INTELIGÊNCIA ARTIFICIAL E VISÃO COMPUTACIONAL APLICADAS À DETECÇÃO DE DOENÇAS DE PLANTAS

Renan Curty da Silva, Felipe Pereira Aguiar Cunha de Almeida, João Pedro Martins Marconsine, Leonardo Peres Moreira, Lucas Jordão Santana Tigre, Simone de Paiva Caetano Bucker Moraes, Willian Bucker Moraes.

Universidade Federal do Espírito Santo/Centro de Ciências Agrárias e Engenharias, Alto Universitário, s/n, Guararema, 29500-000 – Alegre – ES, Brasil, curty.renan@gmail.com, fpalmeida1103@gmail.com, marconsine2407@gmail.com, leo.peres.moreira@hotmail.com, lucasjst01@gmail.com, simonepaivabucker@gmail.com, willian.moraes@ufes.br.

Resumo

O diagnóstico tradicional de doenças de plantas por inspeção visual é subjetivo e frequentemente tardio, o que tem impulsionado o uso de Inteligência Artificial (IA) para automatizar esse processo. Objetivou-se com este trabalho realizar uma revisão narrativa/integrativa sobre as principais abordagens de Deep Learning aplicadas à detecção de doenças de plantas, com base em artigos publicados entre 2022 e 2025, abrangendo redes neurais convolucionais (CNNs), arquiteturas de detecção em tempo real (YOLO), Vision Transformers (ViT) e os desafios de explicabilidade e generalização dos modelos. Os resultados mostram que as CNNs permanecem predominantes, com acurácias médias superiores a 90%, enquanto o YOLO se destaca na detecção em tempo real e os ViTs surgem como alternativa promissora, porém mais exigente computacionalmente. Explicabilidade e generalização a campo permanecem as principais lacunas. Conclui-se que a visão computacional é um pilar da Agricultura 4.0, cuja adoção em larga escala ainda depende de validação agronômica.

Palavras-chave: Tecnologia. Deep Learning. Fitopatologia. Agricultura de Precisão.

Área do Conhecimento: Ciências Agrárias – Agronomia

Introdução

O diagnóstico de doenças de plantas por meio de inspeção visual em campo permanece o método mais difundido no monitoramento fitossanitário, mas seus limites são conhecidos: o diagnóstico depende diretamente da experiência do avaliador, é suscetível a erros de subjetividade e, na maioria das vezes, só é realizado quando os sintomas já são visíveis a olho nu, momento em que os danos econômicos frequentemente já se instalaram. Diante dessa lacuna, a Inteligência Artificial (IA) tem se firmado como uma resposta tecnológica robusta, sobretudo por meio de técnicas de visão computacional baseadas em Deep Learning, capazes de automatizar o diagnóstico fitossanitário e reconhecer, em imagens digitais, padrões sutis que escapam à percepção humana (Jorge *et al.*, 2022).

Essa transformação não ocorreu isoladamente. Nas últimas décadas, a convergência entre o aumento do poder computacional, a disponibilização de bases de dados públicas de imagens de plantas e a evolução de arquiteturas de redes neurais cada vez mais sofisticadas impulsionou uma rápida diversificação das abordagens de IA voltadas à fitopatologia: das redes convolucionais clássicas a modelos de detecção em tempo real e arquiteturas baseadas em atenção (Pacal *et al.*, 2024; Shoaib *et al.*, 2023). Esse ritmo acelerado de inovação, no entanto, também produziu um volume crescente e fragmentado de evidências científicas, o que torna necessário sistematizar o conhecimento disponível, identificando as principais tecnologias empregadas, seus avanços mais recentes e as limitações que ainda restringem sua adoção em escala comercial. É nesse sentido que o presente trabalho se insere: tem como objetivo apresentar uma revisão de literatura sobre a aplicação de técnicas de visão computacional e Deep Learning na detecção e classificação de doenças de plantas, abordando redes neurais convolucionais, arquiteturas de detecção em tempo real (YOLO), Vision Transformers e os desafios relacionados à explicabilidade e à generalização dos modelos, eixos que estruturam as seções seguintes deste estudo.

Metodologia

Este trabalho consiste em uma revisão de literatura de caráter narrativo/integrativo, cujo propósito foi sintetizar e analisar criticamente publicações científicas recentes sobre o uso de Inteligência Artificial na detecção de doenças de plantas. A busca foi realizada em periódicos científicos revisados por pares e indexados em bases como Scopus, Web of Science e SciELO, contemplando os periódicos Scientific Reports, Frontiers in Plant Science, Artificial Intelligence Review e Engineering Research Express, priorizando-se publicações do período de 2022 a 2025, de modo a assegurar a atualidade das informações apresentadas.

Os descritores utilizados na busca, em português e inglês, incluíram: "deep learning plant disease detection", "convolutional neural network plant disease", "YOLO plant disease detection", "Vision Transformer plant disease classification" e "explainable AI agriculture". Foram incluídos artigos originais e revisões sistemáticas que apresentassem resultados quantitativos de desempenho de modelos, disponibilizassem DOI verificável e abordassem diretamente a detecção, classificação ou explicabilidade de modelos de IA voltados à fitopatologia. Foram excluídos materiais sem revisão por pares, como blogs e repositórios não indexados. A partir dos artigos selecionados, as informações foram organizadas tematicamente em quatro eixos analíticos (redes neurais convolucionais; arquiteturas de detecção em tempo real da família YOLO; Vision Transformers; explicabilidade e generalização dos modelos), os quais estruturam a seção de resultados a seguir.

Resultados

As redes neurais convolucionais (CNNs) permanecem a arquitetura mais amplamente empregada na classificação de doenças foliares a partir de imagens digitais. Uma revisão sistemática que analisou 160 estudos publicados entre 2020 e 2024 identificou acurácias médias superiores a 90% na maioria dos anos avaliados, evidenciando também uma diversificação progressiva das arquiteturas utilizadas, que passaram das CNNs clássicas para modelos híbridos e arquiteturas mais robustas de segmentação (Pacal *et al.*, 2024). Essa evolução é acompanhada, contudo, por desafios relacionados ao custo computacional e à necessidade de grandes volumes de dados rotulados, uma vez que os modelos de Deep Learning aprendem características diretamente das imagens, sem depender de engenharia manual de atributos, mas exigem bases de treinamento extensas e balanceadas para atingir desempenho satisfatório (Shoaib *et al.*, 2023).

Para aplicações que demandam resposta em tempo real, como o monitoramento contínuo em campo ou a integração com pulverizadores de precisão, arquiteturas de detecção de objetos da família YOLO (You Only Look Once) têm sido amplamente adotadas. Miao; Meng; Zhou (2024) propuseram o modelo SerpensGate-YOLOv8, que incorpora convoluções dinâmicas do tipo Snake e um módulo de atenção por super-tokens para melhorar a detecção de estruturas foliares finas e alongadas. Avaliado no conjunto de dados PlantDoc, composto por imagens de 13 espécies vegetais coletadas em condições reais de campo, o modelo alcançou mAP@0,5 (mean Average Precision, isto é, a média entre as classes da precisão média calculada com limiar de interseção sobre união, IoU, de 0,5) de 64,9%, o que representa ganho de 3,3 pontos percentuais em relação ao YOLOv8 original (61,6%). Esse resultado é relevante porque, ao contrário de bases de dados com fundo controlado, o PlantDoc reflete a complexidade de cenários agrícolas reais, com variação de iluminação, sobreposição de folhas e ruído de fundo, fatores que costumam reduzir sensivelmente o desempenho dos modelos treinados apenas em ambientes controlados.

Os Vision Transformers (ViT) emergiram como alternativa às CNNs ao utilizarem mecanismos de autoatenção capazes de capturar relações espaciais de longo alcance entre diferentes regiões da imagem. Borhani; Khoramdel; Najafi (2022) propuseram uma abordagem leve baseada em ViT para classificação automatizada de doenças em tempo real, testando três variantes do modelo e obtendo acurácias entre 85,87% e 94,16%. Os resultados indicam que os ViTs não substituem de forma generalizada as CNNs: a evidência disponível indica ganhos de acurácia em cenários específicos, ao custo de maior exigência computacional, o que ainda limita sua adoção em dispositivos embarcados de baixo custo utilizados em campo.

No que se refere à explicabilidade (Explainable AI, XAI), Pai; Balachandra; Kamath (2025) identificaram, em revisão sistemática das aplicações de XAI na agricultura, o uso crescente de técnicas como LIME, SHAP e Grad-CAM para tornar transparentes as decisões de modelos de caixa-preta em tarefas de discriminação de plantas daninhas, detecção de doenças e previsão de produtividade. Os

autores constataram, porém, que a maioria dos estudos carece de validação em condições reais de campo e de frameworks padronizados para avaliar a qualidade das explicações geradas. Outro resultado recorrente, identificado tanto por Pai; Balachandra; Kamath (2025) quanto por Pacal *et al.* (2024), refere-se à queda de desempenho dos modelos quando aplicados fora do contexto de treinamento: modelos treinados em bases de dados de referência, como o PlantVillage, apresentam acurácia elevada em condições controladas, mas seu desempenho cai significativamente quando aplicados a imagens de campo, em razão de variações de iluminação, ocorrência de múltiplas doenças e viés geográfico das bases de dados públicas disponíveis.

Discussão

Os resultados reunidos evidenciam uma trajetória evolutiva clara na aplicação de IA à fitopatologia: das CNNs clássicas, consolidadas como abordagem de referência (Pacal *et al.*, 2024; Shoaib *et al.*, 2023), passando pelas arquiteturas de detecção em tempo real da família YOLO, cuja principal contribuição está na viabilização de sistemas embarcados capazes de operar durante a aquisição de imagens em campo (Miao; Meng; Zhou, 2024), até os Vision Transformers, que representam a fronteira mais recente em termos de capacidade de modelagem de relações espaciais complexas (Borhani; Khoramdel; Najafi, 2022). Essa progressão não configura, contudo, uma substituição linear de tecnologias: cada arquitetura apresenta um equilíbrio distinto entre acurácia, velocidade de inferência e custo computacional, o que torna a escolha do modelo dependente do contexto de aplicação: por exemplo, dispositivos embarcados de baixo custo tendem a favorecer CNNs leves ou YOLO, enquanto cenários com maior disponibilidade computacional podem se beneficiar dos ganhos marginais de acurácia oferecidos pelos ViTs.

Um ponto de convergência importante entre os estudos analisados é a fragilidade da generalização dos modelos. O fato de arquiteturas distintas (CNNs, YOLO e ViT) compartilharem a mesma limitação ao serem confrontadas com imagens de campo reforça que o problema não está apenas na escolha da arquitetura, mas na representatividade e na diversidade das bases de dados utilizadas no treinamento (Pacal *et al.*, 2024; Pai; Balachandra; Kamath, 2025). Esse achado tem implicação direta para a adoção prática dessas tecnologias: soluções validadas exclusivamente em bases controladas, como o PlantVillage, correm o risco de apresentar desempenho insatisfatório quando transferidas para lavouras comerciais, especialmente em regiões tropicais, cujas condições edafoclimáticas e fitossanitárias diferem substancialmente daquelas predominantes nas bases de dados públicas mais utilizadas.

A discussão sobre explicabilidade, por sua vez, aponta para uma lacuna que transcende o desempenho técnico dos modelos e alcança sua legitimidade perante os usuários finais. Ainda que técnicas como LIME, SHAP e Grad-CAM ofereçam caminhos para tornar interpretáveis as decisões de modelos de caixa-preta, a ausência de padronização metodológica identificada por Pai; Balachandra; Kamath (2025) sugere que a área ainda carece de critérios consensuais para avaliar se uma explicação gerada é, de fato, confiável e útil para o tomador de decisão agronômico. Isso reforça a necessidade de que futuras pesquisas não se limitem a reportar ganhos de acurácia, mas também investiguem a robustez, a transparência e a aplicabilidade prática dos modelos propostos.

Conclusão

A presente revisão evidenciou que a visão computacional baseada em Deep Learning representa um dos pilares tecnológicos da Agricultura 4.0 aplicada à fitossanidade, oferecendo ganhos relevantes de velocidade e precisão diagnóstica em relação à inspeção visual tradicional. As CNNs seguem como abordagem consolidada, as arquiteturas YOLO respondem à demanda por detecção em tempo real e os Vision Transformers surgem como alternativa promissora para cenários que exigem maior capacidade de modelagem espacial. Entretanto, a maturidade dessas tecnologias para uso operacional em larga escala ainda depende da superação de desafios relacionados à generalização dos modelos para condições reais de campo, à padronização de métodos de explicabilidade e à ampliação de bases de dados representativas de diferentes culturas, regiões e estádios fenológicos. Reforça-se, assim, que a validação agronômica continua sendo etapa indispensável antes da adoção dessas ferramentas em escala comercial, e que estudos futuros devem priorizar a avaliação de desempenho em condições reais de campo e o desenvolvimento de bases de dados mais diversificadas e representativas.

Referências

BORHANI, Y.; KHORAMDEL, J.; NAJAFI, E. A deep learning based approach for automated plant disease classification using vision transformer. **Scientific Reports**, v. 12, art. 11554, 2022. DOI: 10.1038/s41598-022-15163-0. Disponível em: https://doi.org/10.1038/s41598-022-15163-0. Acesso em: 24 ago. 2026.

JORGE, L. A. de C.; OSCO, L. P.; MARCATO JÚNIOR, J.; RAMOS, A. P. M.; GONÇALVES, W. N.; SPERANZA, E. A.; VAZ, C. M. P.; INAMASU, R. Y. Inteligência artificial aplicada na agricultura de precisão e digital. In: **Agricultura de precisão**: um novo olhar na era digital. Brasília, DF: Embrapa, 2022. p. 51-64. DOI: 10.4322/978-65-86819-38-0.1000082. Disponível em: https://doi.org/10.4322/978-65-86819-38-0.1000082. Acesso em: 24 ago. 2026.

MIAO, Y.; MENG, W.; ZHOU, X. SerpensGate-YOLOv8: an enhanced YOLOv8 model for accurate plant disease detection. **Frontiers in Plant Science**, v. 15, art. 1514832, 2024. DOI: 10.3389/fpls.2024.1514832. Disponível em: https://doi.org/10.3389/fpls.2024.1514832. Acesso em: 24 ago. 2026.

PACAL, I.; KUNDURACIOGLU, I.; ALMA, M. H.; DENIZ, M.; ÜNAL, A.; ALTUN, F.; GENC, A.; GENC, S. A systematic review of deep learning techniques for plant diseases. **Artificial Intelligence Review**, v. 57, art. 304, 2024. DOI: 10.1007/s10462-024-10944-7. Disponível em: https://doi.org/10.1007/s10462-024-10944-7. Acesso em: 16 ago. 2026.

PAI, D. G.; BALACHANDRA, M.; KAMATH, R. Explainable AI in agriculture: review of applications, methodologies, and future directions. **Engineering Research Express**, v. 7, n. 3, art. 032202, 2025. DOI: 10.1088/2631-8695/ae086a. Disponível em: https://doi.org/10.1088/2631-8695/ae086a. Acesso em: 15 ago. 2026.

SHOAIB, M.; SHAH, B.; EL-SAPPAGH, S.; ALI, A.; ULLAH, A.; ALENEZI, F.; GECHEV, T.; HUSSAIN, T.; ALI, F. An advanced deep learning models-based plant disease detection: a review of recent research. **Frontiers in Plant Science**, v. 14, art. 1158933, 2023. DOI: 10.3389/fpls.2023.1158933. Disponível em: https://doi.org/10.3389/fpls.2023.1158933. Acesso em: 25 ago. 2026.

Agradecimentos

Os autores expressam seus agradecimentos ao Laboratório de Epidemiologia e Manejo de Doenças de Plantas Agrícolas e Florestais (LEMP) e ao Núcleo de Desenvolvimento Científico e Tecnológico em Manejo Fitossanitário de Pragas e Doenças (NUDEMAFI), vinculados à Universidade Federal do Espírito Santo – Campus de Alegre, pela disponibilização da infraestrutura necessária à execução deste estudo. Agradecem, igualmente, à Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro concedido para o desenvolvimento da pesquisa.