



Avaliação da eficácia de detecção de doenças em culturas de soja utilizando visão computacional

Evaluation of disease detection effectiveness in soybean crops using computer vision

 Camila Ribeiro de Moraes ¹,  Alessandro Rodrigues e Silva ¹,  Thiago Augusto Mendes ¹,
 Augusto Souza Batista ²

RESUMO

A detecção precoce de doenças em culturas de soja é fundamental para reduzir perdas produtivas. No entanto, métodos tradicionais baseados em inspeções manuais são pouco eficientes. Este trabalho apresenta uma abordagem utilizando redes neurais profundas para classificar doenças em plantas de soja. Modelos pré-treinados (ResNet101, VGG19, DenseNet201 e MobileNetV3 Large) foram ajustados com imagens terrestres e aplicados a imagens aéreas obtidas por veículos aéreos não tripulados. A diferença entre os domínios afetou o desempenho, com o DenseNet201 alcançando os melhores resultados em imagens terrestres (acurácia de 0,9691 e macro-F1 de 0,9690). Ainda assim, o modelo identificou 90% das plantas doentes em imagens aéreas. Conclui-se que a abordagem é promissora, mas exige aprimoramentos no pré-processamento e na adaptação de domínio.

Palavras-chave: Aprendizagem profunda; Veículos aéreos não tripulados; Glycine max.

ABSTRACT

Early detection of diseases in soybean crops is essential to reduce yield losses. However, traditional methods based on manual inspections are inefficient. This study presents a deep neural network approach for classifying diseases in soybean plants. Pre-trained models (Res-Net101, VGG19, DenseNet201, and MobileNetV3 Large) were fine-tuned with ground-level images and applied to aerial images captured by unmanned aerial vehicles. The domain difference affected performance, with DenseNet201 achieving the best results on ground images (accuracy of 0.9691 and macro-F1 of 0.9690). Nevertheless, the model correctly identified 90% of diseased plants in aerial images. It is concluded that the proposed approach is promising but requires improvements in preprocessing and domain adaptation techniques.

Keywords: Deep learning; Unmanned aerial vehicles; Glycine max.

¹ Instituto Federal de Goiás (IFG) - Câmpus Anápolis;

Autor correspondente Camila Ribeiro de Moraes: camilaribeiro0308@gmail.com

Recebido em: 29/10/2025. Aceito em: 20/06/2026.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado. 

1 INTRODUÇÃO

O monitoramento eficaz das plantações é importante para otimizar a produção agrícola e reduzir perdas (Silva *et al.*, 2020). Neste contexto, a implementação de estratégias de monitoramento de baixo custo e eficientes, especialmente para combater doenças, pode ser benéfica para safras.

A agricultura de precisão surge como uma solução candidata, oferecendo uma variedade de técnicas que podem melhorar a gestão das plantações e maximizar a eficiência no uso de recursos, como insumos químicos (pesticidas e herbicidas), água e fertilizantes. Dentre essas técnicas, a pulverização agrícola destaca-se por possibilitar a aplicação eficiente de insumos, contribuindo significativamente para proteção e o desenvolvimento das culturas agrícolas. O uso de drones ou Veículos Aéreos Não Tripulados (VANT), no processo de pulverização agrícola mostra-se eficaz, aumentando a precisão na identificação de áreas críticas e na dispersão de produtos para tratamentos (EMBRAPA, 2022).

Além disso, drones podem alcançar áreas das plantações, no contexto dessa pesquisa, em especial a cultura de soja, inacessíveis pelos métodos tradicionais de pulverização (EMBRAPA, 2022), mantendo os padrões, parâmetros e qualidade já estabelecidos na prática convencional.

Identificar doenças em plantações é essencial para a atividade agrícola (Khakimov *et al.*, 2022). Detectar patologias de maneira ágil e eficaz é uma tarefa adequada para aplicações de ferramentas que compreendem a área de visão computacional. A visão computacional aplicada ao contexto de reconhecimento de patologias em plantações oferece uma abordagem automatizada e eficiente para monitorar, diagnosticar e quantificar estes problemas com o desenvolvimento de diferentes culturas, permitindo uma resposta rápida e precisa para minimizar os danos às colheitas e elevar a produtividade.

Ademais, utilizar drones para a captura de imagens de alta resolução ao sobrevoar plantações possibilita a aplicação de técnicas de visão computacional. Câmeras embarcadas nos drones podem capturar fotografias detalhadas de folhagens, caules e até mesmo do solo (Coelho e Cavichioli, 2021; Zhu *et al.*, 2024; Soares e Schröder, 2025) que permitem uma análise rigorosa de características visuais de diferentes doenças em diversos estágios evolutivos.

Portanto, este artigo investigou técnicas de visão computacional aplicadas no diagnóstico de doenças em culturas de soja, para identificar folhas doentes em mapeamento de áreas de cultivo, bem como criar um plano de voo otimizado para operações defensivas.

2 MATERIAIS E MÉTODOS

A metodologia adotada neste artigo baseou-se nas etapas do método de *Knowledge Discovery in Databases* (KDD) (Fayyad *et al.*, 1996). Inicialmente, realizou-

se a seleção dos dados, composta por imagens de plantas de soja (folhagens) obtidas em condições controladas e de campo. Em seguida, procedeu-se o pré-processamento, etapa responsável pela limpeza, padronização e tratamento das imagens para garantir a qualidade dos dados utilizados. A etapa de mineração de dados consistiu no treinamento e validação de modelos de aprendizado profundo, especificamente Redes Neurais Convolucionais (CNNs) aplicadas à visão computacional, na intenção de identificar padrões visuais relevantes para distinguir folhas saudáveis e doentes. Por fim, na fase de interpretação e avaliação, o conhecimento gerado foi aplicado em campo para subsidiar a elaboração de um plano de pulverização agrícola mais preciso e eficiente.

2.1 OBTENÇÃO E COMPOSIÇÃO DA BASE DE DADOS

Inicialmente, foi montada uma base de dados de imagens de plantas de soja (folhagens) com aspectos saudáveis e patológicos, como manchas de diferentes tonalidades das folhas provocadas por diferentes doenças.

Para as condições de campo, as imagens das folhagens das plantas de soja foram obtidas em março de 2024 e agrupadas em duas classes: plantas de soja saudáveis, que incluíam plantas com folhas apresentando coloração, tamanho, morfologia e desenvolvimento compatíveis com o estágio fenológico observado, sem a presença de sintomas visíveis de estresse, deformidades ou doenças conhecidas e, plantas de soja doentes, que consideravam plantas com folhas de espectro de cores incoerentes, degradação ou quaisquer aspectos que fugissem do padrão normal, como: clorose, necrose, manchas foliares, murcha, descoloração anormal dos tecidos ou outros sintomas visíveis que indicassem comprometimento fitossanitário.

Essa base de dados foi então integrada a outras três bases de repositórios públicos, sendo elas: *Image of Soybean Leaves* (Mignoni, 2021), contendo imagens com orifícios e recortes nas folhagens de soja causados por pragas (lagartas); *Soybean Leaf Image Dataset* (Abdullah, 2024), que contém imagens de folhas de soja saudáveis e com doenças comuns da cultura (como pústula bacteriana, mancha foliar olho de sapo, ferrugem asiática, mancha alvo, mosaico amarelo e oídio) e; a base de Imagens de Sintomas de Doenças de Plantas PDDB (Barbedo *et al.*, 2024), dispondo de imagens de diversas espécies de plantas saudáveis e doentes, incluindo soja, elaborada por pesquisadores da Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

A contribuição de cada uma das três bases ocorreu por meio de uma seleção manual de imagens, priorizando aquelas com a melhor qualidade em termos de iluminação, brilho, resolução e proximidade do zoom para uma visualização e identificação clara e precisa das folhagens envolvidas.

Considerando a base de dados da *Image of Soybean Leaves*, das 6.410 imagens totais, foram selecionadas 896 imagens de plantas saudáveis. Já da base de dados *Soybean Leaf Image Dataset*, com 1.163 imagens, foram selecionadas 110 imagens de plantas saudáveis e 660 de plantas doentes. Por fim, da base de dados PDDB, com 471 imagens, foram selecionadas apenas as imagens de plantas de soja doentes, totalizando 185 imagens.

Assim, a base de dados integrada final de imagens consistiu em 1.851 imagens, sendo 845 de plantas de soja com sintomas de diversas doenças (como antracnose, cretamento foliar de cercospora e ferrugem) e 1.006 de plantas saudáveis.

Uma segunda base de dados com imagens de folhas de plantas de soja doentes e saudáveis foi construída para treinamento e inferência em condições reais. As imagens foram capturadas em 26 de fevereiro de 2025, em uma plantação de soja localizada entre o município de Anápolis e o distrito de Goialândia, estado de Goiás, Brasil (16°23'21.0149" S, 49°5'41.5349" W), utilizando um VANT modelo DJI Mavic 3 Multispectral, a uma altitude de 8 metros.

A Figura 1 ilustra o local da área de estudo obtida pelo drone, em formato de ortomosaico. Para adaptar a detecção da doença cercosporiose, identificada na área de estudo, foram adicionadas 305 imagens da base de imagens de plantas de soja ASDID (Bever *et al.*, 2022) a esta segunda base, já que essa base adicional possui imagens de plantas de soja com cercosporiose.

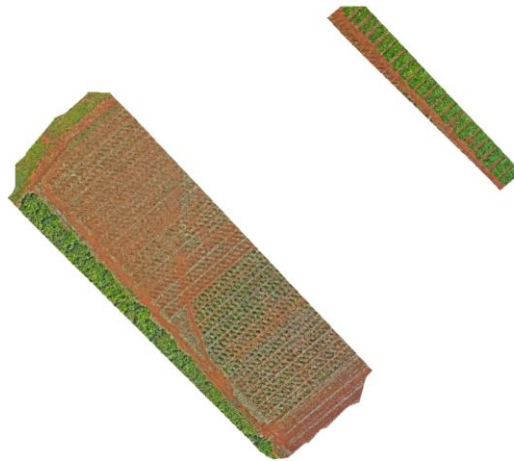


Figura 1. Região da área de estudo.
Fonte: Autoria própria (2025).

2.2 PRÉ-PROCESSAMENTO DAS IMAGENS

As imagens das folhas de plantas de soja da base de dados integrada foram submetidas a um processo de aumento de dados (*data augmentation*), a fim de ampliar a variabilidade do conjunto de treino e melhorar a generalização dos modelos sob as condições não controladas de campo.

O principal benefício dessa abordagem no treinamento de redes neurais é a redução do sobreajuste (*overfitting*), situação em que o modelo memoriza os dados de treinamento em vez de aprender padrões gerais. Como resultado, o modelo melhora sua capacidade de generalização, tornando-se mais apto a lidar com as variações do mundo real, como diferentes condições de iluminação, ângulos, posições e tamanhos das manchas de doenças nas folhas. Todas as transformações foram implementadas utilizando o módulo *transforms* da biblioteca PyTorch (Paszke *et al.*, 2019).

O processamento iniciou-se com transformações espaciais e geométricas. Primeiramente, as imagens das folhas da soja foram submetidas a um corte aleatório

(*Random Resized Crop*) para uma dimensão final de 224×224 pixels, com uma escala de corte variando entre 80 e 100% da área original. Esta abordagem não apenas padroniza o tamanho da entrada para a arquitetura da rede, otimizando o custo computacional, mas também simula diferentes enquadramentos, forçando o modelo a identificar alvos em posições e escalas diferentes.

Adicionalmente, para aumentar a invariância do modelo à orientação, foram aplicadas inversões horizontais e verticais com 50% de probabilidade, rotações aleatórias de até 45° , distorções de cisalhamento (*shear*) de até 20° e alterações de perspectiva. Em conjunto, essas operações simulam a diversidade de ângulos e pontos de vista da captura de imagens em campo.

Além das manipulações geométricas, foram aplicadas transformações para simular a variabilidade fotométrica. Ajustes aleatórios de brilho, contraste e saturação, com um fator de 0,3 prepararam o modelo para operar sob diferentes condições de iluminação, como intensidade solar variável e sombras. Para simular cenários de captura com menor nitidez, como os causados por movimento ou problemas de foco, aplicou-se também um desfoque gaussiano, garantindo que o modelo pudesse extrair características relevantes mesmo de imagens com qualidade inferior.

Como etapa final, os valores dos pixels das imagens foram normalizados. Este processo padroniza a escala dos dados de entrada, subtraindo a média ($\mu = [0,485; 0,456; 0,406]$) e dividindo pelo desvio padrão ($\sigma = [0,229; 0,224; 0,225]$) para cada canal de cor (RGB). O uso desses valores, que são padrão para modelos pré-treinados, é importante para garantir a estabilidade do treinamento e acelerar a convergência da rede neural, quando a rede aprende padrões consistentes dos dados de treinamento.

Em conjunto, essa cadeia de transformações enriqueceu significativamente a variabilidade do conjunto de dados, permitindo que o modelo aprendesse características mais complexas e generalizáveis para a detecção de doenças nas plantações de soja sob condições de campo não controladas.

2.3 ARQUITETURA E TREINAMENTO DOS MODELOS

Foram avaliadas quatro arquiteturas CNNs pré-treinadas com a base de dados de imagens de diversas culturas e doenças da soja, sendo: PDDD-Pretrain (Dong *et al.*, 2023): ResNet101, VGG19, DenseNet201 e MobileNetV3 Large.

A técnica de aprendizado por transferência (*transfer learning*) foi aplicada para otimizar o processo de treinamento, reduzindo o tempo e os recursos computacionais, visto que o pré-treinamento com dados de imagens de plantas facilita a adaptação dos modelos à tarefa de classificação binária deste estudo. Para isso, as camadas de classificação originais foram substituídas por novas camadas totalmente conectadas, adaptadas para a tarefa de classificação binária.

A otimização de hiperparâmetros foi realizada por meio da Otimização Bayesiana (Wu *et al.*, 2019) para determinar a melhor configuração de número de camadas densas, neurônios por camada, taxa de *dropout*, função de ativação, otimizador (Adam ou SGD com *momentum*) e tamanho do lote (*batch size*). O

treinamento foi conduzido utilizando a função de perda de entropia cruzada, um agendador de taxa de aprendizado StepLR (redução a cada 10 épocas) e um critério de parada antecipada (*early stopping*) com *patience* de 50 épocas, e um limite máximo de 500 épocas.

2.4 AVALIAÇÃO DE DESEMPENHO DOS MODELOS

O desempenho dos modelos foi avaliado utilizando validação cruzada estratificada de 10 partições (*10-fold*). A partir da base de dados integrada (1.851 imagens), a divisão foi realizada em duas etapas com a função *train_test_split* da biblioteca *scikit-learn*, mantendo a estratificação por classe.

Na primeira etapa, o conjunto total foi separado em 75% para treino e 25% como restante. Na segunda etapa, esse subconjunto de 25% restante, foi subdividido em 15% para validação, utilizado no ajuste dos hiperparâmetros via Otimização Bayesiana, e 10% para teste, reservado para a avaliação final do modelo em dados não vistos. Obteve-se, assim, uma divisão de 75/15/10 (treino/validação/teste) referente ao total de imagens.

Essa divisão reflete uma prática heurística que se alinha com os princípios de avaliação de modelos, como descrito por Goodfellow *et al.* (2016), que busca equilibrar dois objetivos: reservar a maior parte dos dados para o aprendizado dos padrões (treino), e garantir ao mesmo tempo, conjuntos independentes para o ajuste de hiperparâmetros (validação) e para uma estimativa imparcial do desempenho de generalização (teste), reduzindo o risco de sobreajuste e de avaliação otimista.

Para a avaliação dos modelos foram utilizadas as métricas de acurácia e macro-f1, comumente empregadas em tarefas de classificação supervisionada para aferir tanto o desempenho geral quanto o balanceamento entre as classes (Sokolova e Lapalme, 2009).

A acurácia representa a proporção de predições corretas (positivas e negativas) em relação ao número total de amostras avaliadas, fornecendo uma visão geral da performance do classificador. Sua fórmula é expressa na Equação 1:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

onde: *VP* corresponde a verdadeiros positivos (amostras positivas classificadas corretamente), *VN* a verdadeiros negativos (amostras negativas classificadas corretamente), *FP* a falsos positivos (amostras negativas classificadas incorretamente como positivas) e *FN* a falsos negativos (amostras positivas classificadas incorretamente como negativas).

Apesar de sua simplicidade, a acurácia pode ser enganosa em cenários com classes desbalanceadas. Por essa razão, utilizou-se também a métrica macro-f1, que avalia o desempenho de forma mais completa por ser menos sensível ao desbalanceamento entre as classes. Diferentemente da acurácia, que pode ser dominada pela classe majoritária, a macro-f1 calcula a média aritmética do f1-score

de cada classe, atribuindo a todas igual importância independentemente do número de amostras (Sokolova e Lapalme, 2009).

O f1-score, por sua vez, é a média harmônica entre a precisão (*precision*), a exatidão das predições positivas e, a revocação (*recall*), a capacidade de encontrar todas as amostras positivas, balanceando a capacidade do modelo de acertar ao prever uma classe e de identificar todas as amostras daquela classe. A fórmula da macro-f1 é expressa na Equação 2.

$$\text{Macro-F1} = \frac{1}{N} \sum_{i=1}^N F1_i \quad (2)$$

onde: N é o número total de classes e $F1_i$ é o f1-score da i -ésima classe.

Essa abordagem oferece uma avaliação mais justa do desempenho do modelo em todas as categorias, sendo particularmente útil quando a distribuição de classes é desigual (Sokolova e Lapalme, 2009).

Adicionalmente, a técnica Grad-CAM (Selvaraju *et al.*, 2019) foi empregada para visualizar as regiões de interesse das imagens que mais influenciaram a decisão dos modelos, auxiliando na interpretabilidade dos resultados.

Essa técnica gera um mapa de calor que sobrepõe a imagem original, e destaca as áreas mais relevantes que influenciaram a decisão do modelo para determinada classificação de imagem da folha (planta) da soja doente ou saudável. As regiões com cores mais quentes do mapa de calor indicam os pontos em que a rede neural “observou com maior atenção” as patologias presentes, o que permite visualizar se o modelo classificou a imagem da planta de soja corretamente conforme as características visuais correlatas aos aspectos de treinamento para as plantas consideradas doentes.

2.5 APLICAÇÃO EM CAMPO E GERAÇÃO DO PLANO DE PULVERIZAÇÃO

Para a aplicação prática da solução, o modelo com melhor desempenho (DenseNet201) foi submetido a um segundo ajuste fino, utilizando as imagens aéreas capturadas pelo VANT. O objetivo desse segundo ajuste foi refinar o modelo para duas novas tarefas: a classificação de doenças a partir de uma perspectiva aérea e a identificação específica da doença cercosporiose, que era a única patologia presente na plantação da área de estudo.

A transição das imagens terrestres (treinamento) para as imagens aéreas (inferência) caracteriza um cenário de *domain shift*, em que as distribuições dos dados de treinamento e de inferência divergem em resolução espacial, escala, iluminação, ângulo de captura e composição da cena. Enquanto as imagens terrestres foram capturadas próximas das plantas de soja, com alto detalhamento de textura foliar, as imagens aéreas obtidas pelo VANT a 8 metros de altitude, apresentam menor resolução por folha, maior incidência de oclusão entre plantas e presença de elementos de fundo, como solo e sombra, em cada *tile*.

Para mitigar esse deslocamento de domínio, adotou-se uma estratégia de adaptação supervisionada por *fine-tuning*, na qual as camadas convolucionais iniciais

do DenseNet201, responsáveis por características de baixo nível (bordas, cores, gradientes), foram mantidas congeladas, preservando o conhecimento transferido, enquanto as camadas finais e a camada totalmente conectada foram re-treinadas com as imagens aéreas rotuladas da cultura de soja.

O processo iniciou-se com a geração de um ortomosaico georreferenciado da área de estudo, a partir do processamento das imagens na plataforma DJI SmartFarm Web.

Sobre este mapa, foi calculado o Índice de Vegetação por Diferença Normalizada (NDVI) (Werner *et al.*, 2018) para mascarar áreas sem vegetação e focar a análise apenas nas plantas (Equação 3). Esse índice mensura o quanto de vegetação há em uma imagem por meio do cálculo das bandas do infravermelho próximo (NIR) e do vermelho (RED), a partir da razão entre a diferença e a soma das refletâncias da vegetação no NIR e no RED, ambos com comprimentos de onda descritos na Equação 3.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3)$$

Os valores de refletância utilizados no cálculo do NDVI correspondem a comprimentos de onda do espectro eletromagnético, expressos em micrômetros (μm).

O ortomosaico foi então segmentado em fragmentos (*tiles*) de 224×224 *pixels*, formato compatível com a entrada da rede neural, e o modelo DenseNet201 ajustado realizou a inferência de cada *tile*, identificando as áreas com plantas de soja doentes e plantas de soja saudáveis. Os resultados da inferência foram exportados como um arquivo *shapefile* e manipulados no software QGIS (QGIS, 2021). Nele, os *tiles* classificados como plantas doentes foram selecionados e um *buffer* de 5 metros foi aplicado ao redor dessas áreas para criar uma área de aplicação.

O *shapefile* final, contendo o mapa de pulverização de precisão, foi então gerado, estando pronto para ser importado em sistemas de planejamento de missão de VANTs agrícolas.

3 RESULTADOS E DISCUSSÃO

Os resultados obtidos do desempenho dos modelos CNNs na tarefa de classificação de imagens de folhas de plantas de soja saudáveis e doentes são apresentados de forma sequencial, iniciando pela avaliação comparativa dos modelos, seguindo com a análise do desempenho da arquitetura DenseNet201 após o ajuste fino para imagens aéreas e, por fim, com a validação em campo e o processo de geração do plano de pulverização.

Para o treinamento das diferentes arquiteturas utilizadas, na Tabela 1 são sintetizados e apresentados os resultados do desempenho de cada modelo para o conjunto de teste.

Tabela 1. Resultados dos desempenhos dos modelos

Modelo	Acurácia	macro-F1
Resnet101	0,9433	0,9431
VGG19	0,9639	0,9637
DenseNet201	0,9691	0,9690
MobileNetV3 Large	0,9485	0,9480

Fonte: Autoria própria (2025).

Os resultados mostram que todos os modelos utilizados apresentaram um bom desempenho (Tabela 1), com acurácias superiores a 94%, evidenciando a viabilidade da abordagem proposta. Entre eles, as arquiteturas DenseNet201 e VGG19 se destacaram como as mais eficazes, alcançando acurácias de 0,9691 e 0,9639, respectivamente.

Esse desempenho superior pode ser explicado pela capacidade dessas CNNs em capturar padrões complexos e de alto nível em imagens, como microtexturas nas superfícies foliares ou transições sutis de cor que indicam estágios iniciais de certas doenças. No caso específico da DenseNet201, suas conexões densas favorecem o fluxo de gradientes e a reutilização de características entre camadas, aprimorando o aprendizado de representações mais complexas.

Os resultados obtidos estão em consonância com outros resultados de pesquisas recentes na área. Por exemplo, Yu *et al.* (2022) utilizaram uma rede de atenção residual (RANet) para identificar doenças em folhas de soja e alcançaram uma acurácia de 98,49%. Embora essa acurácia seja ligeiramente superior, o referido estudo utilizou uma base de dados de imagens de plantas de soja não pública e um pré-processamento focado na segmentação do fundo da imagem, o que pode ter simplificado a tarefa de classificação.

Já Amara, Bouaziz e Algergawy (2017), obtiveram 92,88% de acurácia na classificação de doenças em folhas de banana com a arquitetura LeNet, e os modelos DenseNet201 e VGG19 utilizados, apresentaram um desempenho superior, o que reforça a eficácia de arquiteturas complexas para tarefas fitopatológicas.

Em contrapartida, os modelos ResNet101 e MobileNetV3 Large apresentaram um desempenho ligeiramente inferior, com acurácias de 0,9433 e 0,9485, respectivamente. No caso da ResNet101, a presença de conexões residuais, embora essenciais para treinar redes muito profundas, pode induzir um efeito de regularização que limita a capacidade do modelo de se ajustar a padrões excessivamente complexos. Já a MobileNetV3 Large, projetada para eficiência de recursos computacionais para dispositivos móveis, possui uma arquitetura mais concisa. Isso resulta em um modelo com maior viés no dilema viés-variância, restringindo sua capacidade de capturar as nuances mais sutis das imagens, apesar de sua excelente velocidade de convergência por conta do tamanho da arquitetura.

Para ter uma análise mais ampla dos acertos de cada classe dos modelos, o uso de matrizes de confusão foi empregado. A Figura 2a, ilustra a matriz de confusão do modelo ResNet101. Este por sua vez identificou corretamente 86 imagens de plantas consideradas doentes e 97 imagens de plantas saudáveis, no entanto, classificou

erroneamente seis imagens de plantas doentes como saudáveis e cinco imagens de plantas saudáveis como doentes.

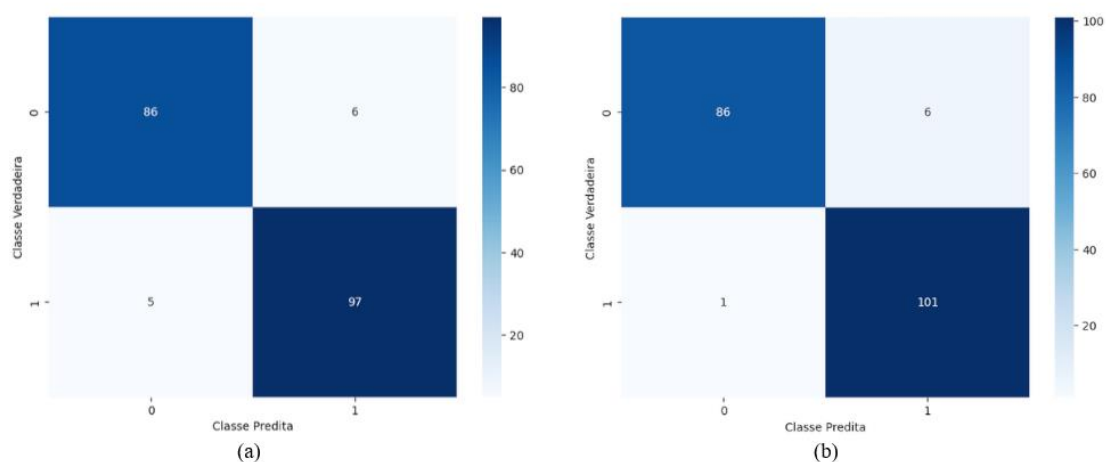


Figura 2. Matriz de confusão: (a) rede ResNet101; (b) rede VGG19.

Fonte: Autoria própria (2025).

A Figura 2b, ilustra o resultado da matriz de confusão da rede VGG19, a qual 101 imagens de plantas saudáveis foram classificadas como de fato saudáveis e 86 imagens de plantas doentes foram classificadas como plantas doentes. No entanto, 6 imagens foram classificadas como imagens de plantas saudáveis, mas estas sendo doentes, e uma única imagem foi classificada como imagem de planta doente, sendo saudável.

A Figura 3a ilustra a matriz de confusão do primeiro *fine-tuning* da DenseNet201, a qual classificou corretamente 100 imagens de plantas saudáveis e 88 imagens de plantas doentes, classificando erroneamente 4 imagens de plantas doentes contra 2 imagens de plantas saudáveis.

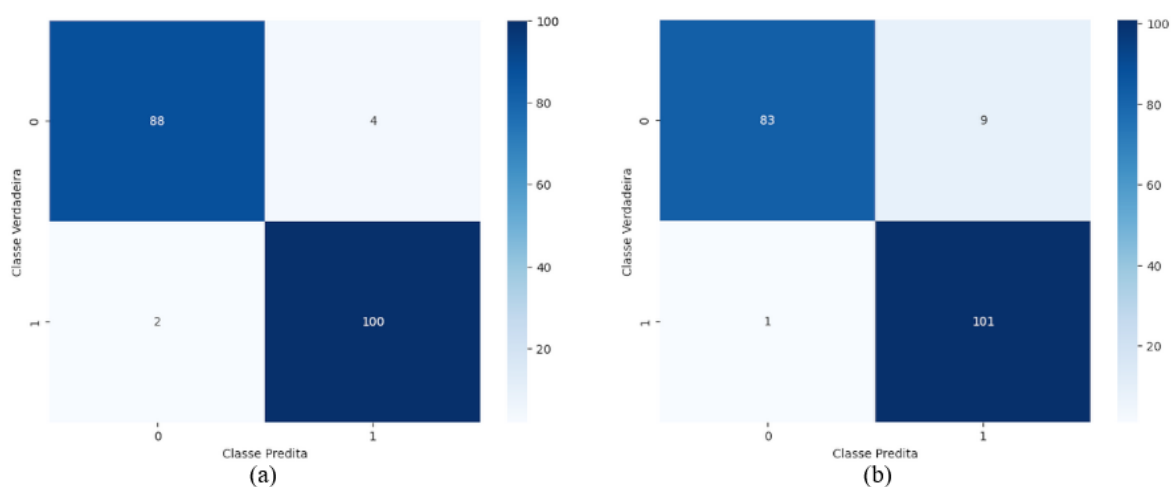


Figura 3. Matriz de confusão: (a) rede DenseNet201; (b) rede MobileNetV3.

Fonte: Autoria própria (2025).

Por fim, a Figura 3b ilustra a matriz de confusão do modelo MobileNetV3. Este modelo, por sua vez, classificou corretamente 101 imagens de plantas saudáveis e 83 imagens de plantas doentes e errou em afirmar que 9 imagens eram de plantas saudáveis e 1 imagem como planta doente.

A análise das matrizes de confusão de todos os modelos revelou uma alta capacidade de identificar corretamente ambas as classes de plantas saudáveis e doentes, com um número reduzido de falsos positivos e negativos, indicando que as características extraídas foram suficientemente discriminativas, validando a abordagem para o conjunto de dados em perspectiva terrestre.

3.1 ADAPTAÇÃO E DESEMPENHO DO MODELO CONSIDERANDO AS IMAGENS AÉREAS

Considerando o melhor desempenho obtido entre os modelos utilizados, o DenseNet201 foi selecionado para a etapa de aplicação em campo, que envolve a inferência das imagens aéreas capturadas por drone. Conforme detalhado na Seção 2.5, o modelo foi submetido a um segundo ajuste fino para adaptá-lo ao domínio aéreo, mitigando o salto de domínio (*domain shift*) decorrente das diferenças de escala, resolução, iluminação e perspectiva entre imagens terrestres de treinamento e as imagens aéreas de inferência.

Após esse ajuste, o modelo alcançou uma acurácia de 0,9237 e um macro-f1 de 0,9217. A queda na acurácia em relação ao desempenho com imagens terrestres era esperada e evidencia a complexidade introduzida pela perspectiva das imagens aéreas. Entretanto, este resultado é promissor, especialmente quando comparado a estudos que utilizam exclusivamente dados de sensoriamento remoto. Qin *et al.* (2021), alcançaram 79% de precisão na detecção de doenças em pinheiros utilizando imagens multiespectrais de drones, o que posiciona nosso resultado de 92,37% com imagens RGB como competitivo para o monitoramento agrícola.

A matriz de confusão resultante do *fine-tuning* (Figura 4), para o conjunto de teste das imagens aéreas, classificou corretamente 135 imagens de plantas doentes e 95 de plantas saudáveis, com um número ainda baixo de erros (11 falsos negativos e 8 falsos positivos), confirmando sua eficácia mesmo após a adaptação de domínio.

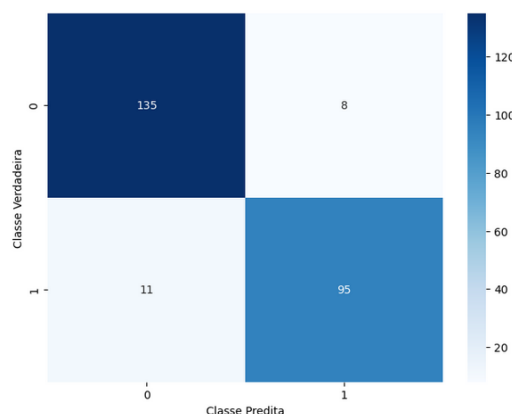


Figura 4. Matriz de confusão da rede DenseNet201 após segundo ajuste fino.

Fonte: Autoria própria (2025).

3.2 VALIDAÇÃO EM CAMPO E APLICAÇÕES PARA PULVERIZAÇÃO

A etapa final consistiu em aplicar o modelo ajustado sobre um ortomosaico de uma plantação de soja para ser gerado um mapa de pulverização. Por conta do desafio biológico de encontrar, na plantação da área de estudo utilizada para este artigo, plantas com doenças e validar a solução, houve a necessidade de simular um cenário com a presença de cercosporiose nas plantas (capturadas pelo mesmo drone em outra plantação de soja). *Tiles* do ortomosaico original foram substituídos probabilisticamente por imagens aéreas rotuladas (40 saudáveis e 30 doentes). O resultado do ortomosaico com a aplicação do modelo pode ser visualizado na Figura 5.

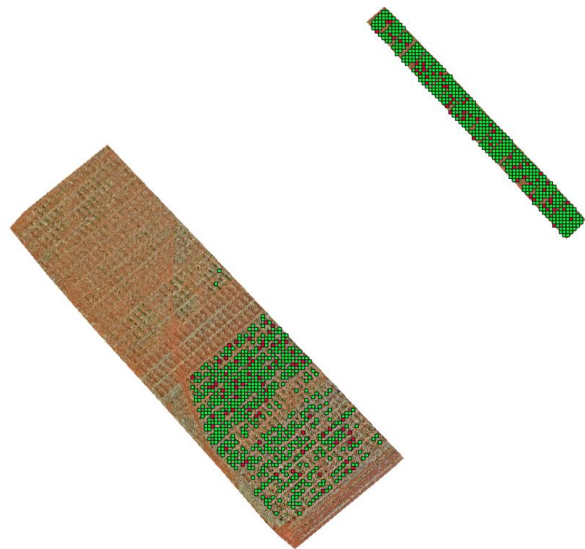


Figura 5. Ortomosaico da região da área de estudo (pontos verdes para plantas saudáveis e vermelhos para plantas doentes).
Fonte: Autoria própria (2025).

A inferência do modelo sobre este conjunto de 70 imagens revelou um desafio para a aplicação prática, enquanto o modelo identificou corretamente 27 das 30 imagens de plantas doentes, ele classificou erroneamente 31 das 40 imagens de plantas saudáveis como doentes (alto número de falsos positivos). Este resultado sugere que, embora o ajuste fino tenha sido eficaz, o modelo tornou-se excessivamente sensível a variações nas imagens aéreas que não correspondem à doença, como sombras, variações de umidade no solo ou diferentes estágios das doenças na planta, interpretando-as como anomalias.

Essa dificuldade para generalizar é um ponto crítico discutido por Lee *et al.* (2020), que observaram que as CNNs, por vezes, não focam nas regiões foliares efetivamente afetadas pela doença, mas em outras características contextuais da imagem para realizar a classificação. No caso deste artigo, o modelo, treinado com imagens terrestres ricas em detalhes de lesões (doenças), pode estar falhando em identificar características de plantas saudáveis na perspectiva aérea, que é menos

detalhada, levando-o a classificar qualquer variação como um sintoma potencial para uma doença.

O ajuste fino supervisionado adotado, embora tenha elevado o desempenho no conjunto de teste aéreo, mostrou-se insuficiente para a generalização em campo aberto, sobretudo pelo volume limitado de imagens aéreas rotuladas disponíveis. Esse comportamento indica que o re-treinamento das camadas finais, isoladamente, não recompõe a representação interna necessária quando a discrepância entre os domínios é acentuada.

A literatura aponta caminhos mais robustos, não esgotados neste estudo: o *domain-adversarial training* (Ganin *et al.*, 2016), que induz a CNN a aprender características invariantes ao domínio a partir de dados-alvo não rotulados, o alinhamento estatístico de características entre os domínios de origem e destino, o uso de pseudorrotulos (*pseudo-labels*) para aproveitar imagens aéreas não rotuladas e o treinamento misto, combinando imagens terrestres e aéreas desde o início, de modo que o modelo aprenda uma única representação que abranja ambas as perspectivas. Adicionalmente, técnicas de pré-processamento que aproximem visualmente os dois domínios, como reamostragem de escala e equalização de iluminação, poderiam reduzir a distância entre as distribuições antes do treinamento.

Apesar dos resultados, a metodologia aplicada para a criação do mapa de pulverização foi bem-sucedida no objetivo de gerar um produto operacional, ou seja, o fluxo converteu as predições do modelo em um *shapefile* georreferenciado, delimitando as zonas de aplicação com um *buffer* de 5,0 metros ao redor de cada região classificada como plantas doente, em formato diretamente importável por sistemas de planejamento de voo de VANTs agrícolas. A aplicação possibilitou transcrever as predições do modelo DenseNet201 (ajustado para o domínio aéreo) em um plano de voo para pulverização, com a delimitação de zonas de interesse em um raio de cinco metros ao se detectar uma região com doença. A esse raio de aplicação não se visa apenas tratar plantas visivelmente doentes, mas também criar uma forma de proteção para conter a disseminação de patógenos. A Figura 6 ilustra o plano de pulverização para o estudo e resultado.

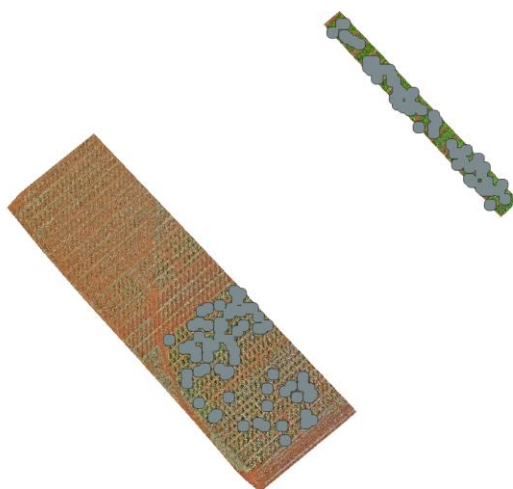


Figura 6. Ortomosaico com o mapa de pulverização (pontos cinzas são as zonas de pulverização com base na inferência do modelo).

Fonte: Autoria própria (2025).

Entretanto, o alto índice de falsos positivos indica que, no estado atual, a aplicação resultaria em um uso de defensivos maior que o necessário, indo parcialmente contra o objetivo de otimização de insumos. Mas em casos reais o maior uso de insumos seria mais benéfico do que uma quantidade ínfima que permitiria a proliferação de patógenos.

A principal contribuição deste artigo reside em demonstrar um fluxo de trabalho completo para o objetivo proposto, desde o treinamento e a comparação dos modelos avaliados até a geração de um mapa de pulverização e ademais, identificar o “salto de domínio” como o principal empecilho tecnológico a ser superado.

4 CONCLUSÃO

O desenvolvimento desse artigo se iniciou utilizando Redes Neurais Convolucionais (CNNs) pré-treinadas para classificar imagens de folhas de plantas de soja saudáveis e doentes, utilizando o ajuste fino (*fine-tuning*) para adaptar os modelos a esta tarefa. Para isso, foram selecionadas imagens de bases de dados de imagens de folhas de plantas de soja capturadas em ambiente terrestre, isto é, fotografadas com câmeras operadas manualmente.

Para a primeira etapa foram avaliados os modelos pré-treinados ResNet101, VGG19, DenseNet201 e MobileNetV3 Large, cada uma com características distintas para a tarefa de classificação. As bases de dados foram selecionadas considerando imagens de boa qualidade, garantindo uma base sofisticada para a adaptação dos modelos. A base de dados construída serve de *benchmark* para comunidade científica para teste de soluções existentes bem como o desenvolvimento e aprimoração de soluções consideradas mais complexas.

Buscou-se aplicar processamentos nas imagens com o objetivo de evidenciar características importantes das folhas de soja da base de dados. Utilizando-se de diferentes técnicas de processamento e com um auxílio visual a partir do algoritmo de Grad-CAM, a qual gerou mapas de calor com indicações das regiões mais relevantes na execução da classificação. Foi utilizado também, por meio da otimização Bayesiana a seleção dos melhores ajustes para os modelos. A avaliação dos experimentos se deu a partir das métricas de acurácia, macro-f1 e matriz de confusão. Os modelos ajustados apresentaram bons resultados, capazes de diferenciar com facilidade folhas de soja saudáveis das folhas doentes.

Com os modelos já estabelecidos, seguiu-se para uma coleta de dados em ambiente real, utilizando imagens e ortomosaicos capturados por drone em uma plantação de soja. Utilizando o *software* QGIS e técnicas para lidar com georreferenciamento foi possível integrar os dados capturados ao fluxo de inferência do modelo selecionado, o DenseNet201. Essa integração demonstra que o fluxo utilizado é viável para implementação em campo, uma vez automatizado todo o processo por meio de uma plataforma.

Na etapa final, foi realizado um segundo *fine-tuning* com o objetivo de adaptar o modelo para as imagens em ambiente aéreo, capturadas pelo drone, nas condições reais de operação. Entretanto ao aplicar o modelo nas imagens aéreas, observou-se uma queda no desempenho, evidenciando a discrepância dos dados utilizados no treinamento (ambiente terrestre) dos dados usados na inferência (ambiente aéreo). Os resultados indicam que o modelo enfrenta dificuldades em generalizar corretamente para imagens aéreas, reforçando a necessidade de estratégias adicionais para essa adaptação.

Assim, foi possível desenvolver um plano de pulverização direcionado às áreas afetadas, permitindo a aplicação precisa de calda fitossanitária sobre as plantas doentes em um raio de cinco metros ao redor. Embora o modelo DenseNet201, baseado em imagens do drone DJI Mavic 3 multiespectral, não tenha alcançado elevada precisão na identificação, o plano gerado demonstra, em análise preliminar, potencial para proteger a plantação de soja.

A aplicação da solução em campo demonstrou desempenho promissor, mas ainda com desafios a serem superados em trabalhos posteriores. Entre estes, destacam-se: a adaptação do modelo para imagens aéreas por meio de ampliação da base de dados, a incorporação de índices multiespectrais, como o NDVI, para maior precisão e, a investigação de diferentes estratégias de transferência de conhecimento, incluindo testes de *ablation*. Além disso, expandir a classificação para múltiplas patologias, desenvolver uma plataforma web que integre todo o pipeline, e otimizar planos de pulverização visando reduzir insumos, custos operacionais e impactos ambientais também são pontos de partida para dar continuidade a este trabalho.

5 AGRADECIMENTOS

O quarto autor agradece o apoio ao programa de pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) do Brasil - Chamada CNPq nº 23/2025 Bolsas de Produtividade do CNPq.

6 REFERÊNCIAS

ABDULLAH. **Soybean Leaf Image Dataset**. 2024. Conjunto de dados de imagens de folhas de soja disponível no Kaggle. Disponível em:

<https://www.kaggle.com/datasets/mamun009/soybean-leaf-image-dataset/data>.

Acesso em: 27 ago. 2025.

AMARA, J.; BOUAZIZ, B.; ALGERGAWY, A. **A deep learning-based approach for banana leaf diseases classification**. In: DATENBANKSYSTEME FÜR BUSINESS, TECHNOLOGIE UND WEB (BTW 2017) – WORKSHOPBAND. Lecture Notes in Informatics (LNI). Bonn: Gesellschaft für Informatik, 2017. p. 79-88.

BARBEDO, J. G. A.; KOENIGKAN, L. V.; HALFELD-VIEIRA, B. A.; COSTA, R. V. de; NECHET, K. L.; GODOY, C. V.; JUNIOR, M. L.; PATRICIO, F. R. A.; TALAMINI, V.; CHITARRA, L. G.; OLIVEIRA, S. A. S.; ISHIDA, A. K. N.; FERNANDES, J. M. C.; SANTOS, T. T.; CAVALCANTI, F. R.; TERAPO, D. **Base de imagens de sintomas de doenças de plantas PDDB**. Redape, 2024. DOI: <https://doi.org/10.48432/XA1OVL>.

BEVERS, N.; SIKORA, E. J.; HARDY, N. B. **Pictures of diseased soybean leaves by category captured in field and with controlled backgrounds: Auburn soybean disease image dataset (ASDID)**. Dryad, 2022. [Dataset]. DOI: <https://doi.org/10.5061/dryad.41ns1rnj3>.

COELHO, W. R.; CAVICHIOLI, F. A. A aplicação de drones na agroindústria de precisão. **Revista Interface Tecnológica**, Taquaritinga, v. 18, n. 1, p. 487-499, jul. 2021. DOI: <https://doi.org/10.31510/infa.v18i1.1108>.

DONG, X.; WANG, Q.; HUANG, Q.; GE, Q.; ZHAO, K.; WU, X.; WU, X.; LEI, L.; HAO, G. PDDD-pretrain: a series of commonly used pre-trained models support image-based plant disease diagnosis. **Plant Phenomics** (Washington, D.C.), v. 5, p. 0054, 2023. DOI: <https://doi.org/10.34133/plantphenomics.0054>.

EMBRAPA. Drones são capazes de melhorar pulverização para controle de pragas da soja. **Embrapa Notícias**, 29 mar. 2022. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/69239452/drones-sao-capazes-de-melhorar-pulverizacao-para-controle-de-pragas-da-soja>. Acesso em: 23 ago. 2025.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AI Magazine**, v. 17, n. 3, p. 37, mar. 1996. DOI: <https://doi.org/10.1609/aimag.v17i3.1230>.

GANIN, Y.; USTINOVA, E.; AJAKAN, H.; GERMAIN, P.; LAROCHELLE, H.; LAVIOLETTE, F.; MARCHAND, M.; LEMPITSKY, V. Domain-adversarial training of neural networks. **Journal of Machine Learning Research**, v. 17, n. 59, p. 1-35, 2016. Disponível em: <https://jmlr.org/papers/v17/15-239.html>. Acesso em: 13 jun. 2026.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Machine learning basics. In: GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Cambridge: MIT Press, 2016. p. 98–164. Disponível em: <https://www.deeplearningbook.org/>. Acesso em: 16 set. 2025.

KHAKIMOV, A.; SALAKHUTDINOV, I.; OMOLIKOV, A.; UTAGANOV, S. Traditional and current-prospective methods of agricultural plant diseases detection: a review.

IOP Conference Series: Earth and Environmental Science, v. 951, n. 1, p. 012002, jan. 2022. DOI: <https://doi.org/10.1088/1755-1315/951/1/012002>.

MIGNONI, M. E. **Images of Soybean Leaves**. Mendeley Data, 2021. Conjunto de dados de folhas de soja, capturadas com smartphones e drones, com foco em lagartas e *Diabrotica speciosa*. Disponível em: <https://data.mendeley.com/datasets/bycbh73438>. Acesso em: 27 ago. 2025.

PASZKE, A.; GROSS, S.; MASSA, F.; LERER, A.; BRADBURY, J.; CHANAN, G.; KILLEEN, T.; LIN, Z.; GIMELSHEIN, N.; ANTIGA, L.; DESMAISON, A.; KOPF, A.; YANG, E.; DEVITO, Z.; RAISON, M.; TEJANI, A.; CHILAMKURTHY, S.; STEINER, B.; FANG, L.; BAI, J.; CHINTALA, S. PyTorch: an imperative style, high-performance deep learning library. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS**, 32, 2019. **Anais** [...]. Red Hook, NY: Curran Associates, 2019. p. 8024-8035. Disponível em: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>. Acesso em: 27 ago. 2025.

QGIS DEVELOPMENT TEAM. **QGIS Geographic Information System**. [S.l.], 2021. Sistema de informação geográfica de código aberto. Disponível em: <https://www.qgis.org>. Acesso em: 29 mar. 2025.

SELVARAJU, R. R.; COGSWELL, M.; DAS, A.; VEDANTAM, R.; PARIKH, D.; BATRA, D. Grad-CAM: visual explanations from deep networks via gradient-based localization. **International Journal of Computer Vision**, v. 128, n. 2, p. 336–359, out. 2019. DOI: <https://doi.org/10.1007/s11263-019-01228-7>.

SILVA, A.; SILVA, B.; SOUZA, C.; AZEVEDO, B.; BASSOI, L.; VASCONCELOS, D.; BONFIM, G.; MANZANO, J.; SANTOS, A. F.; CARNEIRO, F. Irrigation in the age of agriculture 4.0: management, monitoring and precision. **Revista Ciência Agronômica**, Fortaleza, v. 51, n. 5, e20207195, 2020. DOI: <https://doi.org/10.5935/1806-6690.20200083>.

SOARES, R. M.; SCHRÖDER, E. P. **Uso de drones agrícolas no Brasil: da pesquisa à prática**. Londrina: Embrapa Soja, 2025. 83 p. (Embrapa Soja. Documentos, 474). Disponível em: <http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1175882>. Acesso em: 15 set. 2025.

SOKOLOVA, Marina; LAPALME, Guy. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, v. 45, n. 4, p. 427-437, 2009. DOI: <https://doi.org/10.1016/j.ipm.2009.03.002>.

WERNER, F.; SILVA, M. A. de A. e; FERREIRA, A. S.; NEUMAIER, N.; JUNIOR, A. A. B. Dinâmica da cobertura do solo por plantas e NDVI de cultivares de soja em diferentes arranjos espaciais de plantas. **Colloquium Agrariae**, Presidente

Prudente, v. 14, n. 2, p. 183-190, jul. 2018. Disponível em:
<https://journal.unoeste.br/index.php/ca/article/view/2103>. Acesso em: 16 set. 2025.

WU, J.; CHEN, X.-Y.; ZHANG, H.; XIONG, L.-D.; LEI, H.; DENG, S.-H.
Hyperparameter optimization for machine learning models based on Bayesian
optimization. **Journal of Electronic Science and Technology**, v. 17, n. 1, p. 26–40,
2019. DOI: <https://www.sciencedirect.com/science/article/pii/S1674862X19300047>.

YU, M.; MA, X.; GUAN, H.; LIU, M.; ZHANG, T. A recognition method of soybean leaf
diseases based on an improved deep learning model. **Frontiers in Plant Science**, v.
13, 878834, 2022. DOI: <https://doi.org/10.3389/fpls.2022.878834>.

ZHU, H.; LIN, C.; LIU, G.; WANG, D.; QIN, S.; LI, A.; XU, J.-L.; HE, Y. Intelligent
agriculture: deep learning in UAV-based remote sensing imagery for crop diseases
and pests detection. **Frontiers in Plant Science**, v. 15, 2024. DOI:
<https://doi.org/10.3389/fpls.2024.1435016>.