

Comparação de desempenho entre Random Forest e MobileNetV2 para classificação de imagens de folhas de café

Emanuel Aurelio Ferreira de Miranda¹, Romuere Rodrigues Veloso e Silva²,
Humberto José da Silva Júnior¹, Vitor José Ferreira dos Santos de Santana¹

¹PAAD - Núcleo Pesquisas Avançadas em Análise de Dados
Universidade Federal do Piauí (UFPI) - Picos – PI – Brazil

²PAVIC - Pesquisas Aplicadas à Visão e Inteligência Computacional
Universidade Federal do Piauí (UFPI) - Picos – PI – Brazil

{emanuelmiranda.si, romuere, lbertojr, vitorsantos}@ufpi.edu.br

Abstract. *This article presents a study on the classification of coffee leaf images into healthy and unhealthy categories using the RoCoLe image dataset and two different techniques: Random Forest and MobileNetV2. Image features were extracted for training a Random Forest classifier. Subsequently, a comparison with the MobileNetV2 model was conducted, revealing that the latter achieved superior accuracy. These results highlight the promising application of machine learning techniques in outdoor environments and underscore the importance of choosing the most suitable technique based on the dataset and specific application goals.*

Resumo. *Este artigo apresenta um estudo de classificação de imagens de folhas de café em saudáveis e não saudáveis, empregando o conjunto de imagens RoCoLe e duas técnicas distintas: Random Forest e MobileNetV2. As características das imagens foram extraídas para o treinamento de um classificador Random Forest. Em seguida, uma comparação com o modelo MobileNetV2 foi realizada, revelando que este último obteve uma acurácia superior. Esses resultados ressaltam a promissora aplicação de técnicas de aprendizado de máquina em ambientes externos e evidenciam a importância de escolher a técnica mais adequada com base no conjunto de dados e nos objetivos específicos da aplicação.*

1. Introdução

A cafeicultura brasileira, uma das principais atividades do setor agropecuário do país, assume papel fundamental na geração de empregos, recursos e divisas para a economia nacional. Além disso, ela é bastante diversificada, apresentando particularidades regionais que enriquecem ainda mais o seu valor socioeconômico [Fernandes et al. 2012]. A cafeicultura está vivenciando uma fase de grande evolução, resultando em um aumento sistemático da produtividade. Essa evolução está diretamente relacionada à tecnologia, que tem alterado a importância e o manejo das doenças do cafeeiro, como a ferrugem, indicada por ácaros vermelhos e manchas [Patricio and Oliveira 2013].

No âmbito agrícola, identificar problemas em plantações pode representar uma tarefa desafiadora. Diversos fatores estressantes podem manifestar sintomas semelhantes

em diferentes espécies de plantas, o que torna a precisão no diagnóstico uma tarefa complexa. Especificamente, em culturas como o café, a presença de ferrugem nas folhas é uma das principais responsáveis por danos significativos e perdas na produção. Portanto, é fundamental que os produtores tenham à disposição ferramentas capazes de detectar eficientemente a presença de ferrugem e outros problemas, permitindo ações rápidas para minimizar prejuízos [Fenu and Mallocci 2021].

O RoCoLe é um banco de imagens de folhas de café da espécie robusta e desempenha um papel fundamental no treinamento e validação de algoritmos de aprendizado de máquina [Parraga-Alava et al. 2019]. O supramencionado conjunto de imagens abrange informações sobre as condições das folhas de café, desde aquelas em estado saudável até aquelas afetadas por doenças, notadamente a ferrugem, além de fornecer detalhes sobre a gravidade dessas condições. O RoCoLe reúne um total de 1560 imagens, sendo escolhido por ser disponibilizado gratuitamente pelos criadores, tornando-se uma fonte inestimável para estudos visando a compreensão das doenças que afetam a cultura cafeeira.

O conjunto de imagens apresenta desafios significativos que impactam o processo de reconhecimento devido a várias questões, incluindo a ausência de fronteiras claras em torno dos sintomas, a natureza variável dos sintomas de diferentes doenças, a possibilidade de múltiplas doenças com sintomas simultâneos e semelhança de sintomas entre diferentes doenças, sendo a principal a variações nas condições de captura [Dias 2021]. As fotos foram capturas em campo, onde as condições ambientais variam significativamente e não podem ser controladas de maneira precisa. Tal falta de controle sobre as circunstâncias de captura das imagens representa um desafio adicional para o desempenho dos algoritmos de aprendizado e previsão [Fenu and Mallocci 2021].

Este estudo tem como objetivo investigar e comparar o desempenho de duas abordagens realizados com o banco de imagens RoCoLe. No Experimento 1, foi aplicado o classificador *Random Forest* em conjunto com o descritor *LBP Local Binary Pattern*. Enquanto isso, no Experimento 2, a aplicação de um Rede Neural Convolutiva (CNN - *Convolutional Neural Networks*) com a arquitetura *MobileNetV2*. Os experimentos visam a classificação de imagens em duas categorias (Figura 1): saudáveis e não saudáveis, visando o aprimoramento de métodos eficazes de diagnóstico e controle de doenças em plantações e, assim, contribuir com esse setor crucial da economia brasileira.



Figura 1. Exemplos de imagens de folhas de café saudáveis (a) e não saudáveis (b).

A organização deste artigo é definida através das seguintes seções: Seção 2, é apre-

sentado uma análise detalhada dos trabalhos relacionados ao nosso estudo, destacando sua relevância no contexto atual; Seção 3, é descrito a metodologia que foi aplicada na execução do projeto, delineando os passos e abordagens utilizados; Seção 4, são apresentados os resultados obtidos a partir da aplicação da metodologia, oferecendo uma visão clara das resultados alcançadas; e Seção 5, contém as considerações finais e sugestões para trabalhos futuros.

2. Trabalhos Relacionados

Nesta seção, é apresentada uma seleção de trabalhos relevantes no campo do processamento de imagens para a classificação de folhas de café, com foco específico no *dataset* RoCole. É destacado abordagens e contribuições significativas desses trabalhos visando fornecer uma visão abrangente do contexto em que este trabalho se insere. Para embasar esta pesquisa, foram consultados de bases científicas, incluindo o *Google Scholar* e outras, a fim de garantir uma revisão abrangente da literatura disponível. Ademais, é apresentado um quadro comparativo que analise aspectos em relação a esse estudo, ressaltando as principais diferenças e semelhanças entre as abordagens propostas.

A Tabela 1 oferece uma visão geral dos trabalhos relacionados a este artigo, classificando-os com base nos critérios principais: I. Segmentação: Esta categoria avalia se os trabalhos utilizaram técnicas de segmentação de imagens para a classificação das folhas de café. II. Método Utilizado: Aqui, descrevemos qual método ou abordagem foi adotado pelos estudos para a classificação das folhas. III. Comparação entre Métodos: Esta categoria verifica se os trabalhos realizaram comparações entre diferentes métodos de classificação, especialmente quando mais de duas abordagens foram consideradas.

Tabela 1. Quadro comparativo de trabalhos relacionados

Trabalho	I	II	III
[Dias 2021]	JSEG	-	Não
[Tassis et al. 2021]	Mask R-CNN	ResNet50	Não
[Franco 2022]	-	MobileNetV2	Não
[Chowdhury 2021]	-	GIST e SVM	Não
Este artigo	Não	RandomForest e MobileNetV2	Sim

I — Segmentação; II — Método utilizado; III — Comparação entre métodos;

O trabalho de [Dias 2021], trata da segmentação das imagens através do algoritmo JSEG, resultando em 4 escalas de segmentação, para 50 imagens 380x640 do banco. O algoritmo JSEG é um método de segmentação de imagem baseado em regiões que utilizam a análise de textura e a transformada wavelet discreta por decompor imagens em wavelets com diferentes escalas e frequências, permitindo segmentar as imagens em regiões de interesse. Para analisar o resultado da segmentação os autores realizaram o agrupamento dos segmentos por meio da distância Euclidiana entre os histogramas normalizados. No entanto, os autores reconhecem que há limitações na abordagem proposta, como a sensibilidade a variações na iluminação e na textura das imagens.

No trabalho de [Tassis et al. 2021] é proposto um framework, uma abstração que permite o desenvolvimento de aplicações sem precisar se preocupar com a operação de

baixo nível, integrado utilizando diferentes CNN para reconhecer as folhas de café não saudáveis. A segmentação de instâncias, com o Mask R-CNN, é uma técnica que identifica e rotula objetos individuais em uma imagem, enquanto a segmentação semântica, os autores utilizaram UNet e a PSPNet que agrupa os pixels da imagem em regiões semânticas com características visuais semelhantes. A classificação foi realizada pelo ResNet50. Os resultados mostraram que a abordagem proposta foi satisfatória, superando outras técnicas de aprendizado de máquina, como *Random Forest* e SVM.

O trabalho de [Franco 2022] utilizou uma arquitetura baseada no *MobileNetV2*, amplamente conhecida no domínio das redes neurais convolucionais. A arquitetura foi ajustada para a tarefa de classificação de folhas de café do conjunt de imagens *RoCoLe*. Os resultados obtidos revelaram que a abordagem uma acurácia de classificação de 82%, sendo capaz de distinguir folhas saudáveis das infectadas por doenças com alta precisão. Os autores ressaltam que a técnica proposta apresenta potencial para aplicação em outras culturas agrícolas, contribuindo para a identificação de doenças e aprimoramento da produção.

O artigo de [Chowdhury 2021] apresenta uma abordagem utilizando características globais de e 640 imagens do banco de dados (*RoCoLe*). As características foram obtidas por meio da técnica GIST (*Global Image Structure*), que se concentra na extração de informações globais das imagens, incluindo distribuição de texturas, bordas e orientações, sem considerar informações locais. O trabalho utiliza o classificador SVM (*Support Vector Machine*) para a tarefa de classificação, e os resultados impressionam, alcançando uma notável acurácia de 99,4% na identificação de folhas não saudáveis

3. Metodologia

A metodologia adotada neste estudo para a identificação e classificação das imagens, o banco de dados de folhas de café foi dividido em duas categorias: folhas saudáveis e folhas não saudáveis. O estudo não foi empregado qualquer método ou técnica de segmentação nas imagens e foi estruturado em dois experimentos distintos, evidenciando a versatilidade das técnicas utilizadas: Experimento 1 e Experimento 2.

No Experimento 1, foi utilizado o algoritmo *Random Forest* em conjunto com o descritor LBP (*Local Binary Pattern*) para extrair características das imagens e classificá-las em folhas de café saudáveis e não saudáveis. O descritor LBP é um método de textura que se concentra na extração de características locais da imagem, representando-as por meio de um histograma

No Experimento 2, foi utilizada abordagem baseada em uma rede neural convolucional, utilizando o modelo *MobileNetV2* para classificação das imagens. As redes neurais convolucionais são reconhecidas por sua eficácia na identificação e classificação de imagens. O modelo escolhido nesse experimento se destaca por sua arquitetura eficiente e alta precisão, sendo especialmente adequado para aplicações em dispositivos com recursos limitados, como dispositivos móveis.

Na Figura 3, é apresentado o fluxograma da metodologia proposta, que consiste em quatro etapas principais: pré-processamento das imagens, extração de características e classificação. No processo de pré-processamento das imagens, estas foram redimensionadas para uma dimensão padrão e normalizadas, enquanto a extração de características

envolveu o descritor LBP (Experimento 1) e a rede neural *MobileNetV2* (Experimento 2). A etapa de classificação, por sua vez, fez uso dos algoritmos *Random Forest* (Experimento 1) e rede neural convolucional (Experimento 2) para classificar as imagens em folhas saudáveis e não saudáveis.

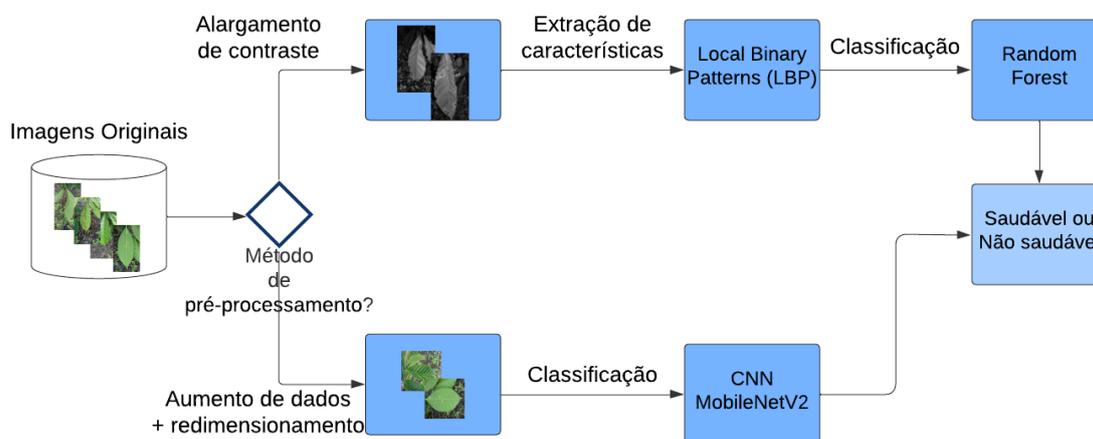


Figura 2. Fluxograma da metodologia.

Na etapa de **pré-processamento**, foram adotadas técnicas específicas para cada um dos experimentos realizados. No Experimento 1, foi aplicado o alargamento de contraste, com o valor de K e E igual a 2,25, para melhorar a qualidade e o contraste das imagens antes da extração de características. Por sua vez, no Experimento 2, optou-se pelo redimensionamento das imagens para uma dimensão de 256x256 pixels, a fim de padronizar as dimensões das imagens para a fase de treinamento do modelo.

Na etapa de **extração de características**, foi adotado o *Local Binary Patterns* (LBP) como descritor de textura para extrair informações das imagens. O LBP baseia-se em uma comparação local de intensidades de pixel em uma janela específica, onde cada pixel é comparado com seus vizinhos. Essa comparação gera uma sequência binária que representa as diferenças de intensidade entre os pixels. Já no caso do Experimento 2, a extração de características é realizada por meio de uma *Convolutional Neural Network* (CNN), que aprende padrões e características relevantes das imagens de forma automática. Essas características são posteriormente utilizadas para treinar o modelo de classificação, permitindo uma identificação mais precisa e confiável das folhas de café saudáveis e não saudáveis.

Na etapa de **classificação**, o Experimento 1 utilizou o *Random Forest*, um algoritmo de aprendizado de máquina baseado em árvores de decisão que é conhecido por sua eficiência em termos de recursos computacionais, menos intensivo em poder de processamento. Por outro lado, no Experimento 2, a classificação foi realizada com o uso de uma rede neural convolucional, mais especificamente o modelo *MobileNetV2*, sendo um modelo de rede neural convolucional leve e eficiente, visto que foi projetado para dispositivos móveis e ambientes com recursos limitados, empregando blocos de convolução e técnicas de regularização para reduzir a quantidade de cálculos necessários, garantindo inferências rápidas e eficientes.

A escolha de aplicar dois métodos diferentes ressalta nossa abordagem diversifi-

cada na busca da melhor solução para a tarefa de classificação de folhas de café saudáveis e não saudáveis, permitindo uma avaliação comparativa do desempenho e das características de cada método. Enquanto o Experimento 1 é valorizado por sua eficiência e facilidade de interpretação, o Experimento 2 sobressai-se ao demonstrar a capacidade de identificar automaticamente padrões complexos nas imagens. A utilização de ambos os métodos enriquece nossa análise e compreensão das soluções propostas neste trabalho, abordando diferentes aspectos da classificação de imagens de folhas de café.

4. Resultados e discussão

De acordo com o desenvolvimento dos experimentos propostos na metodologia e nos resultados obtidos, a Tabela 2 apresenta um resumo das métricas de desempenho alcançadas por cada abordagem. A abordagem com a rede neural convolucional, no Experimento 2, alcançou uma acurácia de 91%, demonstrando sua eficácia na identificação e classificação de imagens de folhas de café saudáveis e não saudáveis. Este resultado sugere que a CNN é uma escolha promissora para tarefas de classificação de imagens, superando o desempenho do Experimento 1, que obteve acurácia de 68%.

Tabela 2. Quadro comparativo dos resultados dos experimentos

Experimento	I	II
Método	RandomForest	MobileNetV2
Acurácia	68%	91%

A tabela 3 possibilita a análise e comparação do melhor resultado obtido neste estudo com as métricas dos trabalhos relacionados previamente levantados. O desempenho significativo alcançado neste artigo reforça a eficácia da abordagem adotada, que combina o uso da rede neural convolucional *MobileNetV2* com a aplicação do alargamento de contraste no pré-processamento das imagens, contribuindo para uma classificação precisa e robusta de folhas de café saudáveis e não saudáveis.

Tabela 3. Quadro comparativo de resultados.

Trabalho	Acurácia
[Dias 2021]	-
[Tassis et al. 2021]	94.25%
[Franco 2022]	82%
[Chowdhury 2021]	99,8%
Este artigo	91%

O gráfico abaixo mostra duas das principais funções usadas para avaliar o desempenho da rede MobileNetV2, que são a acurácia e a função de perda. A função de acurácia mede a proporção de imagens classificadas corretamente pela rede em relação ao total de imagens avaliadas. A função de perda é usada para calcular a diferença entre a saída esperada da rede (rótulo original da imagem) e sua saída real (classificação da imagem feita pela rede). O objetivo é minimizar a perda durante o treinamento da rede, ajustando os

pesos dos neurônios, para que a acurácia possa ser maximizada. Em resumo, a função de acurácia avalia a precisão da rede em classificar as imagens corretamente, enquanto a função de perda ajuda a ajustar a rede para melhorar seu desempenho.

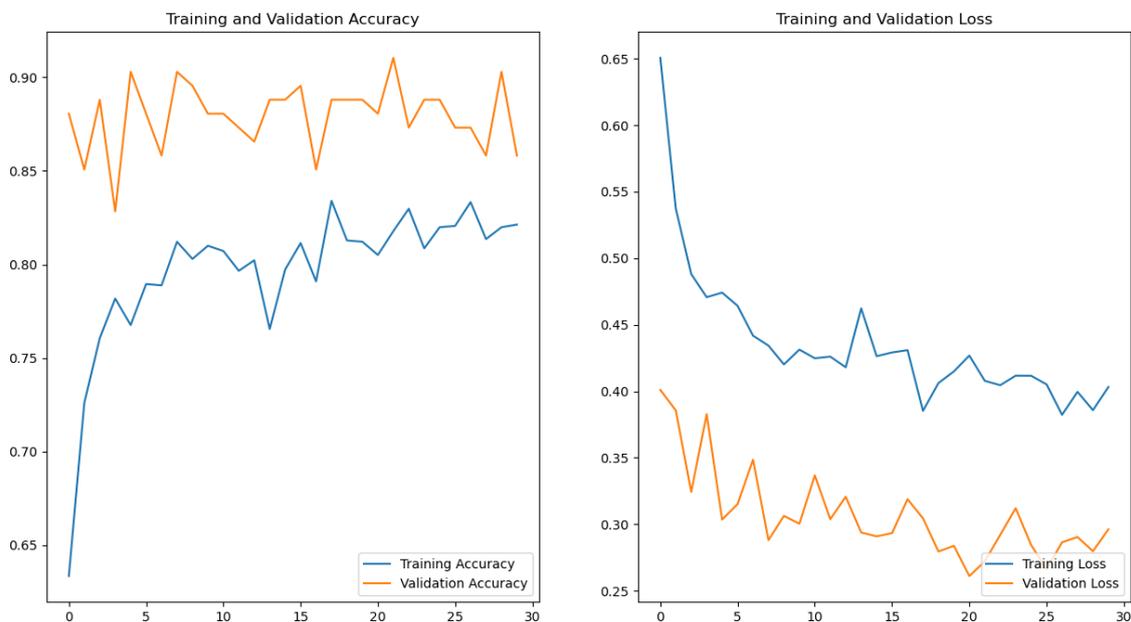


Figura 3. Gráfico de acurácia e função de perda.

A quantidade de épocas escolhida para o treinamento da rede pode influenciar significativamente o seu desempenho. Uma época é definida como um ciclo completo de treinamento da rede usando todo o conjunto de dados de treinamento disponível. Normalmente, quanto maior o número de épocas, maior é a chance da rede se ajustar ao conjunto de treinamento, mas também há um risco de overfitting, onde a rede se torna muito especializada nos dados de treinamento e não generaliza bem para dados desconhecidos. Nos testes realizados neste trabalho, foi possível notar que o treinamento da rede mostra bons resultados entre 22 e 26 épocas, portanto foi definido a quantidade de épocas em 30 e após o treinamento foi realizada uma análise de desempenho do modelo a fim de avaliar a predição para imagens que não foram usadas no treinamento.

5. Conclusão

Neste artigo, foi desenvolvido uma metodologia que possibilitou a comparação de dois experimentos distintos com o objetivo de encontrar a abordagem mais adequada para a identificação de doenças nas folhas de café. Como supramencionado anteriormente, o desafio de reconhecimento é complexo e desafiador devido às variações nas imagens capturadas no banco de dados RoCoLe. Os resultados obtidos revelaram que o Experimento 2, que empregou uma rede neural convolucional, superou o Experimento 1, alcançando uma acurácia de 91%.

Os resultados deste estudo confirmam a eficácia da metodologia proposta, demonstrando um desempenho satisfatório na identificação de doenças nas folhas de café. Essa validação abre caminho para possíveis aplicações em outras áreas, como a detecção de níveis de doenças e categorias adicionais em culturas agrícolas. Este estudo também

destaca a importância da detecção precoce de doenças foliares nas plantas de café, oferecendo um auxílio valioso para os produtores na preservação e tratamento das culturas.

Como perspectivas de estudos futuros, propõe-se o aprimoramento da metodologia realização de novos experimentos em um conjunto de imagens mais amplo e diversificado, com o objetivo de aprimorar o desempenho do modelo, além de explorar outras técnicas de aprendizado de máquina, como o uso de redes neurais mais complexas e o emprego de recursos de pré-processamento e pós-processamento avançados, a fim de identificar níveis de doenças com maior precisão e categorizar efetivamente as condições das folhas de café.

Referências

- Chowdhury, Md e Burhan, U. (2021). Coffee leaf disease recognition using gist feature. *International Journal of Information Engineering and Electronic Business*, 13(2):55–61.
- Dias, Jeferson de Sousa e Saito, J. H. (2021). Coffee plant image segmentation and disease detection using jseg algorithm. In *Anais do XVII Workshop de Visão Computacional*, pages 42–47. SBC.
- Fenu, G. and Mallocci, F. M. (2021). Diamos plant: A dataset for diagnosis and monitoring plant disease. *Agronomy*, 11(11):2107.
- Fernandes, A. L. T., Partelli, F. L., Bonomo, R., and Golynski, A. (2012). A moderna cafeicultura dos cerrados brasileiros. *Pesquisa Agropecuária Tropical*, 42:231–240.
- Franco, Matheus Eloy e Silva, J. E. L. (2022). Redes neurais convolucionais aplicada à identificação de folhas saudáveis de café. *14º JORNADA CIENTÍFICA E TECNOLÓGICA E 11º SIMPÓSIO DE PÓS-GRADUAÇÃO DO IFSULDEMINAS*, 14(1).
- Parraga-Alava, J., Cusme, K., Loor, A., and Santander, E. (2019). Rocol: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. *Data in brief*, 25:104414.
- Patricio, F. R. A. and Oliveira, E. (2013). Desafios do manejo no controle de doenças do café. *Visão Agrícola, Piracicaba*, 12:51–54.
- Tassis, L. M. e. d. S., Tozzi, J. E., and Krohling, R. A. (2021). A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images. *Computers and Electronics in Agriculture*, 186:106191.