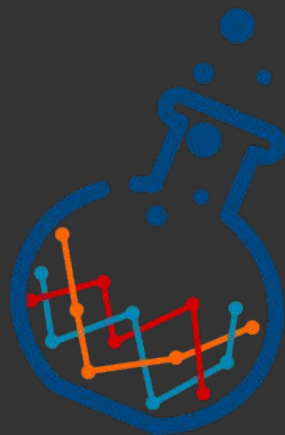


TODATA



LAMIA

Projeto final LAMIA - MOACIR

Mais Outro Algoritmo Classificador de Imagens
da Roça

O seu roceiro virtual

lucasparteka@alunos.utfpr.edu.br
github.com/TODOHARDWARE

INTRODUÇÃO

Após fracassar tentando resolver uma demanda com séries temporais de algo com dados escassos, decidi fazer um projeto mais simples e com dados abundantes.

Desta vez, um algoritmo baseado num desafio do Kaggle onde ele irá classificar doenças em pés de mandioca analisando fotos de suas folhas.

Datasets

Os datasets usados se encontram em:

<https://www.kaggle.com/datasets/visalakshiiyer/cassava-image-dataset>

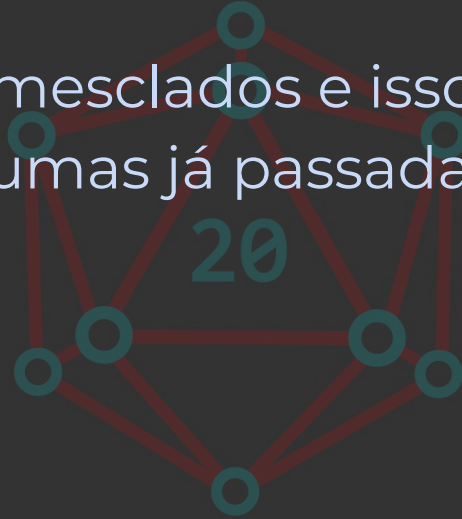
<https://www.kaggle.com/c/cassava-disease/data>

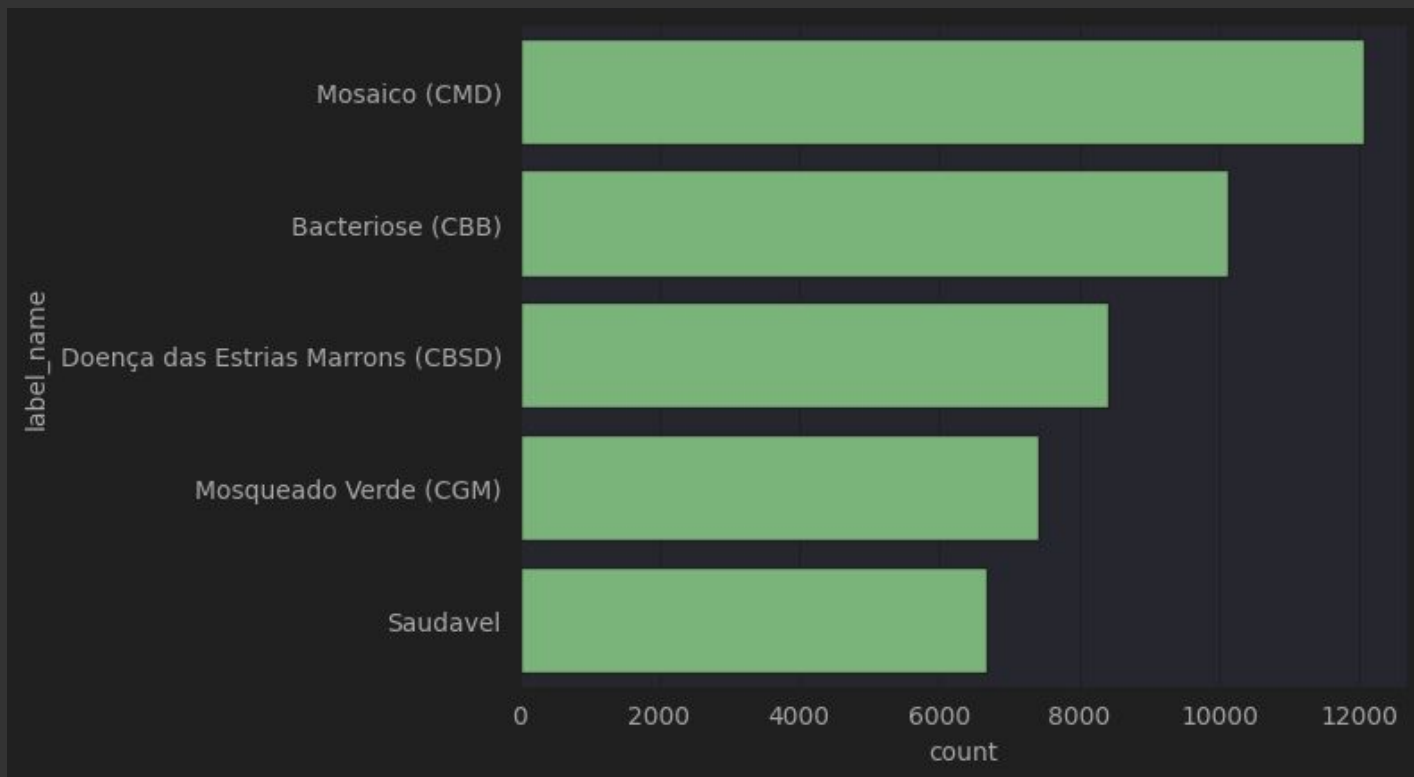
<https://data.mendeley.com/datasets/bwh3zbpkpv/1>

<https://www.kaggle.com/competitions/cassava-leaf-disease-classification/data>

Tratamento dos dados

Os datasets foram mesclados e isso resultou em 44.721 imagens, algumas já passadas pelo processo de augmentation.





Detalhes da rede

Neste projeto será usada a técnica de aprendizado por transferência, onde o nosso algoritmo se baseia em outro pré-treinado.

A nossa vítima será o ResNext-101-32x8d.

Como Framework será usado o PyTorch.

As imagens foram redimensionadas para 224x224.



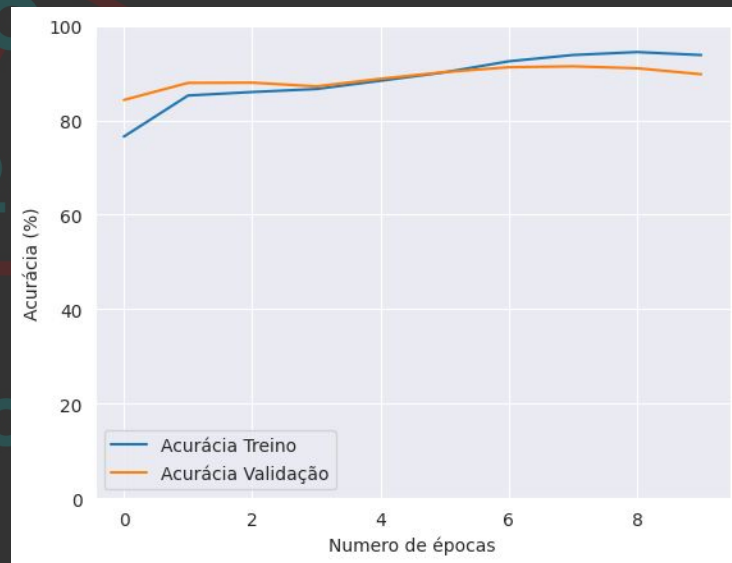
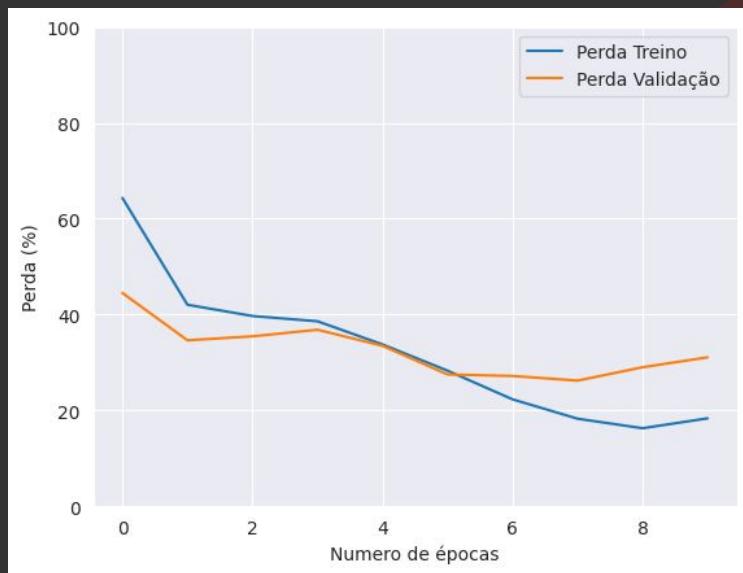
Foi realizado também augmentation usando a biblioteca augmentations.

Para divisão dos dados foi utilizado o Scikit-Learn Train Test Split, com 25% dos dados para validação.

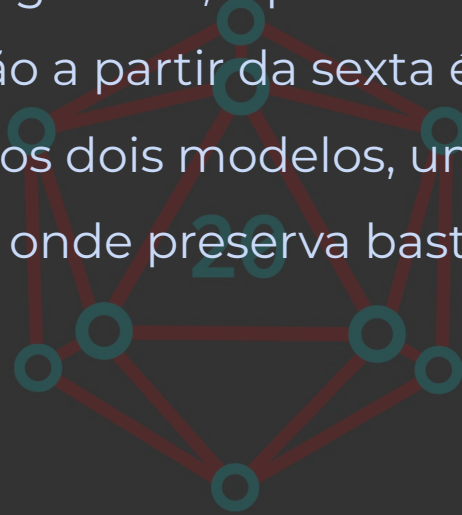
O Learning Rate foi ajustado automaticamente durante o treino usando Circular Learning Rate.

O melhor Batch Size foi de 16 tanto para treino quanto validação e o melhor valor para dropout foi de 30%.

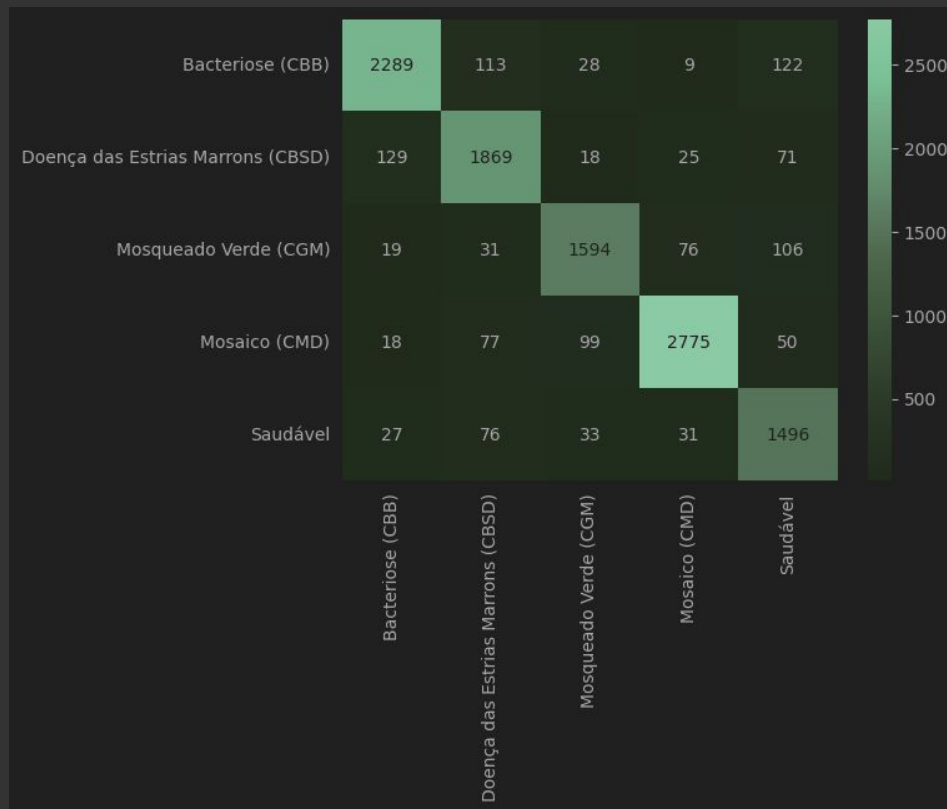
Foram rodadas 10 épocas para analisar o overfitting.



Como podemos ver nos gráficos, a perda do treino começou a ficar menor que a da validação a partir da sexta época, indicando o overfitting. Foram gerados dois modelos, um com 6 épocas e outro apenas com uma época onde preserva bastante peso do modelo ResNext.



Teste com dados de validação 6 épocas:



	precision	recall	f1-score	support
Bacteriose (CBB)	0.92	0.89	0.91	2561
Doença das Estrias Marrons (CBSD)	0.86	0.88	0.87	2112
Mosqueado Verde (CGM)	0.90	0.87	0.89	1826
Mosaico (CMD)	0.95	0.92	0.94	3019
Saudavel	0.81	0.90	0.85	1663
accuracy			0.90	11181
macro avg	0.89	0.89	0.89	11181
weighted avg	0.90	0.90	0.90	11181

Teste com dados de validação uma época:



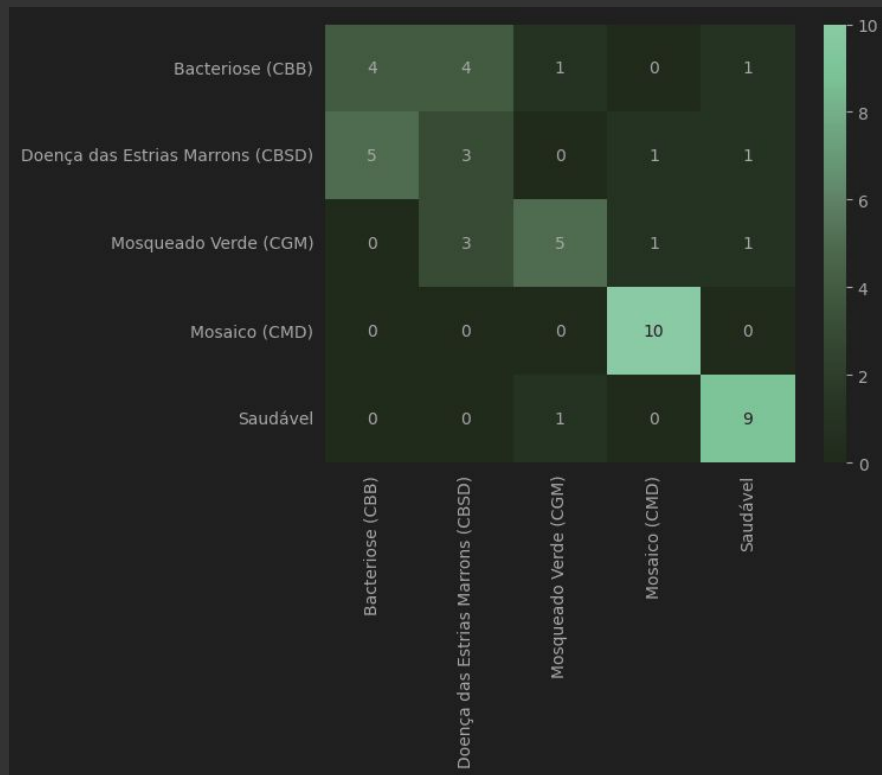
	precision	recall	f1-score	support
Bacteriose (CBB)	0.93	0.77	0.84	2561
Doença das Estrias Marrons (CBSD)	0.80	0.86	0.83	2112
Mosqueado Verde (CGM)	0.85	0.88	0.86	1826
Mosaico (CMD)	0.96	0.88	0.92	3019
Saudável	0.74	0.93	0.82	1663
accuracy			0.86	11181
macro avg	0.85	0.86	0.85	11181
weighted avg	0.87	0.86	0.86	11181

Foram selecionadas 50 imagens da internet, 10 de cada categoria.
Resultados modelo treinado em 6 épocas:



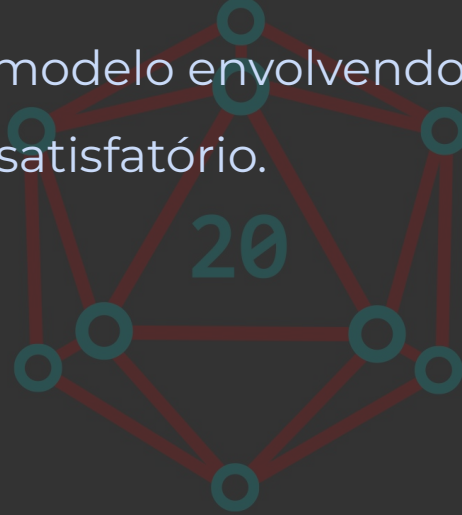
	precision	recall	f1-score	support
Bacteriose (CBB)	0.46	0.60	0.52	10
Doença das Estrias Marrons (CBSD)	0.25	0.20	0.22	10
Mosqueado Verde (CGM)	1.00	0.60	0.75	10
Mosaico (CMD)	0.83	1.00	0.91	10
Saudável	0.91	1.00	0.95	10
accuracy			0.68	50
macro avg	0.69	0.68	0.67	50
weighted avg	0.69	0.68	0.67	50

Foram selecionadas 50 imagens da internet, 10 de cada categoria.
Resultados modelo treinado em uma época:



	precision	recall	f1-score	support
Bacteriose (CBB)	0.44	0.40	0.42	10
Doença das Estrias Marrons (CBSD)	0.30	0.30	0.30	10
Mosqueado Verde (CGM)	0.71	0.50	0.59	10
Mosaico (CMD)	0.83	1.00	0.91	10
Saudável	0.75	0.90	0.82	10
accuracy			0.62	50
macro avg	0.61	0.62	0.61	50
weighted avg	0.61	0.62	0.61	50

Dada a similaridade entre Bacteriose e Doença das Estrias Marrons, houve bastante erro do modelo envolvendo esses dois, mas de resto foi obtido um resultado satisfatório.



Referências

Cassava Leaf Disease Classification; Identify the type of disease present on a Cassava Leaf image

<https://www.kaggle.com/c/cassava-leaf-disease-classification/data>

Pytorch Complete Tutorial | Cassava Leaf Disease Detection | Python

<https://www.youtube.com/watch?v=R7fKir4gtSc>

Cassava Leaf Disease Detection - Image Classification

<https://github.com/aswintechguy/Deep-Learning-Projects/tree/main/Cassava%20Leaf%20Disease%20Detection%20-%20Pytorch%20Image%20Classification>