

INTERPRETABILIDADE NA IDENTIFICAÇÃO DE DOENÇAS EM PLANTAS DE CAFÉ POR MEIO DE IMAGENS E APRENDIZADO PROFUNDO

MARIA EDUARDA MENEGHETTI MOREIRA MARTINS¹, JULIANA APARECIDA FRACAROLLI²

¹ Eng. Agrícola, Mestranda em Engenharia Agrícola, Feagri / Unicamp, Campinas - SP, m097282@dac.unicamp.br.

² Eng. Agrícola, Profª. Doutora, Feagri / Unicamp, Campinas – SP, juliana.fracarolli@unicamp.br.

Apresentado no
LII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2023
18 a 21 de outubro de 2023 – Ribeirão Preto - SP, Brasil

RESUMO: Importante para garantir a segurança alimentar e a sustentabilidade das culturas agrícolas, a detecção precoce de doenças em plantas tem sido tópico de pesquisas com aplicação de aprendizado profundo. A intenção do aprendizado de máquina é de auxiliar e até mesmo substituir a tarefa de detecção feita de forma visual, que demanda muito esforço e em muitos casos pode apresentar dificuldades na detecção das doenças e identificação da sua causa. Contudo, a busca por altas acurácias gera modelos cada vez mais complexos que nem sempre promovem o entendimento humano de suas tomadas de decisão. Para isso, técnicas de interpretabilidade são adicionadas com o objetivo de trazer informações que viabilizem a explicação dos resultados. O projeto visa aplicar duas técnicas em uma rede neural convolucional (CNN) para a identificação e classificação de imagens de doenças em plantas. Ao utilizar o SHAP (*SHapley Additive exPlanations*) e LIME (*Local Interpretable Model-Agnostic Explanations*) na modelagem de doenças na cultura do café espera-se proporcionar a visualização dos principais aspectos, características e variáveis que contribuem para a classificação de doenças da cultura.

PALAVRAS-CHAVE: Doenças em plantas, LIME, SHAP

INTERPRETABILITY IN THE IDENTIFICATION OF DISEASES IN COFFEE PLANT THROUGH IMAGES AND DEEP LEARNING

ABSTRACT: Important to ensure food security and sustainability of agricultural crops, the early detection of plant diseases has been a trending topic of research especially considering the use of deep learning. The purpose of using Machine Learning is to help and even replace the detection task done visually, which demands a lot of effort and in many cases can present bottlenecks in the detection of the disease and identification of its cause. However, the search for high accuracy increasingly generates complex models that do not always promote human understanding of their decision-making. For this, interpretability techniques are added with the aim of bringing information that makes it possible to explain the results. This project adds two techniques in a convolutional neural network (CNN) for the identification and classification of disease images in plants. By applying SHAP (*SHapley Additive exPlanations*) and LIME (*Local Interpretable Model-Agnostic Explanations*) in the machine learning classification it is expected to provide visual information of the main aspects, characteristics and variables that contribute to the classification of diseases in the coffee crop.

KEYWORDS: Plant diseases, LIME, SHAP

INTRODUÇÃO: O controle de doenças de plantas em culturas agrícolas é um grande desafio para garantir a segurança alimentar global e uma agricultura sustentável. A detecção precoce dos sintomas da doença é um dos principais desafios na proteção de safras e na limitação de epidemias (LEE et al., 2020). Essa tarefa milenar tem sido realizada principalmente por monitoramento visual, dependendo do conhecimento de especialistas na avaliação. Entretanto, atualmente, novos métodos de estudo para detecção precoce de doenças de plantas, como modelos de redes neurais convolucionais (CNN, do inglês *Convolutional Neural Network*), têm sido desenvolvidos com o objetivo de otimizar e aumentar a eficiência do processo. Hasan et al. (2020), Habaragamuwa et al. (2021) e Lee et al. (2020) destacam que o aprendizado profundo está desempenhando um papel vital nas tarefas de reconhecimento de imagens agrícolas, como a detecção precoce e classificação de doenças de plantas. Apesar do alto poder de discriminação que o aprendizado profundo proporciona, essa acurácia é obtida ao custo da baixa interpretabilidade (caixa preta) (ZHANG e ZHU, 2018). Dessa forma, entender a lógica e os recursos usados nas tomadas de decisão dos modelos é importante para garantir a verificação e validação dos resultados. A visualização de representações das redes CNN, e redes interpretáveis é uma tendência já identificada e utilizada na área da medicina (VAN DER VELDEN et al., 2020 e GRAZIANI et al., 2020), mas ainda pouco estudada na aplicação na agricultura. O trabalho tem como objetivo avançar no uso das técnicas de interpretabilidade nos modelos de detecção e classificação de doenças em plantas e contribuir na identificação de padrões e variáveis importantes para esse processo.

MATERIAL E MÉTODOS: Com o objetivo de aplicar as técnicas de interpretabilidade SHAP (LUNDBERG e LEE, 2017) e LIME (RIBEIRO et al., 2016) em um modelo desenvolvido para classificação da doença da planta da cultura do café, foi desenvolvida em Python a modelagem de uma rede neural convolucional (CNN). Para validar a utilização das técnicas de interpretação, foi desenvolvido um modelo de classificação simples a partir de uma rede já conhecida com aplicação da técnica de transferência de conhecimento (transfer learning). A rede testada e avaliada como opção eficaz na classificação de plantas foi a rede GoogleNet utilizada por Barbedo (2018) devido ao seu bom desempenho no contexto de reconhecimento de doenças de plantas. Para modelagem, foi utilizado o banco de dados aberto da Embrapa contendo imagens de folhas de café com a presença de doenças, disponível no site do metabuscador de dados de pesquisa da FAPESP (<https://metabuscador.uspdigital.usp.br>). As quatro doenças ocasionadas por fungos e consideradas no projeto foram: a ferrugem, a cercospirose, a phoma e a mantegosa. Após o desenvolvimento da rede de classificação das doenças, foram utilizadas as técnicas de interpretabilidade LIME e SHAP e seus respectivos pacotes de código para interpretação e identificação de quais os parâmetros e partes do objeto em análise são mais relevantes na classificação e detecção precoce de doenças em plantas de culturas agrícolas.

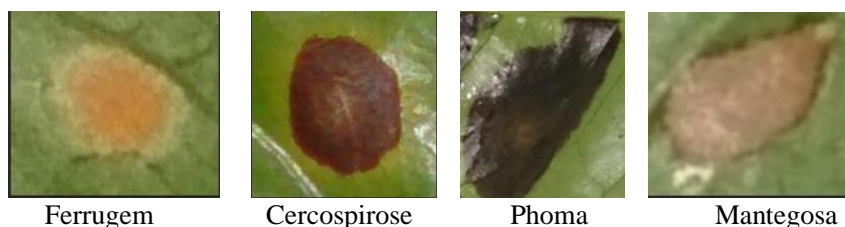


FIGURA 1. Banco de imagens. Fonte: Embrapa (<https://metabuscador.uspdigital.usp.br>)

RESULTADOS E DISCUSSÃO: Neste trabalho foi possível validar a utilização do aprendizado de máquina associada as técnicas de transferência de conhecimento na classificação de imagens para doenças de café, resultando em uma alta acurácia. O modelo desenvolvido com a rede GoogleVNet foi pré-

treinado no conjunto de dados do MobileNet V2 (que utiliza 1,4 milhões de imagens com 1000 classes) e aplicado para o conjunto de dados do presente trabalho, com um banco de imagem de 702 recortes de folhas que apresentam um dos quatro tipos de doenças analisadas. Na Figura 2, apresenta-se um exemplo de predição para a classe de doença “ferrugem” categorizada como classe “0”.

TABELA 1. Resultado de acurácia do modelo utilizado

Modelo	Acurácia	Perda
GoogleVNet	0.9086	0.2449

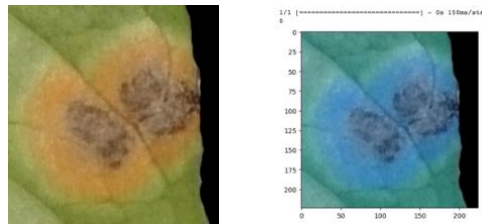


FIGURA 2. Exemplo de predição. Fonte: Produzido pela autora.

O primeiro modelo de interpretabilidade testado foi o modelo local LIME em que foi possível identificar as áreas que mais contribuíram para a tomada de decisão demarcadas em verde e as que menos contribuíram em vermelho, além de uma visão de mapa de calor com os pixels que mais contribuíram em azul até os de menor influência em vermelho escuro (Figura 3). Pode-se observar que a área que visualmente apresenta a doença de fato contribui para a tomada de decisão do modelo e regiões saudáveis da folha foram as consideradas de menor relevância pela rede neural.

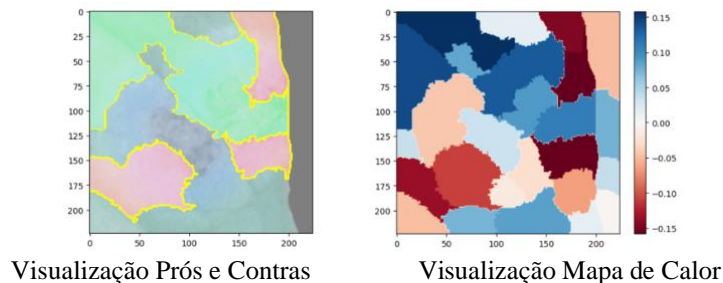


FIGURA 2. Resultados obtidos aplicando o pacote de interpretabilidade LIME. Fonte: Produzido pela autora.

O SHAP, segundo modelo de interpretabilidade avaliado, apresentou resultados menos ilustrativos, em que se encontra uma maior dificuldade na interpretação dos fatores e áreas da imagem que mais contribuíram para a decisão, como apresentado na Figura 3.

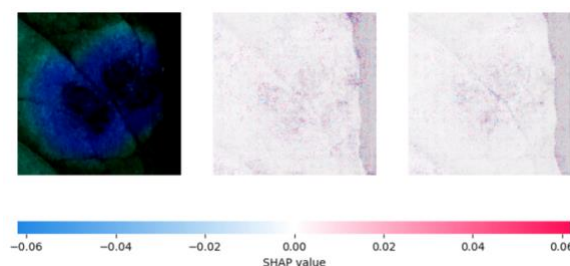


FIGURA 3. Resultados obtidos aplicando o pacote de interpretabilidade SHAP. Fonte: Produzido pela autora.

CONCLUSÕES: O projeto contribui para a área o desenvolvimento de estudos de aplicação de técnicas de interpretabilidade na tarefa de detecção de doenças em plantas. Foi possível validar a utilização da técnica de interpretabilidade LIME, o que contribui para essa área ainda pouco explorada da interpretabilidade aplicada às redes de detecção de doenças de plantas de culturas agrícolas. Para a técnica SHAP os resultados do presente trabalho não foram satisfatórios no sentido de proporcionar maior interpretabilidade ao modelo, não sendo possível inferir regiões da imagem ou características com maior importância.

REFERÊNCIAS:

BARBEDO, J. G. A. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. **Computers and Electronics in Agriculture**, v. 153, p. 45-53, out. 2018

GRAZIANI, M.; ANDREARCZYK, V.; MARCHAND-MAILLET, S.; MULLER H. Concept attribution: Explaining CNN decisions to physicians. **Computers in Biology and Medicine**, v. 123, 103865, ago. 2020.

HABARAGAMUWA, H.; OISHI Y.; TANAKA K. Achieving Explainability for Plant Disease Classification with Disentangled Variational Autoencoders. Fev. 2021.
Disponível em: <arXiv:2102.03082>
Acessado em: 22/02/2021

HASAN, R. I.; YUSUF, S. M.; ALZUBAIDI, L. Review of the State of the Art of Deep Learning for Plant Diseases: A Broad Analysis and Discussion. **Plants**, v. 9(10), 1302, out. 2020.

LEE, S. H.; GOEAU, H.; BONNET, P.; JOLY, A. New perspectives on plant disease characterization based on deep learning. **Computers and Electronics**, v. 170, 105220, mar. 2020.

LUNDBERG, S.; LEE, S. (2017). A Unified Approach to Interpreting Model Predictions. **Advances in Neural Information Processing Systems**, p. 4765–4774, nov. 2017.

RIBEIRO, M.; SINGH, S.; GUESTRIN, C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p. 1135–1144, ago. 2016.
Disponível em: <<https://arxiv.org/abs/1602.04938>>
Acessado em: 23/02/2021

VAN DER VELDEN, B.; RAGUSI, M.; JANSE, M.; LOO, C.; GILHUIJS, K. Interpretable deep learning regression for breast density estimation on MRI. 2020.
Disponível em: <<https://arxiv.org/abs/2012.04336>>
Acessado em: 20/02/2021

ZHANG, Q.; ZHU, S. Visual Interpretability for Deep Learning: a Survey. **Frontiers of Information Technology and Electronic Engineering**, v. 19, p 27 - 39, fev. 2018.