness, aiming at improving mobility, navigation, and object recognition [3–5]. Despite the high technology already available, some gaps remain, particularly in the area of aesthetics and image.

The ways in which we dress and the styles that we prefer for different occasions are part of our identity [6]. Blind people do not have this sense, and dressing can be a difficult and stressful task. Taking advantage of the unprecedented technological advancements of recent years, it becomes essential to minimize the key limitations of a blind person when it comes to managing garments. Not knowing the colors, the types of patterns, or even the state of garments makes dressing a daily challenge. Moreover, it is important to keep in mind that blind people are more likely to have stains on their clothes, as they face more challenges in handling objects and performing simple tasks, such as eating, cleaning and

Appl. Sci. 2023, 13, 1925. <https://doi.org/10.3390/app13031925>

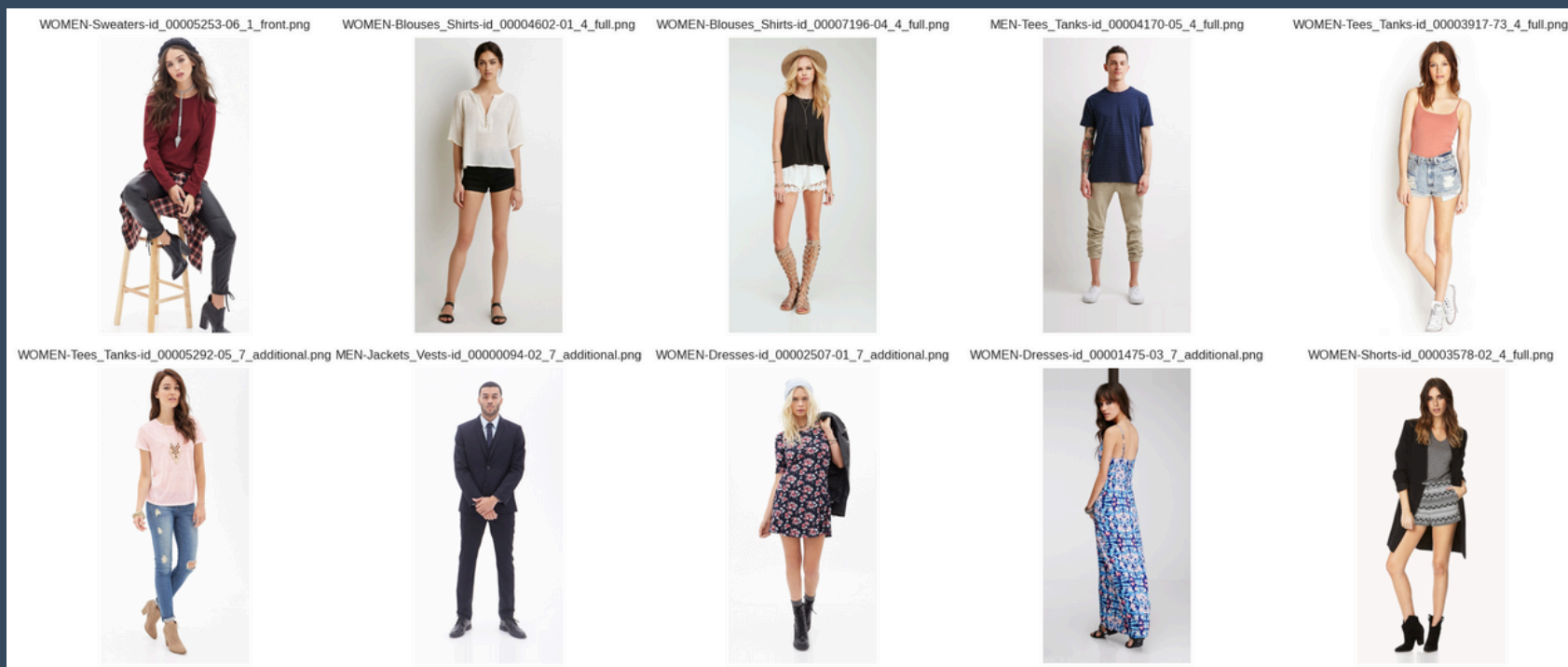
<https://www.mdpi.com/journal/applsci>

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

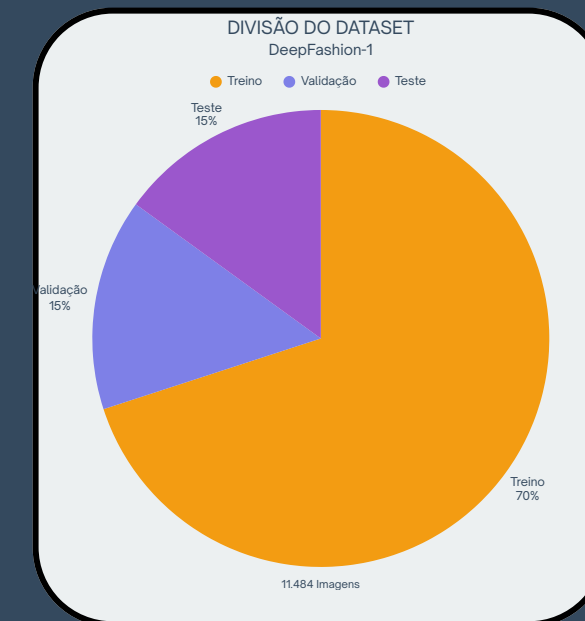


Mosaico das primeiras 10 imagens do dataset DeepFashion-1.

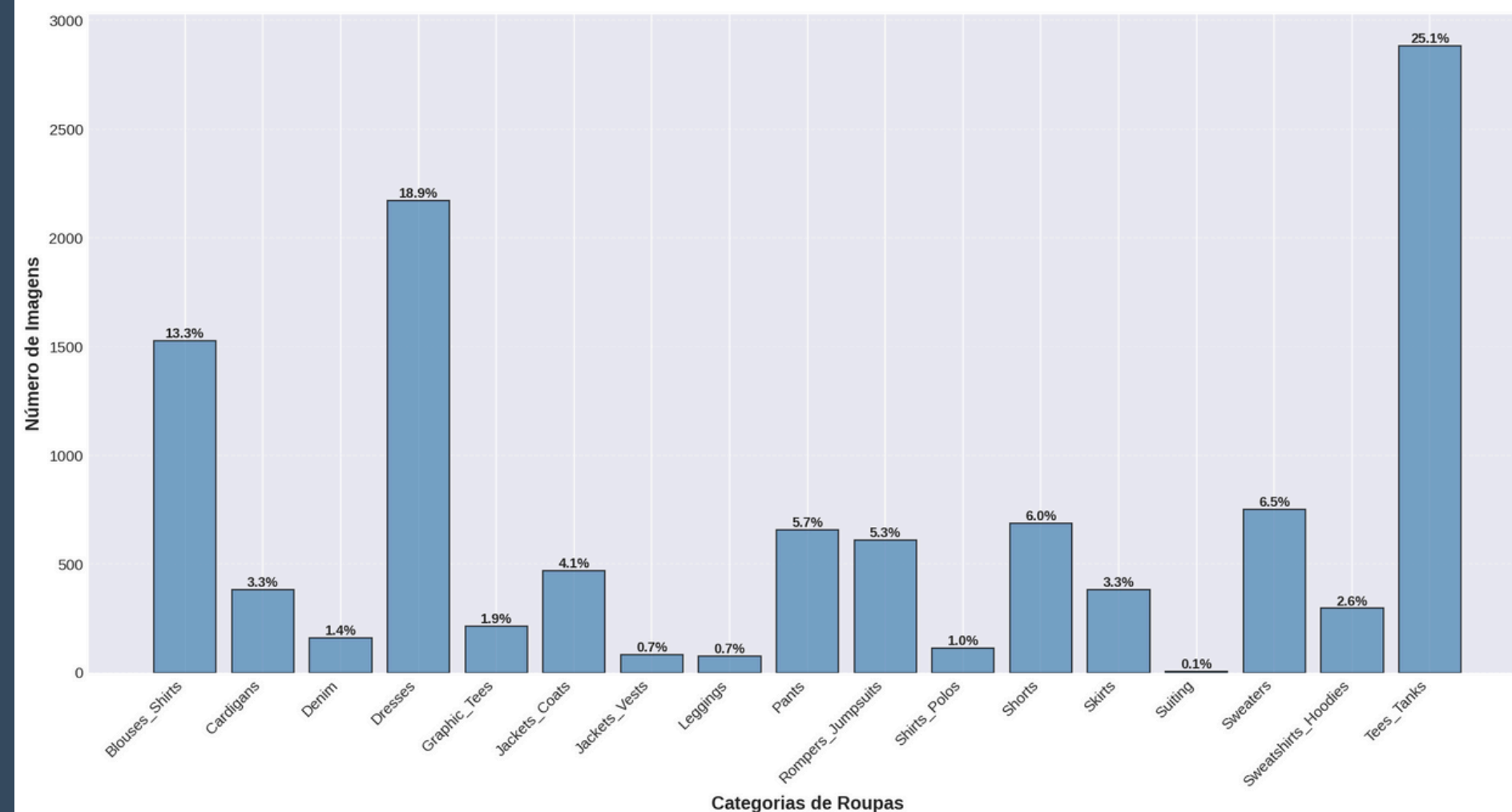
DIVIDINDO DATASET

Divisão realizada:

Treino: 8038 imagens (70.0%)
Validação: 1722 imagens (15.0%)
Teste: 1724 imagens (15.0%)



Distribuição de Imagens por Categoria - Dataset Processado



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.



MEN-Tees_Tanks-id_00000572-01_4_full.png
(Camisetas e Regatas)

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)
GOOGLNET	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

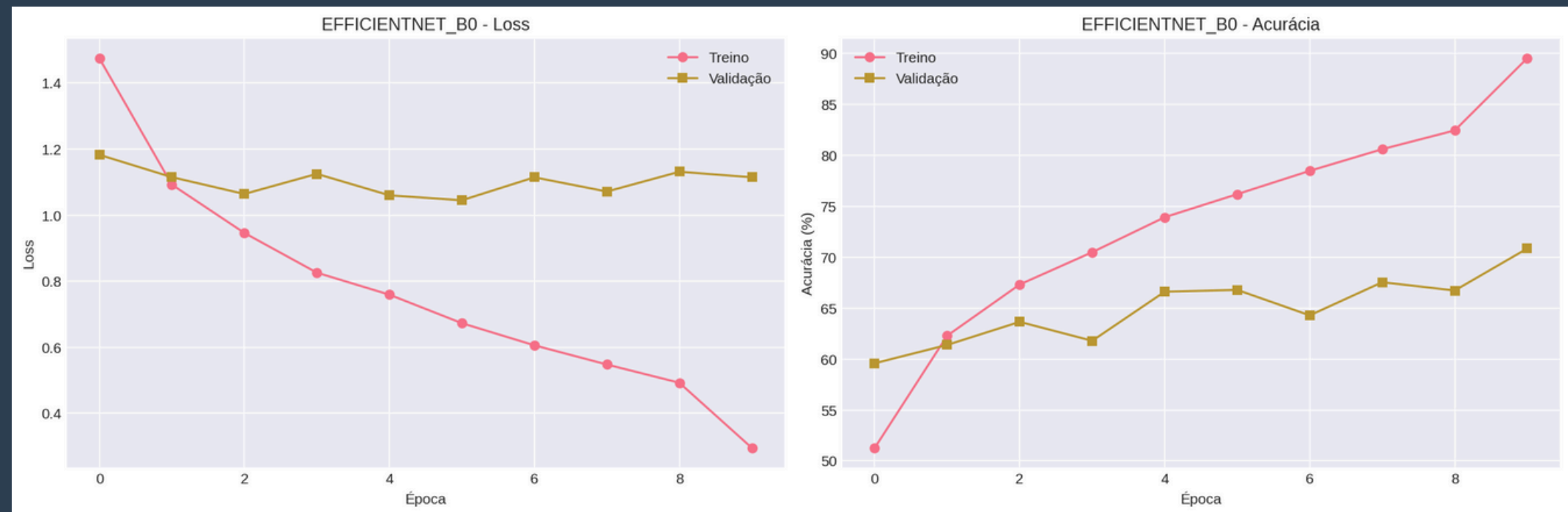
Regularização envolve técnicas como weight decay e dropout, que ajudam a evitar o overfitting, promovendo uma melhor generalização do modelo em dados não vistos.

Otimização do Melhor Modelo

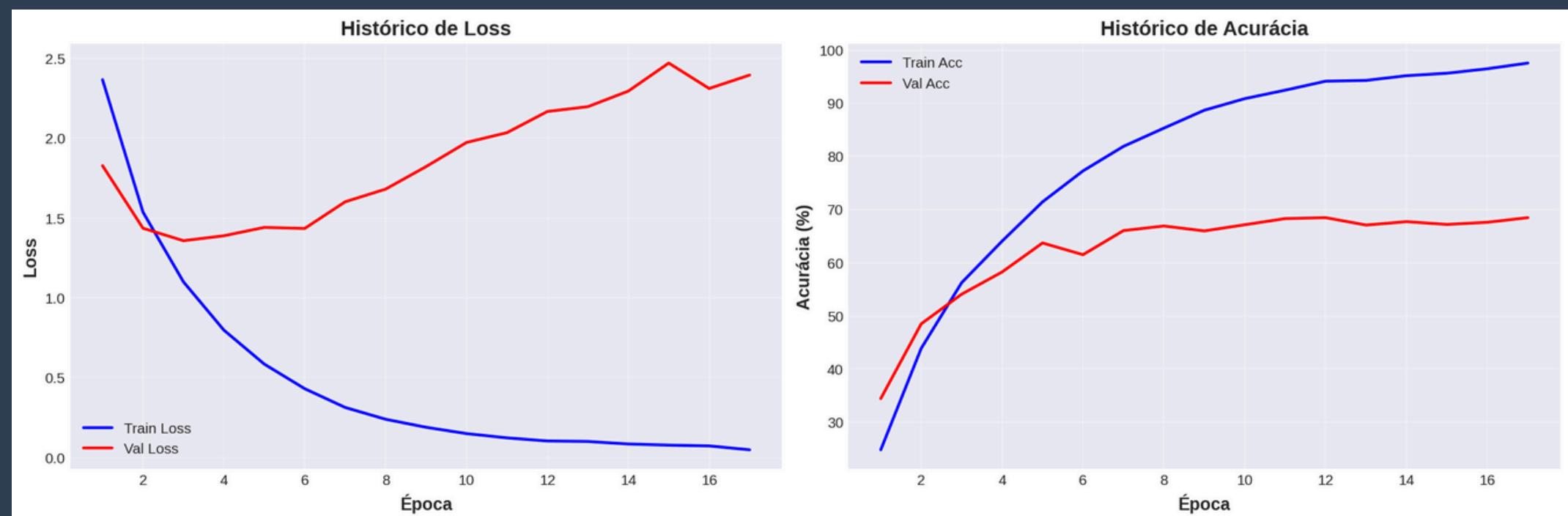
Técnicas aplicadas

- 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.

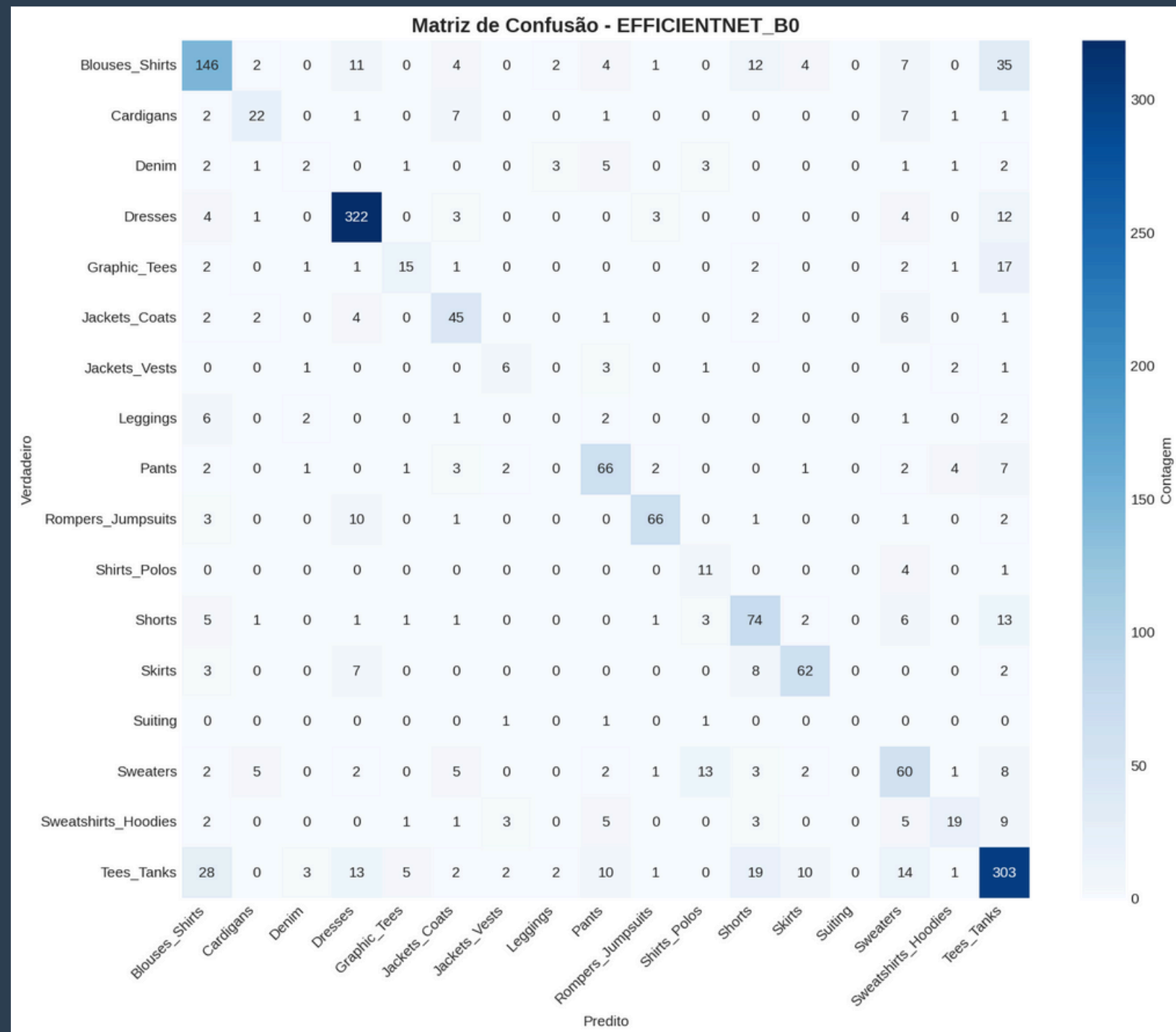


Antes da otimização.

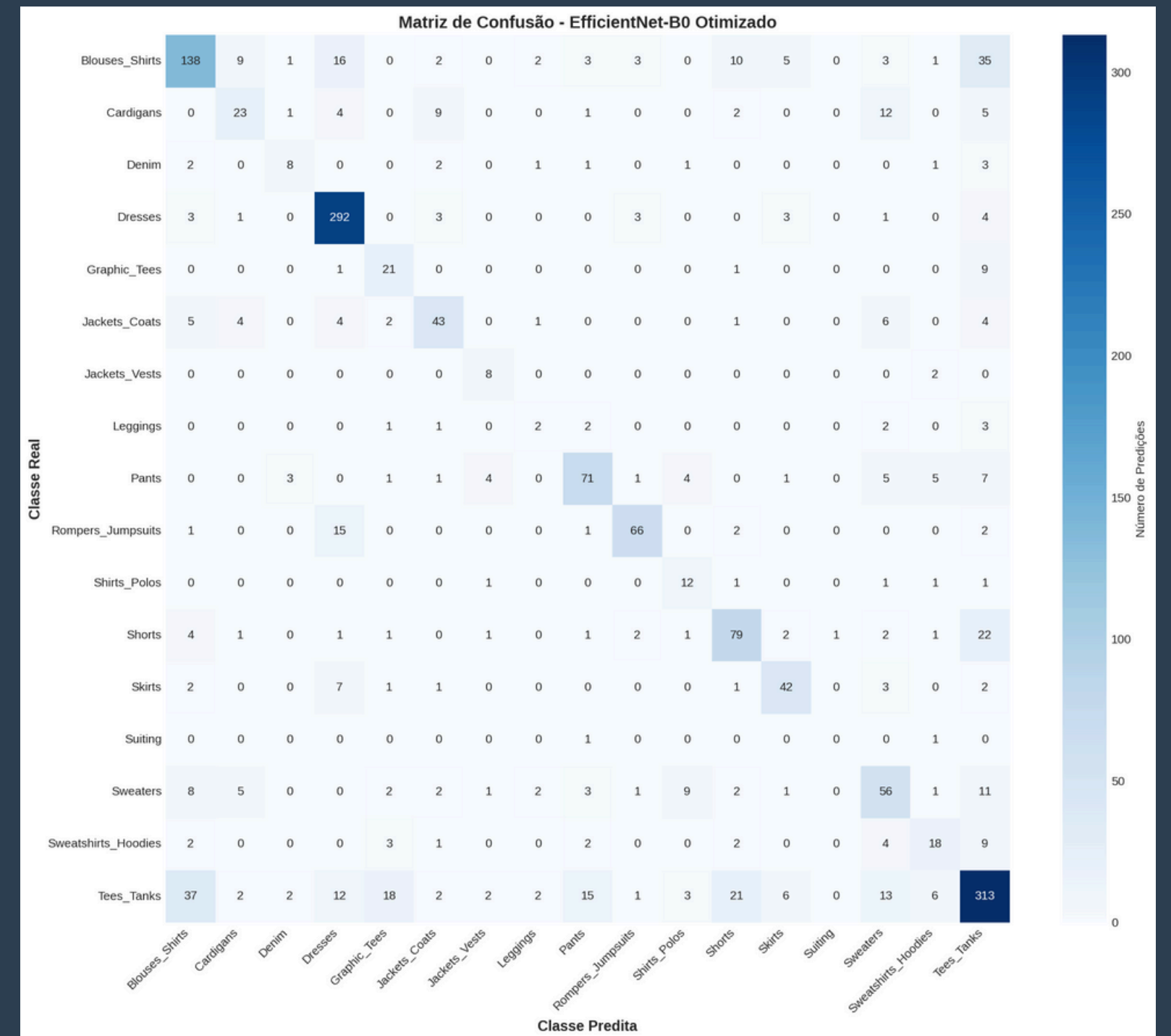


Depois da otimização: Histórico de treinamento do modelo EfficientNet-B0 Otimizado. À esquerda, a curva de Loss mostra a divergência entre a perda de treino (azul) e a perda de validação (vermelho) a partir da Época 4. À direita, a curva de Acurácia mostra a estagnação da validação, confirmando o aumento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

