



Aplicação de redes neurais convolucionais para detecção de doenças em folhas de macieiras

Everton ROMANN ¹ , Rúbia Eliza de Oliveira Schultz ASCARI ^{*1} 

¹ Universidade Tecnológica Federal do Paraná, Pato Branco, PR, Brasil.
*E-mail: rubia@utfpr.edu.br

Submetido em: 28/04/2026; Aceito em: 13/07/2026; Publicado em: 13/08/2026.

RESUMO: O cultivo de maçãs tem grande importância econômica no setor agropecuário brasileiro, especialmente na região Sul do país. No entanto, a produtividade dos pomares é frequentemente comprometida por doenças foliares que, se não tratadas, podem resultar em perdas substanciais. Nesse contexto, os avanços em técnicas de Aprendizado de Máquina têm possibilitado o desenvolvimento de soluções computacionais que auxiliam no diagnóstico agropecuário com maior precisão e agilidade. Este trabalho propõe uma abordagem baseada em Redes Neurais Convolucionais que utiliza aprendizado por transferência para detectar automaticamente sintomas de doenças em folhas de macieira. A metodologia desenvolvida inclui a segmentação e análise de regiões sintomáticas para reduzir o ruído proveniente de áreas saudáveis e direcionar o aprendizado do modelo para sinais relevantes. Um total de 32.382 manchas de sintomas foram extraídas de 1.995 imagens originais, abrangendo cinco classes de distúrbios foliares: glomerela, sarna, danos por herbicidas, deficiência de magnésio e deficiência de potássio. A rede MobileNetV2, treinada por meio de aprendizado por transferência, alcançou um F1-score de 0,926 e 93,8% de acurácia no conjunto de teste reservado. Os resultados indicam um bom desempenho no contexto avaliado, sugerindo o potencial da abordagem como ferramenta de apoio ao diagnóstico da saúde das plantas e à tomada de decisões em campo.

Palavras-chave: aprendizado de máquina; visão computacional; doenças em plantas.

Application of convolutional neural networks for disease detection in apple tree leaves

ABSTRACT: Apple cultivation holds significant economic importance in the Brazilian agricultural sector, especially in the southern region of the country. However, orchard productivity is often compromised by foliar diseases, which, if left untreated, can lead to substantial losses. In this context, advances in Machine Learning techniques have enabled the development of computational solutions that support agricultural diagnostics with greater accuracy and agility. This work proposes an approach based on Convolutional Neural Networks that uses transfer learning to automatically detect disease symptoms in apple leaves. The developed methodology includes segmentation and analysis of symptomatic regions to reduce noise from healthy areas and focus the model's learning on relevant signals. A total of 32,382 symptom patches were extracted from 1,995 original images, covering five foliar disorder classes: glomerella, scab, herbicide damage, magnesium deficiency, and potassium deficiency. The MobileNetV2, trained via transfer learning, achieved a F1-score of 0.926 and 93.8% accuracy on the held-out test set. The results indicate good performance in the evaluated setting, suggesting the approach's potential as a tool to support plant-health diagnosis and field decision-making.

Keywords: machine learning; computer vision; plant disease.

1. INTRODUÇÃO

A maçã, reconhecida por seu alto valor nutricional e medicinal, é uma das frutas mais produtivas e está entre as mais consumidas no mundo (BUENO et al., 2021). No Brasil, o cultivo da macieira é uma atividade de destaque nos estados da região Sul e mantém o país competitivo no mercado externo (BUENO et al., 2021). Entre as frutas de clima temperado cultivadas no Brasil, a maçã é superada apenas pela uva, sendo a segunda mais produzida no país, com destaque para os estados do Rio Grande do Sul e de Santa Catarina como principais regiões produtoras (EPAGRI, 2022).

A produtividade das macieiras é diretamente afetada negativamente pela ocorrência de pragas e doenças, o que é um fator determinante para a qualidade e o rendimento dos pomares (LIU; WANG, 2021). Devido à necessidade de manejo contínuo, estima-se que cerca de um terço dos custos de produção seja destinado a tratamentos fitossanitários para o controle de doenças (ARAUJO et al., 2016). A realização de tratamento adequado auxilia na prevenção da propagação de doenças, melhorando a produtividade agrícola e reduzindo as perdas, mas o monitoramento visual de doenças é demorado e dispendioso (Al-Hiary et al., 2011); portanto, são necessários métodos alternativos.

Nesse contexto, a detecção automática de doenças tornou-se uma área de pesquisa amplamente investigada no domínio da informação agrícola. A Visão Computacional (VC), um campo da Inteligência Artificial (IA) dedicado a extrair informações úteis a partir de dados visuais (Ballard; Brown, 1982), vem sendo empregada nesse contexto e mostra-se uma solução promissora. Por meio do treinamento de algoritmos, a IA e o Aprendizado de Máquina (AM) podem auxiliar na classificação de plantas com base em características desejáveis, na detecção de doenças e de estresses ambientais e na identificação de cruzamentos promissores (RESENDE, 2023).

Este trabalho apresenta um classificador de distúrbios foliares em macieiras, que emprega uma abordagem voltada à identificação automática de regiões que apresentam exclusivamente sintomas de doenças. Ao isolar as áreas sintomáticas, busca-se realizar uma análise mais direcionada dos padrões visuais patológicos por meio de técnicas de processamento de imagens. Contudo, delimita-se que este estudo está estritamente restrito a um cenário controlado em estúdio, sem validação em ambiente real de campo, o que implica em um modelo otimizado para condições padronizadas de fundo e iluminação. Adicionalmente, destaca-se, de forma crítica, que o escopo atual não contempla uma classe de controle para folhas saudáveis; consequentemente, o modelo proposto é adequado para operar com amostras de folhas previamente triadas como sintomáticas. Diante disso, uma Rede Neural Convolucional leve foi construída e treinada com base nessas restrições, visando ao desenvolvimento e à avaliação de um classificador focado na diferenciação de cinco anomalias específicas em condições controladas.

2. DOENÇAS E DISTÚRBIOS FOLIARES DA MACIEIRA

As condições foliares que acometem a macieira constituem importantes fatores limitantes da produção e da qualidade dos frutos, afetando diretamente o rendimento dos pomares (FAORO, 2022). O comprometimento do tecido foliar reduz a superfície fotossintética ativa, impactando processos fisiológicos essenciais, como a assimilação de carbono, a transpiração e o transporte de fotoassimilados, o que pode resultar em menor crescimento vegetativo, atraso no desenvolvimento dos frutos e redução do potencial produtivo (VELHO et al. 2019). Essas condições manifestam-se por sintomas visuais distintos, como alterações de cor, textura e formato das lesões, que refletem o agente causal e o estágio do dano, constituindo um critério relevante para o diagnóstico em campo (FAORO, 2022). Neste trabalho, foram consideradas cinco condições foliares representativas da cultura da macieira: duas doenças fúngicas, Mancha Foliar de Glomerella e Sarna da Macieira, e três distúrbios não infecciosos — Dano por Herbicida, Deficiência de Magnésio e Deficiência de Potássio, cujos sintomas decorrem de fatores de manejo e desequilíbrios nutricionais, brevemente descritos a seguir:

- Mancha Foliar de Glomerella: Doença fúngica altamente agressiva em regiões subtropicais e úmidas, causada por *Colletotrichum* (complexo *C. gloeosporioides*), com registros no Brasil desde 1988 (VELHO et al., 2019). Afeta folhas e frutos, sobretudo na primavera e no verão. Nas folhas, inicia-se com pequenas manchas avermelhadas que evoluem para lesões necróticas com clorose, podendo

coalescer e provocar intensa desfolha sob condições favoráveis (ARAUJO et al., 2016).

- Sarna da Macieira: Doença recorrente em regiões de clima temperado e úmido, favorecida pela alta umidade e por temperaturas amenas (FAORO, 2022). Causada por *Venturia inaequalis*, infecta diferentes órgãos aéreos. Nas folhas, provoca lesões verde-oliva que evoluem para um aspecto acinzentado e aveludado, podendo reduzir a área fotossintética em estágios avançados.
- Dano por Herbicida: Resulta da aplicação inadequada ou da exposição acidental a herbicidas utilizados no manejo. Os sintomas, mais evidentes em folhas jovens, incluem clorose internerval, necrose marginal, enrolamento e queda foliar (GABARDO et al, 2017).
- Deficiência de Magnésio: Distúrbio nutricional frequente em solos ácidos ou com desequilíbrios de potássio e cálcio (NACHTIGALL; CARGNINO, 2012). Manifesta-se inicialmente nas folhas mais velhas, com clorose internerval em “V”, evoluindo para necrose marginal e queda prematura.
- Deficiência de Potássio: Comum em solos arenosos ou pobres em matéria orgânica. Os sintomas iniciam-se nas folhas mais velhas, com clorose marginal, seguidas de necrose e enrolamento, podendo comprometer o crescimento, o tamanho dos frutos e a produtividade (SALLATO et al., 2019).

3. VISÃO COMPUTACIONAL

A definição de Visão Computacional evoluiu progressivamente ao longo das últimas décadas, ampliando seu escopo de atuação. Inicialmente delimitada como a automação de funções visuais humanas em sistemas computacionais (BALLARD; BROWN, 1982), a área passou a incorporar a extração de informações relevantes a partir de mídias digitais para subsidiar a tomada de decisão automatizada (STOCKMAN; SHAPIRO, 2001). Adicionalmente, consolidou-se como o arcabouço metodológico voltado à interpretação de imagens e à reconstrução tridimensional a partir de projeções bidimensionais (FORSYTH; PONCE, 2012). Apesar dessa transição metodológica, o propósito fundamental da VC permanece: capacitar sistemas tecnológicos a perceber, analisar e extrair compreensão de cenas complexas a partir de dados visuais (SZELISKI, 2022).

O avanço contínuo da VC tem ampliado a complexidade de suas abordagens, com a adoção de técnicas cada vez mais sofisticadas, voltadas à resolução de problemas mais desafiadores. No entanto, independentemente das técnicas empregadas, sua estrutura de funcionamento ainda se organiza em torno de etapas fundamentais, que abrangem desde a aquisição de dados visuais até a interpretação e a tomada de decisão com base nas informações extraídas. Conforme Szeliski (2022), esse fluxo compreende etapas que vão desde operações básicas de melhoria da qualidade da imagem até procedimentos mais sofisticados, como a segmentação de Regiões de Interesse (ROIs) e a detecção de padrões específicos.

A partir dessa perspectiva, a análise automática do conteúdo visual passa a depender de modelos que aprendem regularidades diretamente a partir dos dados, o que torna o Aprendizado de Máquina um elemento central na VC contemporânea (SZELISKI, 2022; GOODFELLOW et al., 2016).

4. APRENDIZADO DE MÁQUINA

No âmbito da evolução da Visão Computacional, o Aprendizado de Máquina consolidou-se como um pilar da Inteligência Artificial ao viabilizar a extração de padrões em cenários de alta densidade de dados, caracterizando-se pela capacidade do sistema de otimizar seu desempenho em uma tarefa por meio da experiência, sem necessidade de reprogramação manual (MITCHELL, 1997). Suas metodologias dividem-se em três paradigmas fundamentais (Goodfellow et al., 2016): o Aprendizado Supervisionado, que emprega dados rotulados para treinar modelos de classificação ou regressão; o Aprendizado Não Supervisionado, voltado à identificação de estruturas ocultas em dados não rotulados por meio de agrupamento (*clustering*) e redução de dimensionalidade (Hastie et al., 2009); e o Aprendizado por Reforço, baseado na interação entre agente e ambiente para otimizar decisões via recompensas e penalidades (SUTTON; BARTO, 2018).

A aplicação direta em tarefas de análise de imagens ainda enfrenta desafios devido à complexa estrutura espacial e hierárquica dos *pixels*, que não é adequadamente explorada por algoritmos tradicionais baseados em vetores planos (GOODFELLOW et al., 2016). Além disso, o elevado volume de dados visuais aumenta a complexidade computacional, o que limita o uso de métodos convencionais. Nesse contexto, destacam-se as *Convolutional Neural Networks* (CNNs), desenvolvidas para explorar de forma eficiente as características espaciais e extrair padrões relevantes a partir de imagens digitais.

O avanço das CNNs tem impulsionado significativamente a detecção automática de plantas, possibilitando altos índices de acurácia na identificação de doenças em diferentes culturas. Estudos recentes relatam desempenhos superiores a 99% na classificação de doenças em múltiplas espécies, evidenciando a eficácia das redes neurais profundas como ferramenta de apoio ao diagnóstico agrícola (KAMILARIS; BOLDÚ et al., 2018). Entretanto, persistem desafios práticos, como a variabilidade visual dos sintomas, a semelhança entre diferentes doenças ou estágios de infecção e a presença de ruídos nas imagens, fatores que podem comprometer o desempenho dos modelos (Barbedo, 2018).

5. MÉTODO

A proposta metodológica desenvolvida neste trabalho fundamenta-se nos conceitos de VC, aliados ao Aprendizado Profundo, e emprega uma Rede Neural Artificial (RNA) para a classificação de doenças e anomalias em folhas de macieira. O fluxo geral de processamento segue etapas consolidadas na literatura para o diagnóstico vegetal com imagens digitais, conforme descrito por (PATEL; JOSHI, 2017). Neste trabalho, a análise concentra-se em regiões sintomáticas, identificadas automaticamente na lâmina foliar, para alimentar o classificador final, com o objetivo de reduzir as interferências do fundo e de áreas saudáveis. As subseções a seguir detalham os procedimentos adotados.

5.1. Base de Imagens

O conjunto de dados inicial utilizado neste trabalho foi desenvolvido por Nachtigall (2016) e é composto por imagens de folhas de macieira coletadas em pomares comerciais de Vacaria, no Rio Grande do Sul. O material, que reflete diferentes estágios fenológicos e manejos, foi digitalizado individualmente sobre fundo branco, sob

iluminação artificial e com distância focal fixa, gerando arquivos com resolução uniforme de 4000×2248 *pixels* em orientações variadas. A rotulagem taxonômica e diagnóstica foi discriminada pela injúria predominante e validada por engenheiros agrônomos em cinco classes (Glomerella, Sarna, dano por herbicida, deficiência de magnésio e deficiência de potássio), cujos quantitativos amostrais e exemplares representativos encontram-se sistematizados na Tabela 1 e na Figura 1, respectivamente.

Tabela 1. Quantidade de amostras de folhas de macieira disponíveis no conjunto de dados original para cada sintoma.

Table 1. Number of apple leaf samples available in the original dataset for each symptom.

Classe	Amostras
Mancha da Glomerella	581
Dano por Herbicida	325
Deficiência de Magnésio	355
Deficiência de Potássio	341
Sarna da Macieira	393

Adaptado de Nachtigall (2016).

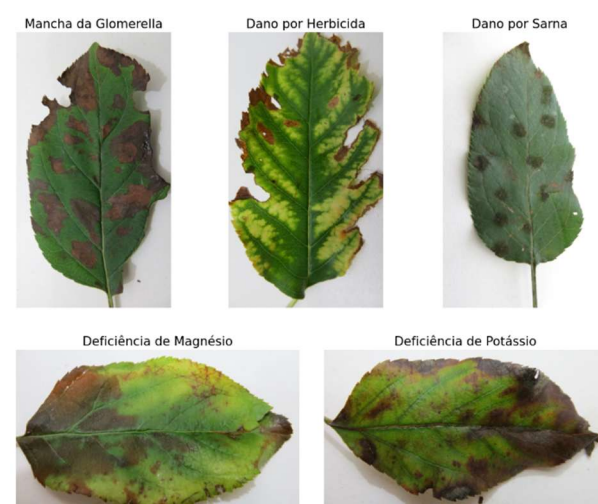


Figura 1. Amostras de folhas de macieira com cinco sintomas distintos, disponíveis no conjunto de imagens utilizado como base deste estudo.

Figure 1. Apple leaf samples showing symptoms of five different symptoms, available in the image set used as the basis for this study.

O conjunto original apresenta um desequilíbrio entre as classes. A fim de mitigar esse problema, foram aplicadas estratégias de compensação por meio de ponderação de classes na função de perda, detalhadas adiante.

5.2. Separação da folha do fundo

Com o objetivo de reduzir interferências do fundo, foi incorporada uma etapa de segmentação da lâmina foliar, motivada pelo desempenho insatisfatório na segmentação direta dos sintomas, especialmente devido a ruídos nas bordas das folhas. Abordagens baseadas em limiarização no espaço *Hue*, *Saturation*, *Value* (HSV) mostraram-se insuficientes para garantir a precisão necessária.

Como alternativa, adotou-se o *Segment Anything Model* (SAM) (Kirillov et al., 2023), uma arquitetura de segmentação com elevada capacidade de generalização e precisão na delimitação de objetos sob variações de cor e forma, já validada em aplicações agrícolas (ZHANG et al., 2024). O processo foi estruturado de forma totalmente automática, com a geração de *prompts* compostos por caixas delimitadoras

e por pontos positivos e negativos, fornecendo informações suficientes para a definição adequada da região foliar.

As máscaras geradas foram avaliadas com base em critérios morfológicos relacionados à área, à centralidade, à regularidade da forma, ao contato com as bordas e à similaridade com mapas de contraste. Quando necessário, pontos adicionais eram inseridos para refinamento da segmentação. A máscara final passou por pós-processamento morfológico e pela remoção do pecíolo, resultando em ROIs mais consistentes para as análises subsequentes. Os exemplos apresentados evidenciam que o método delimita adequadamente a lâmina foliar, suprimindo elementos externos e preservando a ROI.

Para verificar a qualidade das máscaras geradas pelo SAM, foi realizada uma validação quantitativa com base na métrica *Intersection Over Union* (IoU), amplamente utilizada para medir o grau de sobreposição entre regiões segmentadas (ZHAO et al., 2025). Nesse caso, a IoU foi calculada entre a máscara estimada pelo SAM e uma máscara de referência manualmente segmentada, denominada *Ground Truth* (GT). Essa métrica é definida como a razão entre a área de interseção e a área da união.

A área de sobreposição corresponde aos *pixels* coincidentes entre as máscaras comparadas, enquanto a união representa o total de *pixels* classificados como folha em pelo menos uma delas. A razão entre essas áreas, expressa pela IoU, varia de 0 a 1 e quantifica o grau de concordância espacial; valores elevados indicam segmentação precisa (ZHAO et al., 2025). A GT foi construída com o *Label Studio*, a partir da anotação manual de 125 amostras (25 por classe), selecionadas para contemplar a variabilidade do conjunto quanto à orientação, centralização, iluminação, formato e intensidade dos sintomas. As anotações delimitaram exclusivamente a lâmina foliar, excluindo o pecíolo e o fundo.

5.3. Segmentação de regiões sintomáticas

Com as ROIs definidas pela segmentação foliar obtida na etapa anterior, essa fase concentrou-se na identificação automática de regiões sintomáticas no interior da lâmina. A hipótese central é que tais sintomas, ao alterarem a coloração do tecido afetado, tornam-se visíveis por meio da análise cromática da imagem.

A segmentação dos sintomas baseou-se na conversão da imagem para o espaço HSV, cuja separação entre matiz (H), saturação (S) e valor (V) favorece a distinção de alterações cromáticas associadas a doenças foliares. Para cada classe, foram definidas faixas específicas desses canais, determinadas empiricamente a partir de inspeções visuais sistemáticas em amostras reais, conforme apresentado na Tabela 2. Essas combinações foram ajustadas manualmente com o objetivo de maximizar a cobertura sintomática, ao mesmo tempo em que se minimiza a inclusão de áreas saudáveis.

Após a segmentação inicial dos sintomas, foi realizada uma etapa de refinamento composta por duas operações complementares. Primeiramente, aplicou-se o fechamento morfológico para suavizar os contornos e preencher as descontinuidades nas regiões detectadas. O elemento estruturante adotado foi um disco cujo raio era ajustado dinamicamente em função do tamanho da folha, assegurando proporcionalidade à escala da imagem.

Em seguida, empregou-se um filtro de área mínima para remover componentes conexos de pequenas dimensões, frequentemente associados a ruídos cromáticos ou a irregularidades na textura foliar. O limiar foi definido de

forma adaptativa como 0,25% da área total da folha segmentada - o que corresponde, por exemplo, a 1.250 *pixels* em uma folha com 500.000 *pixels*. Esse valor foi estabelecido empiricamente, de modo a equilibrar a sensibilidade na detecção de sintomas reais em folhas menores e a supressão de artefatos irrelevantes em folhas de maior dimensão.

Tabela 2. Intervalos de HSV utilizados para segmentação das amostras de cada classe.

Table 2. HSV ranges used for the segmentation of samples from each class.

Classe	Sintoma / Tom alvo	H	S	V
Mancha da Glomerella	Tons terrosos / Marrom	10 a 35	40 a 255	30 a 150
Dano por Herbicida	Manchas escuras / Saturação média	0 a 20	60 a 160	20 a 110
Deficiência de Magnésio	Clorose amarelada / Necrose	15 a 50	0 a 210	105 a 255
Deficiência de Potássio	Amarelecimento forte	20 a 35	100 a 255	100 a 255
Sarna da Macieira	Lesões escuras, secas e saturadas	5 a 25	120 a 255	10 a 120

5.4. Extração das áreas sintomáticas

A partir das regiões sintomáticas segmentadas, iniciou-se a construção do conjunto de dados de sintomas, destinado a compor a entrada da rede neural convolucional. A estratégia consistiu na conversão das áreas afetadas em unidades menores que preservassem características visuais discriminativas, por meio da extração de recortes quadrados representativos das lesões.

Os recortes foram definidos com dimensões fixas de 224×224 *pixels*, compatíveis com a entrada padrão de CNNs. Para assegurar a relevância das amostras, estabeleceu-se como critério mínimo que ao menos 25% da área de cada janela contivesse *pixels* classificados como sintomáticos, evitando a inclusão predominante de regiões saudáveis. Adicionalmente, implementou-se um controle de sobreposição para impedir a extração repetida da mesma área da folha, bem como um limite máximo de tentativas por imagem, calculado como um múltiplo fixo do número estimado de recortes, o que garante eficiência computacional.

A quantidade de recortes por imagem foi definida proporcionalmente à área útil da lâmina segmentada. Com base em testes preliminares, estabeleceu-se como limite superior uma cobertura de até 65% da área foliar, considerada suficiente para assegurar representatividade sem introduzir redundância excessiva. Ressalta-se que esse percentual constitui um teto teórico, atingido apenas em folhas amplamente acometidas, e não representa a condição predominante do conjunto analisado. No que tange à generalização, esses limiares atuam como filtros de qualidade: ao garantir que cada recorte contenha informação patológica densa (mínimo de 25%) e ao impedir que folhas individuais monopolizem o conjunto de dados com múltiplos recortes idênticos (máximo de 65%), o modelo é induzido a extrair características genéricas das lesões em vez de memorizar espécimes específicos, elevando sua robustez em cenários reais de campo. Quanto ao desbalanceamento entre classes, o teto de 65% afeta o conjunto de dados de forma assimétrica, limitando a geração de amostras em folhas amplamente acometidas - condição comum em doenças fúngicas mais agressivas (como a Mancha Foliar de Glomerella) e rara em anomalias de evolução lenta. Essa

restrição física, embora necessária para evitar o *overfitting*, acentua o desequilíbrio quantitativo entre as classes no conjunto final, o qual foi devidamente tratado e mitigado nas etapas subsequentes do *pipeline* de Aprendizado de Máquina para assegurar a equidade e a estabilidade do treinamento.

A Figura 2 sintetiza as etapas centrais do processo, apresentando, em sequência, a imagem original, a segmentação da lâmina foliar, a segmentação das regiões sintomáticas e o posicionamento das janelas de recorte. Essa última etapa explicita o critério adotado para a seleção das janelas de 224×224 pixels no interior da região previamente segmentada.



Figura 2. Etapas de segmentação e seleção das regiões de interesse empregadas nas amostras. À esquerda, a amostra original; em seguida, o resultado da segmentação da folha; posteriormente, a representação da segmentação dos sintomas; e, por fim, os limites das áreas sintomáticas a serem extraídas, destacados em vermelho. Figure 2. Stages of segmentation and selection of regions of interest applied to the samples. On the left is an original sample, followed by the result of the leaf segmentation, then the representation of the symptom segmentation, and finally, the boundaries of the symptomatic areas to be extracted, highlighted in red.

5.5. Preparação do conjunto de dados e pré-processamentos

Como etapa inicial, definiu-se uma semente aleatória única, sincronizando os geradores de números aleatórios do Python, do NumPy e do PyTorch. Dessa forma, operações estocásticas do treinamento - como a inicialização de pesos e a ordem de apresentação dos exemplos - passaram a seguir o mesmo padrão em cada execução, garantindo a comparabilidade dos resultados e a previsibilidade analítica.

Outro aspecto metodológico tratado foi o *data leakage*, caracterizado pela influência indevida de dados de validação ou de teste no treinamento, o que inflaciona artificialmente as métricas de desempenho (KAPOOR; NARAYANAN, 2023). Como o conjunto é composto por múltiplos recortes provenientes da mesma folha, uma divisão aleatória poderia distribuir amostras correlatas entre as diferentes partições. Para mitigar esse risco, utilizou-se o identificador *leaf_id* presente nos nomes dos arquivos, registrando caminho, classe e *leaf_id* em um manifesto tabular e aplicando particionamento por grupos. Assim, todos os recortes de uma mesma folha foram mantidos em uma única partição.

Com reprodutibilidade e agrupamento assegurados, definiu-se a divisão do conjunto de 32.382 recortes sintomáticos. Adotou-se o método de validação cruzada *k-fold* (5 folds), e o método *Holdout*, com 70% dos dados para treinamento (22.655 amostras), 15% para validação (4.829) e 15% para teste (4.898). Em ambas as etapas, utilizou-se o identificador *leaf_id* para manter os recortes de uma mesma folha em uma única partição. Essa configuração preserva exemplos inéditos para avaliação imparcial, equilibrando o volume de dados de aprendizado com as amostras reservadas para validação e teste. Garantiu-se compatibilidade com a

arquitetura MobileNetV2, cujos modelos pré-treinados requerem tensores RGB de 224×224 pixels. Os recortes já foram extraídos nesse formato e não requerem redimensionamento adicional.

Durante o carregamento das imagens de treinamento, aplicaram-se técnicas de *data augmentation*, incluindo inversões horizontais e verticais, rotações de até 20° e ajustes moderados de brilho, contraste e saturação, conforme estratégias consolidadas na literatura (SHORTEN; KHOSHGOFTAAR, 2019). Essas transformações geométricas e fotométricas foram aplicadas dinamicamente a cada época, ampliando a variabilidade do conjunto e reduzindo o risco de *overfitting*. As partições de validação e teste permaneceram inalteradas, sendo submetidas apenas à conversão para tensor e à normalização dos canais de cor.

5.6. Treinamento da Rede Neural

Concluída a preparação das entradas, procedeu-se à configuração e ao treinamento da CNN responsável pela classificação dos sintomas. Para este trabalho, foi escolhido o modelo MobileNetV2, proposto por (Sandler et al., 2018), por se tratar de uma arquitetura leve e eficiente, mesmo em dispositivos com restrições computacionais.

O modelo foi inicializado com pesos previamente treinados no conjunto ImageNet, uma base com mais de um milhão de imagens rotuladas em mil categorias, amplamente utilizada em redes neurais convolucionais. Essa estratégia caracteriza o uso de *transfer learning*, técnica que consiste em reutilizar representações visuais aprendidas em tarefas anteriores, reduzindo a necessidade de grandes volumes de dados rotulados e acelerando a convergência em novos problemas. Isso permitiu que a MobileNetV2 se adaptasse ao domínio específico das folhas de macieira sem partir do zero, aproveitando filtros já otimizados para extrair padrões genéricos como contornos, texturas e formas. O uso de *transfer learning* é especialmente eficaz quando se busca desempenho competitivo com baixo custo de treinamento e dados moderados, como ocorre neste trabalho.

Como consequência direta do uso de *transfer learning*, tornou-se necessário preparar a rede para lidar com as classes específicas deste trabalho sem comprometer os filtros convolucionais ajustados no ImageNet. Para isso, aplicou-se uma etapa inicial de *warm-up*, com cinco épocas em que apenas a camada de saída é treinada, enquanto o restante da arquitetura permanece congelado. Essa etapa funciona como um mecanismo de adaptação gradual, permitindo que a nova camada aprenda a interpretar adequadamente as representações herdadas antes da atualização completa da rede, reduzindo instabilidades e o risco de sobreajuste. Dessa forma, evitam-se perturbações nas camadas profundas nos primeiros ciclos de aprendizado, preparando a estrutura do modelo para a fase posterior de especialização.

O *fine-tuning*, que corresponde à etapa principal de treinamento da rede, foi iniciado após a conclusão do *warm-up*. Nessa fase, todas as camadas da MobileNetV2 foram destravadas, permitindo que o modelo passasse a ajustar não apenas a camada de saída, mas também os filtros convolucionais previamente aprendidos no ImageNet. Nessa etapa, as representações internas da rede começam, de fato, a se especializar na extração de características visuais associadas ao recorte dos sintomas.

Também nessa etapa, foi necessário reduzir o efeito do desbalanceamento entre as classes e, para isso, empregou-se *CrossEntropyLoss* com o parâmetro *weight*, atribuindo peso

maior às categorias menos frequentes e menor às mais abundantes. O cálculo seguiu o método padrão da literatura de aprendizagem profunda, expresso pela equação $w_c = \frac{1}{N_c}$, em que w_c representa o peso atribuído à classe c e N_c corresponde à quantidade total de amostras disponíveis para aquela categoria específica no lote de treino. Esses pesos foram calculados a partir das contagens de amostras por classe no conjunto de treino e mantidos fixos ao longo de todo o processo. As partições de validação e de teste permaneceram inalteradas, sem reamostragem nem ponderação nessas etapas.

O ciclo de treinamento foi configurado para até 50 épocas, valor definido com base em observações práticas durante o desenvolvimento. Para evitar *overfitting*, aplicou-se a técnica de *early stopping*, interrompendo o processo caso a métrica de *F1-score* no conjunto de validação não apresentasse melhora após seis épocas consecutivas.

Nas duas fases descritas, os hiperparâmetros foram ajustados para promover a estabilidade, controlar o sobreajuste e garantir a compatibilidade com a arquitetura utilizada. Durante o *warm-up*, a taxa de aprendizado foi fixada em 0,001, valor suficientemente alto para permitir ajustes rápidos na camada de saída, ainda isolada do restante da rede. Já no *fine-tuning*, esse valor foi reduzido para 0,0003, com o intuito de preservar as representações aprendidas no ImageNet e evitar atualizações agressivas nos filtros convolucionais mais profundos, conforme sugerido por (SMITH, 2017).

Para refinar ainda mais o processo de otimização ao longo das épocas, foi empregado o agendador *ReduceLROnPlateau*, que monitora a perda de validação e reduz a taxa de aprendizado em 50% quando essa métrica deixa de melhorar após três épocas consecutivas, com um tempo de espera de uma época e valor mínimo de 0,000001. Essa abordagem, inspirada nas diretrizes propostas por (Smith, 2017), permite um refinamento mais suave e eficiente, adaptando dinamicamente o ritmo de aprendizagem do modelo.

O otimizador utilizado foi o AdamW, com *weight decay* de 0,0001, o que ajuda a conter o crescimento excessivo dos pesos e favorece a generalização. Por fim, o *batch size* foi fixado em 64, um valor compatível com a capacidade de processamento disponível.

5.7. Medidas de avaliação de desempenho

A fase final do trabalho consistiu na avaliação do modelo desenvolvido, por meio de validação cruzada *k-fold* sobre um conjunto formado pela concatenação dos dados de treino e de validação, e de avaliação final no conjunto de teste, reservado e mantido isolado desde a definição das partições.

Para a definição das métricas, estabelecem-se inicialmente quatro parâmetros fundamentais: Verdadeiro Positivo (VP), quando o modelo classifica corretamente uma amostra da classe de interesse; Falso Positivo (FP), quando atribui essa classe a uma amostra que não lhe pertence; Falso Negativo (FN), quando deixa de reconhecer uma amostra da classe de interesse; e Verdadeiro Negativo (VN), quando identifica corretamente que a amostra não pertence a essa classe.

Com base nesses elementos, constrói-se a matriz de confusão (exemplificada na Figura 3), organizada em formato 2×2, na qual as linhas representam os rótulos reais e as colunas as classes previstas. Os acertos situam-se na diagonal principal (VP e VN), enquanto os erros (FP e FN) ocupam a

diagonal secundária, o que permite identificar claramente os tipos de acerto e de erro do modelo.

		Valor Predito	
		Positivo	Negativo
Valor Real	Positivo	VP	FN
	Negativo	FP	VN

Figura 3. Exemplo ilustrativo de uma matriz de confusão 2×2. A diagonal principal, na cor verde, apresenta os acertos do modelo, e a diagonal secundária, na cor rosa, apresenta os erros.

Figure 3. An illustrative example of a 2×2 confusion matrix. The main diagonal, in green, shows the model's correct predictions, and the secondary diagonal, in pink, shows the errors.

Adaptado de Sokolova; Lapalme (2009).

A partir dessa estrutura, derivam-se as principais medidas de desempenho utilizadas neste trabalho, conforme descrito a seguir.

- *Precision*: mede a proporção de amostras classificadas como positivas que, de fato, são corretas. Dada pela Equação 1, avalia a confiabilidade das previsões positivas do modelo.

$$\text{Precision} = \frac{\text{VP}}{\text{VP} + \text{FP}} \quad (01)$$

- *Recall*: representa a proporção de amostras positivas reais corretamente identificadas pelo modelo (Equação 2).

$$\text{Recall} = \frac{\text{VP}}{\text{VP} + \text{FN}} \quad (02)$$

- *F1-score*: corresponde à média harmônica entre precisão e *recall*, como exibida na Equação 3, sendo uma medida que equilibra ambas em um único valor, tornando-a especialmente útil quando há desbalanceamento entre as classes.

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (03)$$

- Acurácia: mede a proporção de previsões corretas (positivas e negativas) em relação ao total de amostras avaliadas. Exibida na Equação 4, é uma medida global de desempenho, embora sensível à distribuição das classes.

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{FP} + \text{FN} + \text{VN}} \quad (04)$$

6. RESULTADOS

Os resultados correspondem às saídas finais do fluxo de processamento proposto, após a segmentação das folhas, a extração das regiões sintomáticas e a classificação das amostras pela CNN. As análises a seguir abrangem a composição do conjunto gerado, o desempenho do modelo no conjunto de teste e os principais padrões observados nas métricas, nas curvas de treinamento e na matriz de confusão.

6.1. Composição do conjunto gerado

Os resultados da etapa de separação da folha do fundo são apresentados na Figura 4 e Tabela 3. Foi obtido uma IoU média global de 0,949, com médias macro por classe entre 0,934 e 0,975 e baixos desvios-padrão, indicando alta concordância entre as máscaras automáticas e as anotações

manuais. Esses valores confirmam a consistência do método e a adequação das segmentações às etapas subsequentes.

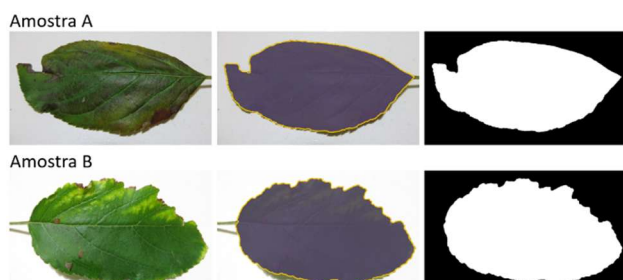


Figura 4. Resultado da segmentação de folhas de macieira realizada com o algoritmo SAM. Em cada linha, à esquerda está a imagem original. Ao centro, a sobreposição da máscara estimada, com a área segmentada em roxo e o contorno gerado pelo SAM destacado em amarelo. À direita, a respectiva máscara binária.

Figure 4. Result of the segmentation performed on apple leaves, using the SAM algorithm. In each row, the original image is on the left. In the center is the overlay of the estimated mask, with the segmented area in purple and the SAM-generated contour highlighted in yellow. On the right is the corresponding binary mask.

Tabela 3. Resultado da avaliação, via IoU, da folha segmentada, calculada entre a máscara estimada pelo SAM e a máscara de referência manualmente segmentada.

Table 3. Result of the IoU-based evaluation of the segmented leaf, calculated between the mask estimated by SAM and the manually segmented reference mask.

Classe	Amostras	IoU média	Desvio padrão
Mancha da Glomerella	25	0,941	0,041
Dano por Herbicida	25	0,945	0,092
Deficiência de Magnésio	25	0,953	0,086
Deficiência de Potássio	25	0,975	0,011
Sarna da Macieira	25	0,934	0,012
Média macro	125	0,949	

A Figura 5 exemplifica a combinação entre os canais para identificar as ROIs: (a) amostra de folha com deficiência de magnésio; (b) a máscara correspondente da folha gerada pelo SAM que restringe as demais operações a esta área; (c) a extração da clorose indicada pelo amarelamento da folha; (d) detecção das áreas necróticas ajustadas para realçar as tonalidades marrons características da deficiência; (e) a máscara final resultante da combinação das segmentações de (c) e (d).

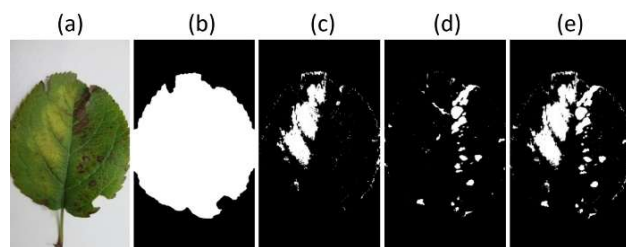


Figura 5. Combinação de canais HSV para identificar as ROIs. (a) folha com deficiência de magnésio; (b) máscara correspondente gerada pelo SAM; (c) extração da clorose, indicada pelo amarelamento da folha; (d) detecção das áreas necróticas, ajustada para realçar as tonalidades marrons; (e) máscara final resultante da combinação das segmentações de (c) e (d).

Figure 5. Combination of HSV channels to identify the ROIs. (a) leaf with magnesium deficiency; (b) corresponding mask generated by SAM; (c) extraction of chlorosis indicated by leaf yellowing; (d) detection of necrotic areas adjusted to highlight the brown tones;

(e) final mask resulting from the combination of the segmentations in (c) and (d).

A partir das regiões sintomáticas segmentadas, procedeu-se à etapa de extração de sintomas. Foram usadas janelas de 224×224 pixels, resultando em 32.382 recortes, distribuídos por classe conforme a Tabela 4. Observa-se um aumento expressivo no número de recortes na classe Sarna da Macieira, que também apresenta a maior média por imagem, enquanto a classe Deficiência de Magnésio registra a menor média. Esse padrão pode ser explicado tanto pela morfologia dos sintomas quanto pelos critérios adotados na segmentação. A Figura 6 ilustra exemplos de sintomas extraídos de cada uma das cinco classes analisadas.

Tabela 4. Distribuição dos recortes sintomáticos obtidos a partir das amostras disponíveis no conjunto de dados original, por classe.

Table 4. Distribution of symptomatic patches obtained from the samples available in the original dataset, separated by class.

Classe	Amostras	Recortes	Média por imagem
Mancha da Glomerella	581	9245	15,91
Dano por Herbicida	325	4507	13,87
Deficiência de Magnésio	355	3300	9,30
Deficiência de Potássio	341	5428	15,92
Sarna da Macieira	393	9902	25,20
Média macro	1995	32382	16,23

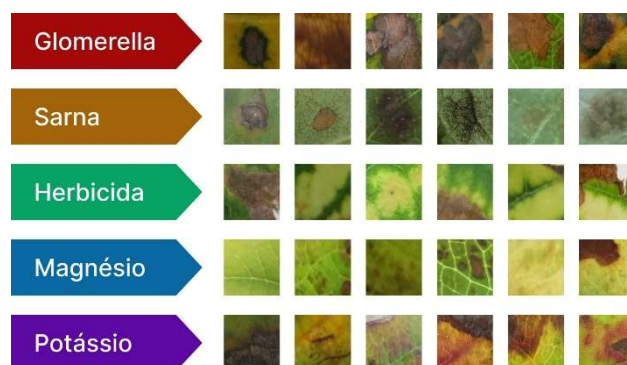


Figura 6. Exemplos de amostras do conjunto de dados de sintomas resultante do processo de segmentação de ROI a partir das amostras originais. Em cada linha há seis amostras diferentes da classe descrita à esquerda.

Figure 6. Example samples from the symptom dataset resulting from the ROI segmentation process applied to the original samples. Each row contains six different samples of the class described on the left.

Na Sarna da Macieira, as folhas tendem a ser maiores e apresentam lesões mais numerosas e dispersas, o que aumenta o número de janelas que atingem o limiar mínimo de ocupação sintomática e são mantidas após os filtros, resultando em maior quantidade de recortes por imagem. Em contraste, na Deficiência de Magnésio, as folhas são menores e a clorose ocorre de forma mais localizada, reduzindo a proporção de janelas que superam o limiar e, consequentemente, o número de recortes. Nas demais classes, observa-se um comportamento intermediário, decorrente do equilíbrio entre extensão das áreas sintomáticas e quantidade de componentes preservados pelos filtros. A validação da qualidade da segmentação cromática procedeu-se de forma indireta, fundamentando-se no impacto dos recortes gerados sobre o desempenho do

classificador final, a rede MobileNetV2. Adotou-se como critério de sucesso a premissa de que a geração automática de recortes imprecisos ou que contenham excesso de tecido saudável confundiria a rede neural, resultando em métricas de classificação inferiores.

6.2. Desempenho e validação da CNN

Utilizando a MobileNetV2 pré-treinada, o classificador foi ajustado no conjunto de treino, com monitoramento no conjunto de validação, mantendo-se um conjunto de teste reservado para a avaliação final. A Tabela 5 apresenta a distribuição das amostras utilizadas por classe em cada partição: treino, validação e teste.

Tabela 5. Distribuição das amostras utilizadas por classe em cada partição nas etapas de treino, avaliação e teste.

Classe	Treino	Validação	Teste	Total
Mancha da Glomerella	6415	1367	1463	9245
Dano por Herbicida	3198	682	627	4507
Deficiência de Magnésio	2292	489	519	3300
Deficiência de Potássio	3777	805	846	5428
Sarna da Macieira	6973	1486	1443	9902
Total	22655	4829	4898	32382

As Figuras 7 e 8 mostram, respectivamente, a evolução da acurácia e da *loss* ao longo das épocas de treino e de validação. A linha tracejada marca o início do *fine-tuning*. O ponto destacado na Figura 7 indica a época correspondente ao melhor modelo, selecionado pelo critério de *early stopping*, com base no máximo *F1-macro* no conjunto de validação. O *checkpoint* selecionado pelo *early stopping* foi obtido na época 41. Esse modelo, avaliado no conjunto de teste, atingiu um *F1-score* de 0,926 e uma acurácia global de 0,938. A classe Sarna da Macieira apresentou o melhor desempenho individual, com *F1-score* de 0,9705, enquanto a classe Deficiência de Magnésio teve o menor, com 0,8868. A média macro das medidas de *precision*, *recall* e *F1-score* superaram 92,5%, demonstrando a boa capacidade do modelo em generalizar para dados não vistos, mesmo diante do desbalanceamento entre as classes. A Tabela 6 detalha os valores obtidos para cada medida em cada classe, bem como a quantidade de amostras de cada classe e as médias de cada medida, os quais resumem o desempenho geral do modelo.

Para complementar a análise, as matrizes de confusão das predições no conjunto de teste são apresentadas na Figura 9(a), que mostra a contagem absoluta de classificações, enquanto a Figura 9(b) apresenta a versão normalizada por linha, indicando a proporção de acertos e erros em relação ao total de amostras.

Como complemento à avaliação final realizada no conjunto de teste, aplicou-se uma validação cruzada *k-fold* com $k = 5$ sobre um conjunto formado pela concatenação dos dados de treino e validação, mantendo o conjunto de teste completamente isolado, a fim de estimar a estabilidade do modelo diante de diferentes partições dos dados. Nesse processo, o conjunto de dados é dividido em *k* subconjuntos de tamanhos semelhantes. Em cada iteração, um desses subconjuntos é reservado para validação, enquanto os demais compõem o conjunto de treino. O processo é repetido *k* vezes, de modo que cada partição atue uma única vez como

validação. A Tabela 7 apresenta as métricas obtidas por partição, bem como os valores médios e os respectivos desvios padrão. O *F1-score* apresentou média de 0,918, desvio padrão de 0,009 e variou de 0,905 a 0,931. A acurácia teve média de 0,932 e desvio de 0,008. As métricas de *precision* e *recall* também apresentaram valores de média macro próximos de 0,92, com baixa dispersão.

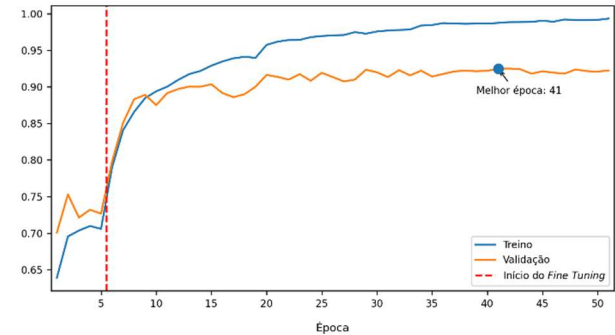


Figura 7. Evolução da acurácia durante as etapas de treinamento e validação, com destaque para a melhor época indicada pelo critério de *early stopping*, com base no *F1-macro* máximo no conjunto de validação. A linha tracejada indica o início do processo de *fine tuning*.

Figure 7. Evolution of accuracy during the training and validation stages, highlighting the best epoch determined by the early stopping criterion based on the maximum macro-F1 score on the validation set. The dashed line indicates the start of the fine-tuning process.

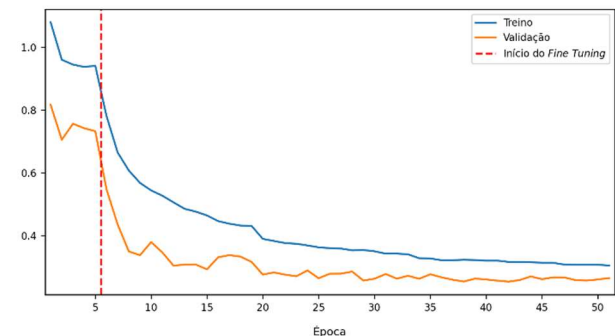


Figura 8. Evolução da função de perda (*loss*) ao longo das etapas de treinamento e validação, com destaque para a melhor época indicada pelo critério de *early stopping*, com base no *F1-macro* máximo no conjunto de validação. A linha tracejada indica o início do processo de *fine tuning*.

Figure 8. Evolution of the loss function during training and validation stages, highlighting the best epoch determined by the early stopping criterion based on the maximum macro-F1 score on the validation set. The dashed line indicates the start of the fine-tuning process.

Tabela 6. Métricas obtidas por classe no conjunto de teste, e a média macro de toda as classes.

Table 6. Metrics obtained by class in the test set, and the macro average of all classes.

Classe	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	Qtd amostras
Mancha da Glomerella	0,9615	0,9398	0,9506	1463
Dano por Herbicida	0,9224	0,8915	0,9067	627
Deficiência de Magnésio	0,8834	0,8902	0,8868	519
Deficiência de Potássio	0,8868	0,9444	0,9147	846
Sarna da Macieira	0,9722	0,9688	0,9705	1443
Média macro	0,9253	0,9270	0,9259	

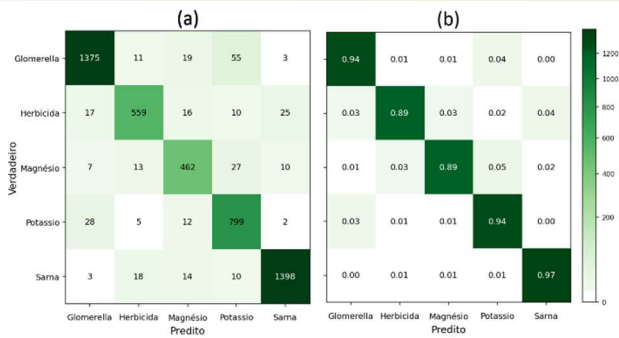


Figura 9. Matrizes de Confusão (a) original e (b) normalizada. A matriz original apresenta a contagem absoluta de classificações no conjunto de teste, enquanto a matriz normalizada apresenta a proporção de acertos e erros em relação ao total de amostras.

Figure 9. Confusion matrices: (a) original and (b) normalized. The original matrix shows the absolute count of classifications in the test set, while the normalized matrix shows the proportion of correct and incorrect classifications relative to the total number of samples.

Tabela 7. Métricas obtidas por partição na validação cruzada k -fold com $k=5$, e média macro e desvio padrão de todos os folds.

Table 7. Metrics obtained by partitioning in k -fold cross-validation with $k=5$, and the macro average and standard deviation of all folds.

Partição	<i>F1-Score</i>	<i>Acurácia</i>	<i>Precision</i>	<i>Recall</i>
1	0,912	0,928	0,909	0,916
2	0,922	0,939	0,926	0,918
3	0,905	0,921	0,905	0,906
4	0,920	0,932	0,923	0,918
5	0,931	0,942	0,934	0,928
Média macro	0,918	0,932	0,919	0,917
Desvio padrão	0,009	0,008	0,010	0,008

7. DISCUSSÃO

As matrizes de confusão indicam que a maioria dos exemplos é corretamente classificada, evidenciada pela predominância de valores na diagonal principal. O alcance de uma acurácia global de 93,8% e um *F1-score* macro de 0,926 no conjunto de teste permite concluir, indiretamente, que as regiões sintomáticas extraídas se mostraram altamente representativas e adequadas à tarefa. No entanto, a classe Mancha da Glomerella confunde-se com a Deficiência de Potássio, o que pode ser atribuído à semelhança visual entre os sintomas de ambas as classes. De forma semelhante, parte das amostras de Dano por Herbicida foi incorretamente atribuída à classe Sarna da Macieira.

Considerando a validação cruzada k -fold realizada, os resultados indicam que o modelo mantém desempenho consistente em diferentes divisões dos dados, reforçando sua estabilidade e capacidade de generalização. A análise das métricas globais, por classe e nas diferentes partições da validação cruzada indica desempenho estável e coerente com o conjunto de dados, evidenciando boa capacidade de generalização do modelo entre os folds.

Este trabalho constitui uma evolução metodológica direta em relação ao estudo pioneiro de Nachtigall (2016). A abordagem original alimentava as redes neurais com imagens brutas de folhas, expondo-as a ruídos ambientais e alcançando acurácias entre 93,3% e 97,3%. Em contrapartida, esta pesquisa introduziu uma etapa intermediária de segmentação para isolar estritamente as manchas patológicas, eliminando fundos complexos, variações de iluminação e tecidos saudáveis. Embora a acurácia geométrica dessa segmentação não tenha sido medida diretamente, sua eficácia

foi comprovada pelo desempenho do classificador MobileNetV2, que se mostrou resiliente às interferências do cenário de captura.

Adicionalmente, mitigou-se o severo desbalanceamento de classes do conjunto de dados original por meio do parâmetro *weight* na função *CrossEntropyLoss*, que ponderou a punição matemática de forma inversamente proporcional à abundância das amostras no treino. Esse ajuste resultou em um *F1-score* macro de 0,926, confirmando alta estabilidade e equilíbrio na identificação de todas as cinco patologias, independentemente do volume original de imagens por categoria.

Por fim, sob o aspecto da eficiência computacional, enquanto Nachtigall (2016) utilizou arquiteturas densas treinadas do zero, com longo tempo de convergência, este estudo adotou a MobileNetV2 combinada com *Transfer Learning*. Essa estratégia elevou a acurácia global para 93,8% com uma infraestrutura computacional leve, viabilizando o emprego do modelo em dispositivos de borda e em sistemas de tomada de decisão em tempo real no campo. Os resultados reforçam que a estratégia baseada na segmentação automática dos sintomas, na geração de recortes sintomáticos e no uso de uma arquitetura leve é viável para aplicações reais, equilibrando desempenho, custo computacional e reprodutibilidade.

A calibração manual e empírica dos intervalos de cor no espaço HSV, que se mostrou altamente eficiente para padronizar e tornar a MobileNetV2 robusta ao ruído do banco de dados, estabelece a base para futuros módulos de adaptação de iluminação em tempo real, mitigando variações extremas, como sombreamento denso ou reflexos solares diretos no campo. Da mesma forma, a elevada precisão alcançada na classificação das cinco patologias consolida a escolha de uma arquitetura otimizada e de baixo custo computacional como solução de borda viável, demonstrando que um modelo leve é capaz de entregar alta assertividade sem exigir hardware pesado. Esse sucesso abre caminho para que investigações subsequentes ampliem o escopo analítico da rede para cenários de campo ainda mais complexos, como a inclusão de amostras de folhas perfeitamente saudáveis no espectro de triagem, a identificação de coinfeções simultâneas ou o tratamento de interferências causadas por poeira e resíduos de defensivos na superfície foliar.

8. CONCLUSÕES

Este trabalho desenvolveu uma abordagem eficiente para a classificação de cinco distúrbios em folhas de macieira, fundamentada na premissa de construir um classificador leve baseado em recortes segmentados dos sintomas, em vez da folha inteira. Utilizando a arquitetura MobileNetV2 refinada por técnicas de *transfer learning* e *data augmentation*, o modelo final alcançou desempenho expressivo no conjunto de testes, registrando uma acurácia global de 93,8% e um *F1-score* macro de 0,926. Esses resultados confirmam o sucesso do método proposto e validam a viabilidade prática do sistema como solução robusta e de baixo custo computacional.

Como perspectivas de trabalhos futuros, propõe-se a expansão do banco de dados com a inclusão de classes adicionais, como folhas saudáveis, e a inserção de maior variabilidade de iluminação e de cenários de campo, permitindo testar a robustez da segmentação em ambientes externos dinâmicos. Adicionalmente, recomenda-se a realização de um estudo comparativo entre diferentes

arquiteturas de redes convolucionais leves sob o mesmo protocolo de treino, com o objetivo de mapear com precisão o melhor equilíbrio entre o desempenho preditivo e a eficiência computacional para dispositivos de borda.

9. REFERÊNCIAS

- AL-HIARY, H.; BANI-AHMAD, S.; REYALAT, M.; BRAIK, M.; ALRAHAMNEH, Z. Fast and accurate detection and classification of plant diseases. **International Journal of Computer Applications**, v. 17, n. 1, p. 31–38, 2011. <https://doi.org/10.5120/2183-2754>
- ARAUJO, L.; MEDEIROS, H. A. de; PASA, M.; SILVA, F. N. da. Doenças da macieira e da pereira. **Informe Agropecuário**, v. 37, n. 291, p. 61-74, 2016.
- BALLARD, D. H.; BROWN, C. M. **Computer vision**. Englewood Cliffs, NJ: Prentice Hall, 1982. 523p.
- 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. 46-53, 2018. <https://doi.org/10.1016/j.compag.2018.08.013>
- BUENO, M. P.; SILVA, A. C.; NUNES, A. L. de P. F.; SARDINHA, A. C.; LIMA, P. T. dos S.; SILVA, J. F. Análise da comercialização da cadeia produtiva da maçã brasileira: produção, importação e exportação no período 2015 a 2019. **Brazilian Journal of Development**, v. 7, n. 4, p. 34061-34078, 2021. <https://doi.org/10.34117/bjdv7n4-048>
- EPAGRI; CEPA. **Síntese anual da agricultura de Santa Catarina 2021-2022**. Florianópolis: Instituto CEPA/EPAGRI, 2023. 180p. Disponível em: https://docweb.epagri.sc.gov.br/website_cepa/publicacoes/Síntese_2021_22.pdf
- FAORO, I. D. (Org.). **Maçãs do grupo “Gala” no Brasil**. Florianópolis: Epagri, 2022. 305p. Disponível em: <https://agapomi.com.br/storage/2026/03/livro-macas-do-grupo-gala-no-brasil.pdf>
- FORSYTH, D.; PONCE, J. **Computer Vision: A Modern Approach**. London: Pearson, 2012. 793p.
- GABARDO, G. C.; KRETZSCHMAR, A. A.; PETRI, J. L.; COUTO, M.; HAWEROTH, F. J.; SILVA, C. D. S. Taxa fotossintética em macieiras tratadas com metamitron. **Revista Eletrônica Científica da UERGS**, v. 3, p. 617-633, 2017. <https://doi.org/10.21674/2448-0479.33.617-633>
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge: MIT Press, 2016. 800p.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: Data mining, inference, and prediction**. 2. ed. New York: Springer, 2009. 764p.
- KIRILLOV, A.; MINTUN, E.; RAVI, N.; MAO, H.; ROLLAND, C.; GUSTAFSON, L.; XIAO, T.; WHITEHEAD, S.; BERG, A. C.; LO, W.-Y.; DOLLÁR, P.; GIRSHICK, R. Segment anything. In: IEEE/CVF International Conference on Computer Vision (ICCV), 2023, Paris. **ArXiv**, p. 3992-4003, 2023. <https://doi.org/10.48550/arXiv.2304.02643>
- KAPOOR, S.; NARAYANAN, A. Leakage and the reproducibility crisis in machine-learning-based science. **Patterns**, v. 4, n. 9, e100804, 2023.
- KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. **Computers and Electronics in Agriculture**, v. 147, p. 70-90, 2018. <https://doi.org/10.1016/j.compag.2018.02.016>
- LIU, J.; WANG, X. Plant diseases and pests detection based on deep learning: a review. **Plant Methods**, v. 17, e22, 2021. <https://doi.org/10.1186/s13007-021-00722-9>
- MITCHELL, T. **Machine Learning**. New York: McGraw-Hill, 1997. 421p.
- NACHTIGALL, G. R.; CARGNINO, C. Nutrição mineral em macieiras. **Jornal da AGAPOMI**, n. 220, p. 6-7, 2012.
- NACHTIGALL, L. G. **Classificação de distúrbios em folhas de macieiras utilizando redes neurais convolucionais**. 84f. Dissertação [Mestrado em Computação] – Universidade Federal de Pelotas, Pelotas, 2016.
- PATEL, A.; JOSHI, B. A survey on the plant leaf disease detection techniques. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 6, n. 1, p. 229-231, 2017.
- RESENDE, R. T.; BRONDANI, C.; CHAVES, L. J. O melhoramento na era de agricultura de precisão. In: RESENDE, R. T.; BRONDANI, C. (Eds.). **Melhoramento de Precisão**. Brasília: Embrapa, 2023. p. 13-39.
- SALLATO, B.; DUPONT, T.; GRANATSTEIN, D. Tree fruit soil fertility and plant nutrition in cropping orchards in central Washington. Pullman, WA: Washington State University Extension, 2019.
- SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. MobileNetV2: Inverted residuals and linear bottlenecks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018. **ArXiv**, p. 4510-4520, 2018. <https://doi.org/10.48550/arXiv.1801.04381>
- SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. **Journal of Big Data**, v. 6, e60, 2019. <https://doi.org/10.1186/s40537-019-0197-0>
- SMITH, L. N. Cyclical learning rates for training neural networks. In: IEEE Winter Conference on Applications of Computer Vision. **ArXiv**, p. 464-472, 2017. <https://doi.org/10.48550/arXiv.1506.01186>
- SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing and Management**, v. 45, p. 427-437, 2009. <https://doi.org/10.1016/j.ipm.2009.03.002>
- STOCKMAN, G.; SHAPIRO, L. G. **Computer vision**. Upper Saddle River, NJ: Prentice Hall, 2001. 609p.
- SUTTON, R.; BARTO, A. **Reinforcement Learning: An Introduction**. Cambridge: MIT Press, 2018. 444p.
- SZELISKI, R. **Computer Vision: Algorithms and Applications**. Cham: Springer, 2022. 979p.
- VELHO, A. C.; STADNIK, M. J.; WALLHEAD, M. Unraveling *Colletotrichum* species associated with Glomerella leaf spot of apple. **Tropical Plant Pathology**, v. 44, p. 197–204, 2019. <https://doi.org/10.1007/s40858-018-0261-x>
- ZHANG, W.; DANG, L. M.; NGUYEN, L. Q.; ALAM, N.; BUI, N. D.; PARK, H. Y.; MOON, H. Adapting the segment anything model for plant recognition and automated phenotypic parameter measurement. **Horticulturae**, v. 10, n. 4, e398, 2024. <https://doi.org/10.3390/horticulturae10040398>

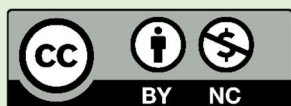
ZHAO, L.; OLIVIER, K.; CHEN, L. An automated image segmentation, annotation, and training framework of plant leaves by joining the SAM and the YOLOv8 models. **Agronomy**, v. 15, n. 5, e1081, 2025. <https://doi.org/10.3390/agronomy15051081>

Agradecimentos: Os autores agradecem à CAPES, ao CNPq e à UTFPR pelo apoio à pesquisa.

Contribuições dos autores: E.R.: metodologia, investigação ou coleta de dados, redação (original); R.E.O.S.A.: conceituação, administração ou supervisão, redação (revisão e edição). Todos os autores leram e concordaram com a versão publicada do manuscrito.

Disponibilidade de dados: Os dados desta pesquisa poderão ser obtidos por e-mail, mediante solicitação ao autor correspondente.

Conflito de interesses: Os autores declaram não haver conflitos de interesses.



Copyright: © 2026 by the authors. This article is an Open-Access article distributed under the terms and conditions of the Creative Commons **Attribution-NonCommercial (CC BY-NC)** license (<https://creativecommons.org/licenses/by/4.0/>).