

Classification of Coffee Biotic Stresses Using Convolutional Neural Networks and Enhanced Image Preprocessing Techniques

Daniel Imai Yamakawa¹, Pedro Moises de Sousa², Pedro Henrique Aparecido Damaso de Melo³

Instituto de Ciências Exatas e Tecnológicas

Universidade Federal de Viçosa (UFV) - Campus Rio Paranaíba

Rodovia MG230, Km7, Caixa Postal 22 - 38.810-000 - Rio Paranaíba - MG - Brasil

¹daniel.yamakawa@ufv.br, ²pedromois@ufv.br ³pedro.henrique.melo@ufv.br

Abstract—Coffee is a big commodity in Brazil, however it's production is threatened by several biotic stresses. The objective of this paper is the classification of these biotic stresses that attack the leaf, such as leaf miner, rust, cercospora, and brown leaf spot, comparing several methods of image preprocessing to reach a better performance and finding the best one for this task. Using neural networks, part of artificial intelligence, we obtained good results using the networks ResNet50, MobileNetV2 and AlexNet using images with filters such as clahe, gaussian, wavelet and Graythresh. The networks achieved good result, varying from 93% to 98% using certain filters. This paper can aid farmers in the classification of these biotic stresses and can also help future researchers in this area, demonstrating the impact of each filter used in this research and it's benefits in this type of task.

I. INTRODUÇÃO

O café é uma das bebidas mais apreciadas do mundo inteiro. No período de 2020/2021, o Brasil se destacou como o maior exportador da bebida, com 38.12 milhões de sacas de café exportadas, o que representa 29.6% das exportações mundiais nesse período, destacando sua importância na produção da bebida [1]. Porém, diversos desafios no plantio do café afetam seriamente a qualidade dos grãos e sua produção, como os estresses bióticos. Entre eles estão a cercospora, a doença da phoma, a ferrugem e o bicho-mineiro [2] que se manifestam nas folhas causando em um grave prejuízo aos plantadores de café.

Atualmente, a classificação desses estresses bióticos é realizada presencialmente, o que pode ser um obstáculo para sua conclusão, dependendo de uma visita de um especialista na plantação de café para realizar esse processo. Não podemos contar com a disponibilidade de um profissional a todo momento, e além disso a classificação manual está sujeita a erros humanos [2].

A tecnologia está evoluindo rapidamente e está sendo utilizada para automatizar processo de classificação [3]. Como exemplo de trabalhos na área de agricultura temos [4], um sistema de detecção de doenças do milho, [5] identificação das doenças da soja e [6] predição da produção de algodão, mostrando a eficiência e eficácia dessa tecnologia. A aplicação dessa inovação também é utilizada na classificação dos estresses bióticos do café, obtendo resultados confiáveis e rápidos, o que se torna essencial para solucionar esse prob-

lema, podendo agilizar a compra e aplicação de produtos que combatem esses estresses bióticos, além de outras providências necessárias [7].

As Redes Neurais Convolucionais (CNNs) são parte da inteligência artificial, amplamente utilizadas para a classificação de imagens agrárias, mostrando resultados precisos como no estudo [8] que obtém uma alta acurácia na classificação dos estresses bióticos bicho-mineiro, ferrugem, cercospora e a doença da phoma. Alguns estudos como [9] deixam o processo ainda mais prático, dependendo de um smartphone para realizar a classificação das imagens, demonstrando seu potencial na abordagem. Em [2] houve a criação de uma arquitetura própria para a realização da classificação das doenças do café, e o estudo [8] avalia o grau de severidade das doenças.

Em vista de resolver os problemas citados, esse trabalho tem como objetivo desenvolver e estudar abordagens utilizando CNNs e técnicas de processamento de imagens. Ferramentas para reconhecer os estresses bióticos apresentados nas folhas como o bicho-mineiro, doença da phoma, ferrugem e cercospora para auxiliar na classificação e no combate contra essas anormalidades, utilizando diversos métodos para aumentar a taxa de acerto.

II. REFERENCIAÇÃO TEÓRICA

Para identificar as doenças das folhas de café é necessário absorver diversas características. O deep learning e o machine learning, são áreas da Inteligência Artificial e se encaixam muito bem para a tarefa.

A. Redes Neurais Convolucionais (CNNs)

A ideia principal do CNN é a obtenção de características de uma entrada (geralmente uma imagem) e quanto mais a camada se aprofunda, mais complexos são os detalhes extraídos. É importante apontar que esse processo não é barato falando sobre o custo computacional, então quando é utilizada uma grande base de dados usa-se o GPU [10]. Existe vários modelos de CNNs, os mais famosos para a identificação de doenças agrárias e são o AlexNet [11], criado em 2012 sendo revolucionário para a época com resultados bons e consistentes, o VGGNet [12] com uma taxa de erro de 7,32%, uma das melhores apresentadas em 2014, GoogLeNet [13] que

ganhou a ILSVRC 2015 com uma taxa de erro de 6.7%, e o ResNet [14] proposto pela *Microsoft Research Asia* e tem um dos melhores históricos para reconhecer doenças na área agrícola e na área de medicina [10].

1) **Arquitetura de uma CNN:** As CNNs consistem em uma sequência de diferentes tipos de camadas para diferentes tarefas. Normalmente uma CNN possui uma camada convolucional, uma camada de ativação da função, uma camada de pooling, uma camada totalmente conectada e uma camada de eliminação. Todas essas camadas juntas formam uma arquitetura funcional de CNN, que podem incluir também outras camadas opcionais como de normalização e eliminação para resolver o problema do *overfitting*, um problema comum que ocorre quando o modelo se acostuma muito com uma situação e não consegue generalizar [10].

2) **Camadas Convolucionais:** As camadas convolucionais são a base fundamental de uma CNN que usa operações convolucionais. Seus parâmetros são filtros chamados *kernels*. A tarefa principal dessas camadas são a detecção das características em regiões locais e mapear sua aparência para uma feature map, que é obtida em cada filtro na camada repetindo o mesmo processo algumas vezes até completar a imagem [10].

3) **Camada de Pooling:** Nas CNNs a próxima camada da convolucional e de ativação é a camada de *pooling*. Essa camada pega todos os mapas da camada convolucional e deixa-os resumidos. Há algumas técnicas de *pooling*, e uma delas é o *max-pooling*, onde sai o maior valor que seja possível entrar. Por exemplo, a entrada é uma região 2 x 2, o *max-pooling* irá maximizar os quatro valores. Essas operações eliminam data que não são necessárias para o funcionamento da CNN e preserva as características realmente importantes em resoluções menores [10].

4) **Camadas totalmente conectadas:** A camada final que une as características com seus estados ocultos na primeira camada totalmente conectada. É possível usar mais de uma camada totalmente conectada para melhorar o desempenho computacional. As camadas totalmente conectadas possuem a maioria dos parâmetros. Por exemplo, se cada uma de duas camadas totalmente conectadas possuir 4096 unidades ocultas, as conexões entre elas teriam mais de 16 milhões de pesos. A natureza das camadas totalmente conectadas pode depender diretamente da aplicação utilizada, uma camada totalmente conectada de uma aplicação de classificação diferenciaria de uma utilizada na segmentação por exemplo [6].

B. Wavelet em 2 dimensões

Wavelets foram mostrados como poderosas fundações para processar sinais e análise de imagens em mais de uma resolução. Uma função de transformação junta todas esses wavelets para formar uma função ou imagem. A transformação wavelet em duas dimensões são usadas frequentemente no pré-processamento de imagens e compressão. É necessário separar a função de uma dimensão wavelet em três coeficientes, sendo eles horizontal (H), vertical (V) e diagonal (D).

$$\psi^H(x, y), \psi^V(x, y), \psi^D(x, y)$$

Esses coeficientes medem diversas variações, primeiramente aplicada a primeira dimensão a todas horizontais, verticais e finalmente as diagonais. A imagem original é dividida em quatro sub-imagens: LL (*Low-Low*): A parte de baixa frequência na direção horizontal e vertical. LH (*Low-High*): A parte de baixa frequência na direção horizontal e alta frequência na direção vertical. HL (*High-Low*): A parte de alta frequência na direção horizontal e baixa frequência na direção vertical. HH (*High-High*): A parte de alta frequência em ambas as direções. Após a decomposição em componentes a imagem original é reconstruída.

III. TRABALHOS RELACIONADOS

O estudo [9] tem como objetivo facilitar o reconhecimento das doenças do café que afetam a folha utilizando CNNs através de um aplicativo. Foi utilizado o banco de dados BRACOL, um banco de dados público com um total de 1747 imagens de café tipo arábica e foi dividido em classes de folha saudáveis, bicho-mineiro, ferrugem, doença da phoma e cercospora, com adição de algumas imagens de outro trabalho e técnicas como *data collection*, *labelling* e *data augmentation* para deixar esse banco de imagens mais robusto. As arquiteturas utilizadas nesse trabalho foram a UNet [15] e o PSPNet [16] para a segmentação e AlexNet, GoogLeNet, VGG16 e ResNet50 [14] para a classificação de sintomas. O melhor resultado foi obtido pela ResNet50 com acurácia de 97.07%, precisão de 96.85% e *recall* de 96.69%.

Em [17] o estudo é centrado em diagnosticar as doenças do café Cercospora e Ferrugem comparando duas diferentes abordagens, uma abordagem estudando a textura da folha e outra abordagem utilizando uma CNN. Foi utilizado um banco de dados privado do Instituto Federal Sul de Minas - Muzambinho-MG com 750 subimagens de cada classe após o *data augmentation* ser aplicado. Para a primeira abordagem foi utilizado *Gray Level Co-occurrence Matrix (GLCM)* e o *Local Binary Pattern (LBP)*, enquanto o CNN utilizado é a arquitetura AlexNet, treinada com 20 épocas, learning rate de 0.001. No experimento utilizando o atributo da textura foi obtido 0.933 de sensibilidade no melhor experimento, enquanto no experimento utilizando a AlexNet foi obtido 0.980 sensibilidade mostrando que ambos os métodos são confiáveis para auxiliar trabalhadores de café com o AlexNet performando um pouco melhor [10].

No estudo [8] temos a classificação de doenças do café junto com o grau de severidade, estudando o bicho-mineiro, a ferrugem, a cercospora a doença da phoma. Sua base de dados é composta por 1747 imagens de café arábica com 372 imagens que mostram mais de uma doença, e outras 2147 imagens cortadas somente com os sintomas de cada doença e com adição de 575 imagens de outro banco de dados para os sintomas. As arquiteturas utilizadas foram adaptadas para o *multi-task* que seria realizar duas tarefas ao mesmo tempo, a classificação da doença e o grau de severidade da mesma. Utilizando as arquiteturas AlexNet, GoogLeNet, VGG16, ResNet50 e a MobileNetV2 [18] obtemos a acurácia da classificação de doença de, respectivamente, 96.58%, 96.82%, 95.60%, 97.07%

e 96.33% utilizando o banco de dados dos sintomas, e com o multi-task tendo o banco de dados normal temos 86.9%, 82.94%, 86.51%, 86.51% e 84.13% para o grau de severidade de cada doença [10].

IV. METODOLOGIA

Algumas etapas são necessárias para que seja possível um sistema de reconhecimento dos estresses bióticos do café conforme mostra a Fig. 1.

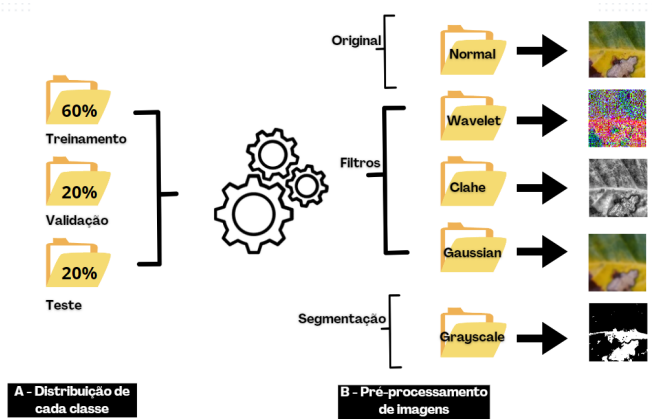


Fig. 1: Fases de pré-processamento

A. Distribuição de cada classe

Foi utilizado bancos de imagem públicos do café arábica chamados JMunBEN e JMunBEN2 para o teste, treinamento e validação [19]. É um banco de imagens feito especialmente para pesquisas relacionadas com reconhecimento de doenças e classificação da folha de café tipo arábica, contendo um total de 58.555 imagens divididas em 5 classes (doença de phoma, cescóspora, ferrugem, bicho-mineiro, folhas saudáveis) de 128x128 pixels e arquivos do tipo JPG.

O banco de imagens foi dividido em três partes: Treinamento contendo 60% das imagens, teste contendo 20% e a validação 20%. É fundamental essa divisão para o funcionamento do código como mostrado na Fig. 2.

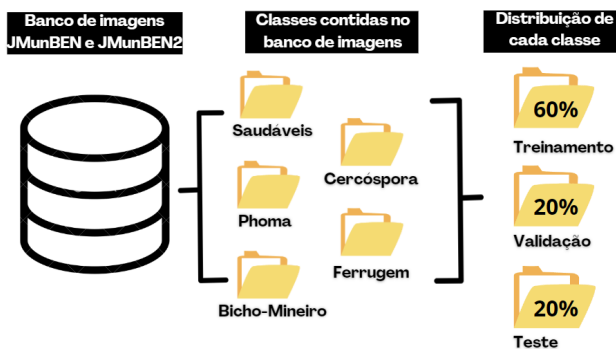


Fig. 2: Distribuição das imagens

B. Pré-processamento de imagens:

Técnicas para a melhorar a qualidade e quantidade das imagens foram utilizadas no próprio banco de imagens e filtros contendo diferentes técnicas para melhorar o resultado foram aplicadas separadamente.

1) **Cropping**: Cada imagem nesse banco de imagens foi verificada para ter certeza que todas estavam da mesma forma. Uma ferramenta de corte foi utilizada para remover áreas não necessárias da imagem para que fosse possível reduzir o tempo de treino no processamento da imagem [19].

2) **Data Augmentation**: O *data augmentation* foi realizado para fins de aumentar a quantidade de imagens disponíveis, rotacionando e invertendo-as de uma maneira para dar ênfase às áreas de interesse.

3) **Gaussian**: O filtro *Gaussian* é uma função utilizada para "borrar" as imagens a fins de remover deturpações e detalhes da imagem [20].

Ele aplica uma "média" dos pixels vizinhos centrais da imagem, proporcionando uma suavização na imagem e preservando as bordas melhor que outros filtros parecidos com ele [20].

É possível mudar os valores de quanto os detalhes serão reduzidos e nesse caso foi aplicado o sigma 2 em seu banco de dados [20].

4) **Clahe**: O CLAHE (Contrast Limited Adaptive Histogram Equalization) é um variante do AHE (*Adaptive Histogram Equalization*). A ampliação do contraste é feita pela vizinhança de um pixel e transformada pelo declive da função de transformação [21].

O CLAHE funciona em pequenas áreas de uma imagem ao invés dela completa, então dentro de cada bloco a técnica é aplicada. Para minimizar a ampliação de deturpações em excesso, é possível definir o limite de contraste. O limite aplicado foi de valor 0.02 [21].

5) **Wavelet**: O filtro de imagem wavelet foi utilizado para decompor a imagem em sinais de diferentes escalas, servindo para a remoção de ruído, compressão de imagem e mudando sua resolução [2].

Foi aplicado o filtro wavelet em 2 dimensões para obter fins de testes e comparação com outros filtros de imagens. Esse filtro contém diversas famílias e a família aplicada no dataset foi a *coif5*.

6) **Graythresh**: O *Graythresh* é uma função utilizada para calcular a binarização ideal de uma imagem utilizando o método otsu [22].

É implementado o método otsu para otimizar o limiar e então separando entre duas classes de pixels para finalmente for convertido em uma imagem binária.

O método *Graythresh* foi utilizado no dataset para segmentação de imagens e binarização da mesma, podendo facilitar a extração de características das imagens e identificação de áreas de interesse.

C. Modelos

Os modelos utilizados conforme a Fig. 3 o ResNet50, a MobileNetV2 e o AlexNet para a classificação dos estresses

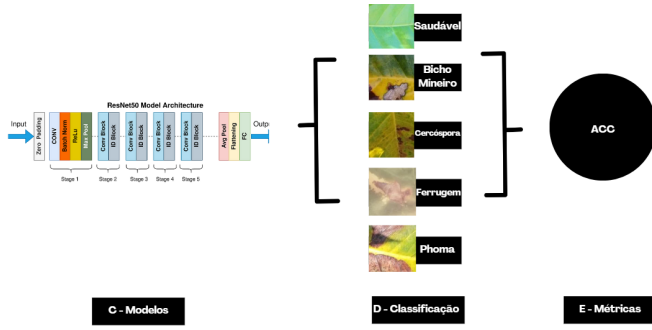


Fig. 3: Classificação das imagens

bióticos que afetam as folhas de café, aplicados cada um com todos os datasets de diferentes filtros de imagens.

O ResNet50 foi aplicado modificando a camada final para a classificação multiclasse, com o optimizer Adam, criterion CrossEntropyLoss e o learning rate 0.001, com alguns experimentos sendo pré-treinada e outros não, com 100 épocas treinadas. Obtém números ótimos nesse tipo de tarefa [14].

MobileNetV2, evolução direta da MobileNet, foi projetado para ser uma rede convolucional leve e eficiente, com a função de ativação ReLu6. Foi modificada uma das camadas para a classificação multiclasse e optimizer Adam, criterion CrossEntropy com um learning rate de 0.001, pré-treinado com seus experimentos rodados em 100 épocas treinadas, obtendo bons resultados [18].

O modelo AlexNet foi utilizado modificando uma de suas camadas para a classificação multiclasse com optimizer adam e criterion CrossEntropy, com um learning rate de 0.001 e o experimento utilizando 100 épocas, sendo um modelo confiável em diversas tarefas [11].

D. Classificação

A imagem é classificada em cinco diferentes classes, sendo elas saudável, bicho-mineiro, cercóspora, ferrugem e doença de phoma.

E. Métricas utilizadas

São aplicadas métricas nos resultados para poder ter uma dimensão melhor da performance de cada modelo em cada filtro de imagem, podendo avaliar e ajudar a interpretar os resultados, facilitando a comparação entre os diferentes métodos.

Esses componentes a seguir são utilizados para o cálculo de diferentes métricas

1) **Acurácia (ACC):** Proporção de previsões corretas feitas por um modelo em relação ao total de previsões. Calcula-se como o número de previsões corretas dividido pelo total de previsões.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

2) **Precisão (PRE):** Proporção de previsões positivas corretas feitas pelo modelo em relação ao total de previsões positivas realizadas. Usada para medir a exatidão das previsões positivas.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3) **Recall (REC):** Proporção de verdadeiras positivas que o modelo conseguiu identificar em relação ao total de casos positivos reais. Consegue medir a capacidade do modelo de encontrar todas as instâncias positivas.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

V. Resultados Obtidos

Foi realizado um teste de 100 épocas utilizando os modelos ResNet50, MobileNetV2 e AlexNet com os resultados apresentados na tabela 1. Os modelos ResNet50, MobileNetV2 e AlexNet obtiveram resultados de 98,01%, 98,01% e 98,25% respectivamente. Em geral o modelo MobileNetV2 obteve melhores resultados que as outras redes do experimento. A rede AlexNet obteve bons resultados porém não conseguiu um número tão alto utilizando o filtro da wavelet conforme a tabela 1.

TABELA I: Rede neural e filtros aplicados

Arquitetura	Filtro Aplicado	ACC (%)	PRE (%)	REC (%)
ResNet50	Normal	97,87%	98,00%	97,00%
	Clahe	97,54%	97,00%	97,00%
	Gaussian	98,01%	98,00%	98,00%
	Graythresh	96,68%	97,00%	96,00%
	Wavelet	93,01%	93,00%	93,00%
MobileNetV2	Normal	97,93%	98,00%	98,00%
	Clahe	97,41%	97,00%	97,00%
	Gaussian	98,01%	98,00%	98,00%
	Graythresh	96,68%	97,00%	97,00%
	Wavelet	94,64%	95,00%	95,00%
AlexNet	Normal	97,56%	98,00%	98,00%
	Clahe	97,43%	98,00%	98,00%
	Gaussian	98,25%	98,00%	98,00%
	Graythresh	96,29%	96,00%	96,00%
	Wavelet	77,43%	78,00%	77,00%

A aplicação dos filtros de imagens apresentaram resultados promissores. Nas médias dos experimentos obtemos números maiores no filtro Gaussian porém com bons números nos filtros Clahe e a segmentação Graythresh como mostrado na tabela 1.

Na tabela 2 é mostrado a média do loss de cada filtro. É apresentada a média mais alta no filtro Wavelet devido aos resultados inferiores a outras arquiteturas e a rede AlexNet obtendo números menores comparados com as outras. Os melhores resultados de loss condizem com a melhor acurácia

TABELA II: Loss média com cada filtro aplicado em todas as redes neurais

Filtros de imagens	Média da acc	Média da loss
Normal	97,78%	0,1082
Clahe	97,46%	0,1196
Gaussian	98,09%	0,1066
Graythresh	96,55%	0,158
Wavelet	88,36%	0,3877

do filtro de imagem, com o filtro Gaussian apresentando melhores resultados.

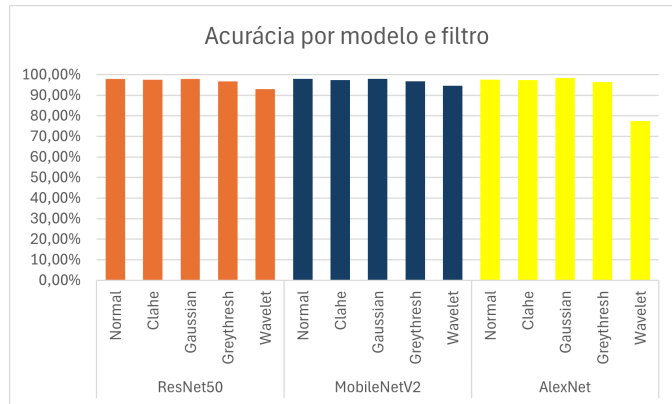


Fig. 4: Gráfico da acurácia

A Fig. 4 mostra um gráfico sobre o modelo, a acurácia obtida e o filtro de imagens utilizadas. Nela nota-se que o valor do filtro wavelet obteve menores resultados enquanto os outros filtros de imagens tiveram resultados similares.

VI. Discussão dos resultados

As redes neurais convolucionais são um método confiável para a classificação de estresses bióticos das folhas do café. Elas conseguem auxiliar de modo rápido e eficiente uma grande quantidade de imagens que não seria possível classificar em tão pouco tempo utilizando outros meios. Esse trabalho avaliou o desempenho das arquiteturas ResNet50, AlexNet e MobileNetV2 nessa tarefa, utilizando os bancos de dados JMunBEN e JMunBEN2 [19]. O filtro Gaussian obteve o melhor desempenho nessa tarefa, com o menor loss e a maior acurácia comparados com os outros filtros em geral, enquanto a imagem com técnica de Wavelet foi obtido a menor acurácia mas mesmo assim foi obtido um desempenho ótimo nas redes ResNet50 e MobileNetV2. A média de todos os filtros foram 98,09% utilizando o filtro Gaussian, 97,78% com a base de dados normal, 97,46% com o filtro Clahe, 96,55% com o método de segmentação Graythresh e finalmente 88,36% com o filtro Wavelet que acabou abaixando muito por causa de seu desempenho na rede AlexNet.

Os resultados obtidos da precisão e do *recall* são similares com os resultados da acurácia, com o filtro gaussian apresentando o melhor desempenho entre os outros filtros, como

apresentado na tabela 2. A wavelet acabou apresentando o pior resultado tanto na precisão quanto no *recall* devido ao maior dificuldade dos modelos de classificarem as imagens com esse filtro aplicado.

Na tabela 3 é comparado os resultados do trabalho [17] e o trabalho [9] com o modelo do artigo aplicado com os filtros. Foi apresentado melhores acurácias das arquiteturas ResNet50, MobileNetV2 e AlexNet, com o melhor resultado da acurácia do trabalho [17] da arquitetura ResNet50 sendo 95,63%, comparando ao trabalho proposto com a base normal e com os filtros Clahe e Gaussian obtendo respectivamente 97,87%, 97,54% e 98,01%. A arquitetura MobileNetV2 do trabalho [17] obteve 94,84% comparando os resultados com o trabalho proposto sendo 97,93% na base normal, 97,41% no filtro Clahe e 98,01% no filtro Gaussian. O modelo AlexNet proposto pelo trabalho [17] chegou a 92,46% enquanto o trabalho proposto chegou a 97,56% na base normal, 97,43% no filtro Clahe e 98,25% no filtro Gaussian. O trabalho [9] obteve 98,00% na sensibilidade e o trabalho proposto obteve 98,25% de acurácia, 98,00% de precisão e 98,00% de recall no modelo AlexNet.

VII. Conclusão

Concluimos que os filtros aplicados no banco de dados apresentam uma acurácia bem alta e podem ser confiáveis na tarefa de classificação dos estresses bióticos de ferrugem, bicho-mineiro, mancha-de-phoma e cercóspora para auxiliar fazendeiros no plantio do café arábica. Obtemos resultados melhores que trabalhos passados com o uso de filtros de imagens, mostrando a diferença de métricas utilizando-os.

Em futuros trabalhos será continuado e expandido as técnicas de imagens, testando em mais épocas e condições diferentes com outras redes neurais a fins de testar outros filtros de imagens e métodos de segmentação. Uma limitação desse modelo é a classificação de somente um estresse biótico por folha, porém é possível haver mais de uma instância em uma só folha.

TABELA III: Comparação das métricas obtidas pelos três melhores filtros propostos com metodologia do estado da arte.

Arquitetura	Filtro Aplicado	Meu modelo			[17], 2020			[9], 2019
		ACC (%)	PRE (%)	REC (%)	ACC (%)	PRE (%)	REC (%)	Sensibilidade (%)
ResNet50	Normal	97,87%	98,00%	97,00%	95,63%	94,12%	92,70%	-
	Clahe	97,54%	97,00%	97,00%				
	Gaussian	98,01%	98,00%	98,00%				
MobileNetV2	Normal	97,93%	98,00%	98,00%	94,84%	93,24%	91,79%	-
	Clahe	97,41%	97,00%	97,00%				
	Gaussian	98,01%	98,00%	98,00%				
AlexNet	Normal	97,56%	98,00%	98,00%	92,46%	89,37%	90,02%	98,00%
	Clahe	97,43%	98,00%	98,00%				
	Gaussian	98,25%	98,00%	98,00%				

VIII. ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. Daniel Imai Yamakawa received a scholarship from PIBIC/CNPq

REFERENCES

- [1] god politica agricola. (00) Política agrícola. [Online]. Available: <https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/cafe/cafeicultura-brasileira#:~:text=Atualmente%2C%20o%20caf%C3%A9%20%C3%A9%20relevante,ao%20produtor%20e%20sua%20atividade.>
- [2] F. Rodrigues, F. Patrício, E. Oliveira, and A. Paula, *Desafios do manejo no controle de doenças do café - F I TOSSA N I DA D E*, 03 2019, pp. 52 – 54.
- [3] R. Sharma, “Artificial intelligence in agriculture: A review,” in *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2021, pp. 937–942.
- [4] S. Mishra, R. Sachan, and D. Rajpal, “Deep convolutional neural network based detection system for real-time corn plant disease recognition,” *Procedia Computer Science*, vol. 167, pp. 2003–2010, 2020, international Conference on Computational Intelligence and Data Science. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S187705092030702X>
- [5] N. Bevers, E. J. Sikora, and N. B. Hardy, “Soybean disease identification using original field images and transfer learning with convolutional neural networks,” *Computers and Electronics in Agriculture*, vol. 203, p. 107449, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169922007578>
- [6] D. Tedesco-Oliveira, R. Pereira da Silva, W. Maldonado, and C. Zerbato, “Convolutional neural networks in predicting cotton yield from images of commercial fields,” *Computers and Electronics in Agriculture*, vol. 171, p. 105307, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169919319878>
- [7] J. Kranz, *Measuring Plant Disease*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1988, pp. 35–50. [Online]. Available: https://doi.org/10.1007/978-3-642-95534-1_4
- [8] J. G. Esgario, R. A. Krohling, and J. A. Ventura, “Deep learning for classification and severity estimation of coffee leaf biotic stress,” *Computers and Electronics in Agriculture*, vol. 169, p. 105162, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169919313225>
- [9] J. G. Esgario, P. B. de Castro, L. M. Tassis, and R. A. Krohling, “An app to assist farmers in the identification of diseases and pests of coffee leaves using deep learning,” *Information Processing in Agriculture*, vol. 9, no. 1, pp. 38–47, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2214317321000044>
- [10] S. A. A. I. K. M. Arif Wani, Farooq Ahmad Bhat, *Advances in Deep Learning*. Singapore: Springer Singapore, 2019.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
- [12] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015.
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [15] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.
- [16] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid scene parsing network,” 2017.
- [17] L. X. Boa Sorte, C. T. Ferraz, F. Fambrini, R. dos Reis Goulart, and J. H. Saito, “Coffee leaf disease recognition based on deep learning and texture attributes,” *Procedia Computer Science*, vol. 159, pp. 135–144, 2019, knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 23rd International Conference KES2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050919313468>
- [18] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” 2019.
- [19] J. Jepkoech, D. M. Mugo, B. K. Kenduiyo, and E. C. Too, “Arabica coffee leaf images dataset for coffee leaf disease detection and classification,” *Data in Brief*, vol. 36, p. 107142, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352340921004261>
- [20] J. P. D’Haeyer, “Gaussian filtering of images: A regularization approach,” *Signal Processing*, vol. 18, no. 2, pp. 169–181, 1989. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0165168489900480>
- [21] S. Pizer, R. Johnston, J. Erickson, B. Yankaskas, and K. Muller, “Contrast-limited adaptive histogram equalization: speed and effectiveness,” in *[1990] Proceedings of the First Conference on Visualization in Biomedical Computing*, 1990, pp. 337–345.
- [22] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.