

Aplicação do ResNet-50 para Classificação Binária de Folhas Cítricas

Luis Ricardo Sevilha Cavano¹

Mauricio Duarte²

RESUMO

Este trabalho apresenta a aplicação da arquitetura ResNet-50 para a classificação binária de folhas cítricas saudáveis e infectadas pelo Huanglongbing (HLB). A citricultura no Brasil é uma atividade agrícola de grande importância, sendo o país o maior exportador mundial de suco de laranja. O HLB, causado por bactérias do gênero *Candidatus Liberibacter*, é considerado a doença mais destrutiva da citricultura global, exigindo diagnósticos rápidos para um manejo eficaz.

A metodologia empregou o "Citrus Plant Dataset", que contém imagens de folhas cítricas saudáveis e infectadas por diversas doenças, incluindo o HLB. As imagens foram pré-processadas com redimensionamento, conversão de cores para o espaço YCrCb, equalização de histograma no canal Y, e filtragem bilateral para realçar detalhes e preservar bordas. O treinamento utilizou a ResNet-50 com aprendizado por transferência, reaproveitando pesos previamente treinados no ImageNet. Para otimização e prevenção de overfitting, foram aplicadas técnicas como dropout, global average pooling, ajuste fino (fine-tuning) das camadas do backbone, early stopping e ajuste dinâmico da taxa de aprendizado.

Os resultados demonstraram uma redução constante e significativa da função de perda (loss) em ambos os conjuntos de dados, sem sinais de overfitting, indicando a eficácia das estratégias de regularização. A acurácia de treinamento e validação apresentou um crescimento expressivo e se manteve estável nas épocas finais. A análise da diferença entre as acurácias (gap) indicou um bom ajuste e equilíbrio do modelo ao longo do treinamento.

Em conclusão, o estudo demonstrou que a combinação da ResNet-50 com aprendizado por transferência e técnicas de pré-processamento de imagens permite desenvolver um modelo de classificação binária altamente acurado (99,26%) e com

¹ Discente em Big Data no Agronegócio na FATEC Pompeia, Pompeia-SP,

² Docentes do curso Big Data no Agronegócio, FATEC Pompeia, Pompeia-SP.

baixo erro de generalização para detecção de HLB em folhas cítricas, mesmo com um dataset reduzido.

Palavras-chave: ResNet-50, Rede neural convolucional, Huanglongbing

1 INTRODUÇÃO

A citricultura representa uma importante atividade agrícola no Brasil, sendo a maior exportação de suco de laranja no mundo (IBGE, 2023). O que faz com que as doenças que afetam a citricultura no Brasil tenham grande relevância, dentre as doenças que veem causando grandes prejuízos a produção de citros no Brasil podemos destacar o Huanglongbing (HLB), também conhecido como greening. Essa doença é causada por bactérias do gênero *Candidatus Liberibacter*, é considerada atualmente como a mais destrutiva da citricultura global, levando a queda prematura de frutos e a morte das plantas (BOVÉ, 2006).

A identificação rápida do greening é de suma importância para o manejo adequado da doença e para a redução de impactos econômicos, porém a utilização de meios tradicionais para o diagnóstico, como inspeções visuais e testes laboratoriais, tendem a serem demorados além de requerer mão de obra especializada (SPITALERI et al., 2022).

Dado o contexto, técnicas de visão computacional juntamente ao aprendizado de máquina e as redes neurais convolucionais (CNNs), vem se provando como ferramentas promissoras para realizar diagnóstico de doenças em plantas (MOHAMMED et al., 2021). De modo especial as CNNs dado que elas possuem a capacidade de extrair e aprender padrões visuais complexos de imagens o que facilita a detecção de anomalias em folhas, frutos e caules.

Apesar destes avanços recentes, a aplicação de visão computacional na agricultura enfrenta um desafio prático significativo, conforme apontado por importantes revisões da área: a dependência de grandes volumes de dados de treinamento (KAMILARIS; PRENAFETA-BOLDÚ, 2018). Arquiteturas de redes profundas, que são as mais eficazes na extração de padrões, geralmente demandam muitas imagens rotuladas por especialistas, um recurso escasso e de alto custo para doenças específicas como o HLB. Como resultado, e servindo de motivação para este estudo, identifica-se uma lacuna na literatura sobre estratégias para adaptar e validar modelos profundos em cenários realistas de poucos dados para o diagnóstico em citros.

Dentre as inúmeras arquiteturas de CNNs utilizadas a ResNet-50 possui destaque devido a sua capacidade de introduz conexões residuais para permitir o treinamento eficiente de redes profundas, melhorando o desempenho em tarefas de classificação visual (HE et al., 2016).

Este trabalho se propõe ao desenvolvimento de uma modelo baseado no ResNet-50 para classificação binária de folhas cítricas em saudáveis e infectadas pelo Huanglongbing. O processo envolve a aplicação de pré-processamentos de imagens

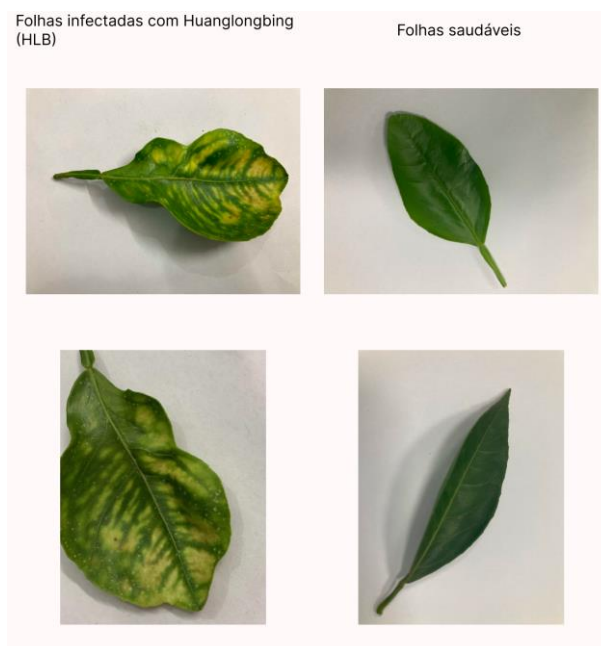
como redimensionamento, conversão de cores para o espaço YCrCb, equalização de histograma e filtragem bilateral, além do uso de estratégias como regularização por dropout, fine-tuning, early stopping e monitoramento por TensorBoard para evitar overfitting e melhorar a generalização do modelo (SRIVASTAVA et al., 2014; GOODFELLOW et al., 2016).

Este estudo busca demonstrar que apesar de poucas amostras é possível obter um desempenho robusto ao combinar boas práticas de pré-processamento, aprendizado por transferência e redes convolucionais profundas. Resultados quantitativos como acurácia, perda e estabilidade durante o treinamento serão utilizados para validar a eficácia da abordagem.

2 MATERIAL E MÉTODOS

O conjunto de dados utilizados foi o Citrus Plant Dataset, obtido no site Mendeley Data (Rauf et al., 2019). Ele contém 150 imagens de frutas cítricas e 609 imagens de folhas cítricas divididas entre saudáveis e infectadas por diversas doenças. Após fazer o baixar e extrair o conteúdo, as imagens de interesse foram separadas em pastas diferentes de acordo com suas classes, ao todo foram obtidas 594 imagens de folhas saudáveis e 254 imagens de folhas infectadas pelo Huanglongbing, posteriormente, foi realizada uma etapa de pré-processamento.

Figura 1 – Exemplo de imagens de folhas obtidas do conjunto de dados Citrus Plant Dataset



Fonte: Elaborada pelo autor (2025).

O pré-processamento realizado nas imagens consiste na aplicação do redimensionamento das imagens para a resolução de 224 x 224 pixels, para garantir

a compatibilidade com a rede neural ResNet-50 e diminuir os custos computacionais do treinamento sem qualquer perda considerável de informação.

Em sequência, as imagens foram convertidas de cor RGB para YCrCb, o que permite a separação de informações da luminância (canal Y) dos componentes de cromaticidade (Cr e Cb). A conversão torna-se relevante pelo fato de a maior parte das informações das imagens estarem contidas na luminância, que é responsável pela percepção de contraste e nitidez. De tal modo, foi aplicada a equalização de histograma de forma exclusiva no canal Y, a fim de realçar detalhes pouco visíveis e as bordas da folha, o que torna mais fácil a detecção de padrões possivelmente patológicos que sejam discretos. Gonzalez e Woods (2018) destacam que a equalização de histograma é especialmente eficaz em imagens com iluminação desigual ou contraste reduzido, como frequentemente ocorre em amostras obtidas em ambientes agrícolas.

Figura 2 – Exemplo de imagens de folhas obtidas do conjunto de dados Citrus Plant Dataset após o pré-processamento.



Fonte: Elaborada pelo autor (2025).

Tendo sido finalizada a equalização, foi aplicada uma filtragem bilateral, que é uma técnica para realizar suavização, reduzindo os ruídos ao mesmo tempo que preserva as bordas. A filtragem bilateral foi aplicada para realçar as regiões uniformes da folha sem que se percam as bordas, o que garante que a estrutura morfológica da folha é mantida como pode ser observado na Figura 2. Por fim, as imagens pós-processadas são salvas em um diretório diferente que será utilizado para treinar o modelo.

O treinamento do modelo de classificação foi realizado utilizando o ResNet-50 (He et al., 2016), que é uma arquitetura amplamente utilizada em aplicações de visão computacional, por ser capaz de detectar padrões espaciais em diferentes

profundidades utilizando blocos residuais para isso. A implementação do ResNet-50 para este estudo também utilizou transferência de aprendizado (transfer learning). Nesse processo, os pesos do ResNet-50, previamente treinado no conjunto de dados do ImageNet, são reaproveitados como pesos iniciais do modelo; a transferência de aprendizado permite acelerar o treinamento e aumentar o desempenho do modelo em cenários com poucos dados disponíveis para treinamento.

Ao iniciar o treinamento, as partes convolucionais da ResNet-50 são mantidas em seu estado original, o que permite que apenas as camadas superiores sejam treinadas. O modelo foi estruturado para que a saída do backbone, porção responsável pela extração hierárquica de características das imagens, fosse conectada a uma camada de global average pooling. Em seguida, foram adicionadas camadas densas com ativação ReLU, normalização por lotes e camadas com dropout com taxa de 50%, que ocasionam desligamento aleatório de neurônios durante o processo de treinamento, o que diminui a ocorrência de overfitting, situação em que o modelo memoriza o conjunto de treino sem generalizar bem para novos dados.

O treinamento inicial foi realizado, com 80% das imagens de cada classe no conjunto de dados para treinamento e 20% para validação e por um período de 15 épocas com o otimizador Adam, com uma taxa de aprendizado (learning rate) de 0,0001. O treinamento também contém o mecanismo de early stop, o qual interrompe automaticamente o treinamento quando a acurácia de validação não melhora após X épocas. Neste estudo, o early stop realiza a interrupção do treinamento se a acurácia de validação não melhorar durante 5 épocas. Também foi implementado ao treinamento o ajuste automático do learning rate se a acurácia de validação permanecer estável por um período de 3 épocas, o ajuste automático do learning rate contribui para refinar a otimização do modelo em fases avançadas do treinamento, segundo (Smith, 2017). Para acompanhar o desempenho do modelo durante o treinamento, utilizou-se o TensorBoard, uma ferramenta para visualização que realiza o registro de métricas do treinamento. Durante o processo de registro das métricas, foram salvos os pesos da rede sempre que se obtinha um novo melhor valor de acurácia de validação, a fim de garantir que o melhor desempenho do modelo fosse sempre salvo. Finalizando essas etapas do processo, foi realizado o processo de ajuste fino (fine-tuning); nesse processo, as 50 camadas do backbone que foram mantidas congeladas durante o treinamento são descongeladas para permitir um refinamento delas. O refinamento é realizado com um learning rate de $(1e-5)$ por mais 5 épocas; o ajuste fino tem como objetivo adaptar os filtros convolucionais da rede aos padrões específicos das imagens contidas no dataset.

Após o término do treinamento foi realizada uma avaliação do modelo de classificação, utilizando de métricas quantitativas, análise temporal do desempenho no decorrer do treinamento e validação cruzada quantitativa para isso foi utilizado a leitura dos logs gerados durante o treinamento com as métricas e curvas de aprendizagem. O principal foco dessa análise é garantir que o modelo possui estabilidade e generalização adequada das classes ao longo do treinamento.

A principal métrica utilizada neste estudo para a avaliação do modelo foi a acurácia, definida como a razão entre as classificações corretas e o número total de amostras para validação. A acurácia foi obtida a partir do conjunto de testes, que foi isolado dos processos de treinamento e validação, tal como Kohavi (1995) recomenda para evitar viés na estimativa da generalização do modelo. A acurácia foi escolhida para ser a principal métrica de avaliação devido a abordagem deste estudo - classificação binária entre folhas cítricas saudáveis e infectadas pelo Huanglongbing - e a distribuição balanceada dos dados, o que torna a acurácia uma métrica apropriada.

A avaliação se deu por meio do método de `model.evaluate()` da API do Tensor Flow, que calcula as métricas predefinidas durante a compilação do modelo, aplicando lote a lote ao conjunto de testes. Também se utilizou a função de perda (loss) a fim de monitorar e aferir o quão bem o modelo consegue minimizar o erro de predição a partir da função de custo `sparse_categorical_crossentropy`. Ao combinar essas duas métricas pode-se avaliar a qualidade das predições (acurácia) e o grau de confiança e penalização de erros (função de perda), o que apresenta uma visão holística do desempenho apresentado pelo modelo (Goodfellow et al., 2016).

O processo de validação verificou os seguintes comportamentos:

Overfitting que é identificado quando a acurácia no conjunto de treinamento aumenta progressivamente enquanto a acurácia do conjunto de validação, não apresenta um aumento correspondente ao do treinamento. Neste estudo utilizamos a diferença entre a acurácia do treinamento e da validação sendo maior que 10% como critério para caracterizar o overfitting. Esta prática está de acordo com que Srivastava et al. (2014) propõem;

Underfitting que pode ser identificado quando a acurácia do treinamento e da validação estão abaixo de um valor mínimo de 70%, o que indica incapacidade representacional do modelo ou quantidade de dados e ou épocas de treino insuficientes;

Instabilidade é identificada pela variação absoluta média entre as acurácias de validação em épocas consecutivas, variações que sejam superiores a 5% são consideradas indicador de instabilidade, a instabilidade pode decorrer de dados com ruído, learning rate elevado ou arquitetura inadequada.

Os três comportamentos utilizados para avaliar o modelo são calculadas por um script em Python que realiza a leitura dos logs e realiza os cálculos dessas métricas.

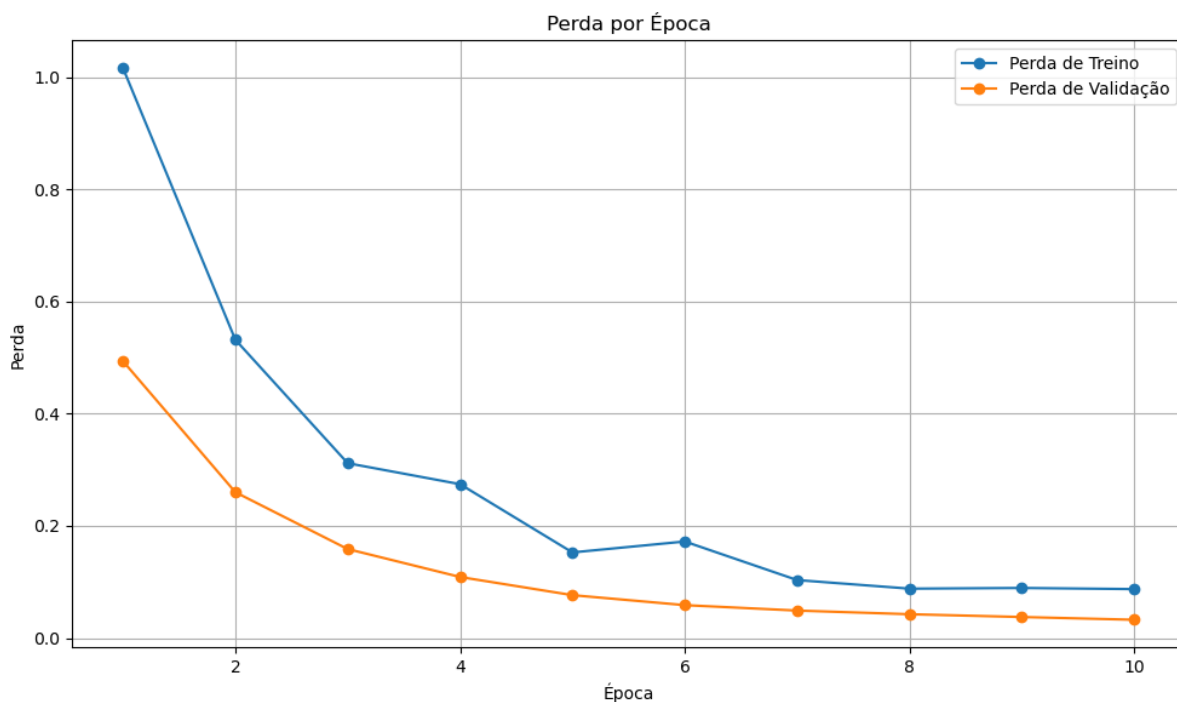
Após o processo de fine-tuning o modelo é reavaliado para comparar os resultados do treinamento inicial com o do fine-tuning, o que permite que se avalie também o desempenho da personalização do backbone.

3 RESULTADOS E DISCUSSÃO

Os resultados basearam-se na avaliação do modelo de classificação binária baseada na arquitetura ResNet-50 com o desempenho do modelo sendo avaliado através das métricas de acurácia, função de perda (loss) e a diferença entre a acurácia do treinamento e da validação (gap). O estudo destaca a estabilidade e a capacidade de generalização.

A Figura 3 demonstra a evolução do loss para os conjuntos de treinamento e validação durante um período de 10 épocas. É observado uma redução constante e significativa do loss em ambos os conjuntos. O loss do treinamento inicia-se por volta de 1,02 reduzindo-se até 0,09. Já a perda do conjunto de validação se inicia em 0,49 reduzindo até se estabilizar em 0,03.

Figura 3 – Gráfico da função de perda (loss) dos conjuntos de treinamento e validação em função da época.



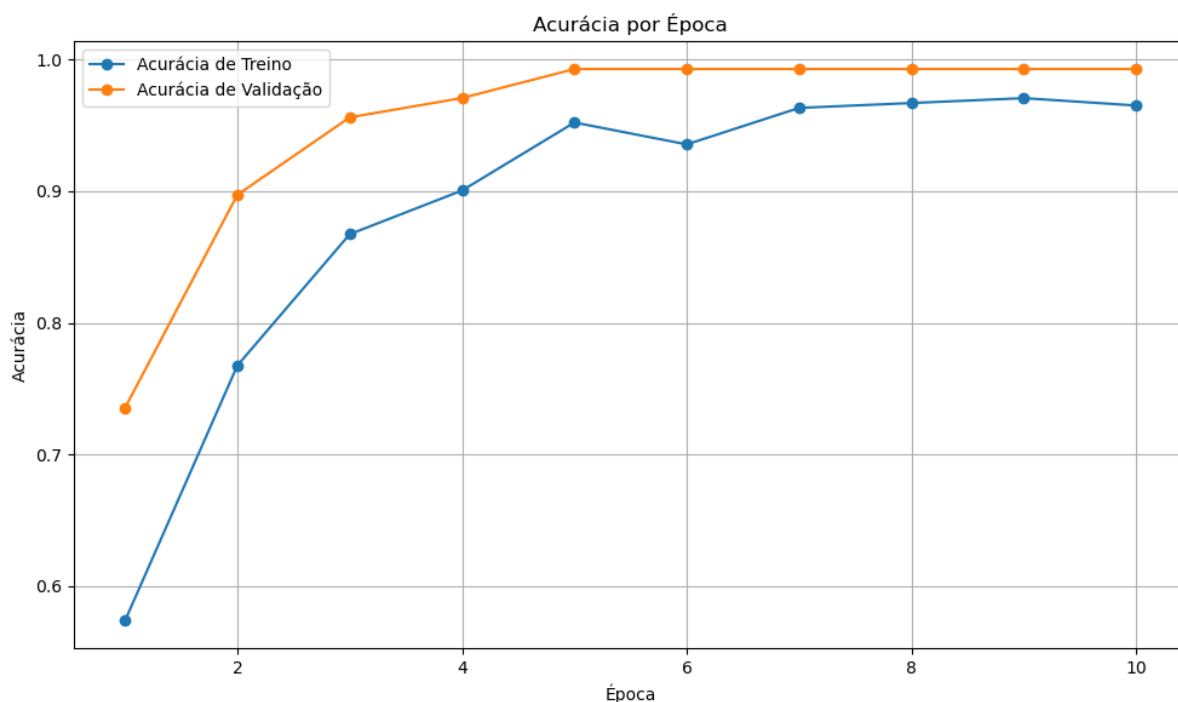
Fonte: Elaborada pelo autor (2025).

O fato de não haver aumento do loss durante as últimas épocas do conjunto de validação indica que o modelo não apresenta sinais de overfitting, dado que a função de custo continuou a diminuir de forma coerente, esse comportamento do loss aponta para a eficácia da estratégia de regularização com dropout, também com os ajustes automáticos do learning rate e o uso do mecanismo de early stop.

A figura 4 demonstra a evolução da acurácia por época para o conjunto de treinamento e o conjunto de validação. A acurácia de treinamento inicia-se em 57% e

crece continuamente, até atingir aproximadamente 97% durante a última época. Já a acurácia do conjunto de validação apresenta um crescimento notavelmente mais expressivo saindo de 73% inicialmente até atingir surpreendentes 99,26% a partir da quarta época de onde passou a permanecer estável.

Figura 4 – Gráfico da acurácia dos conjuntos de treinamento e validação em função da época.

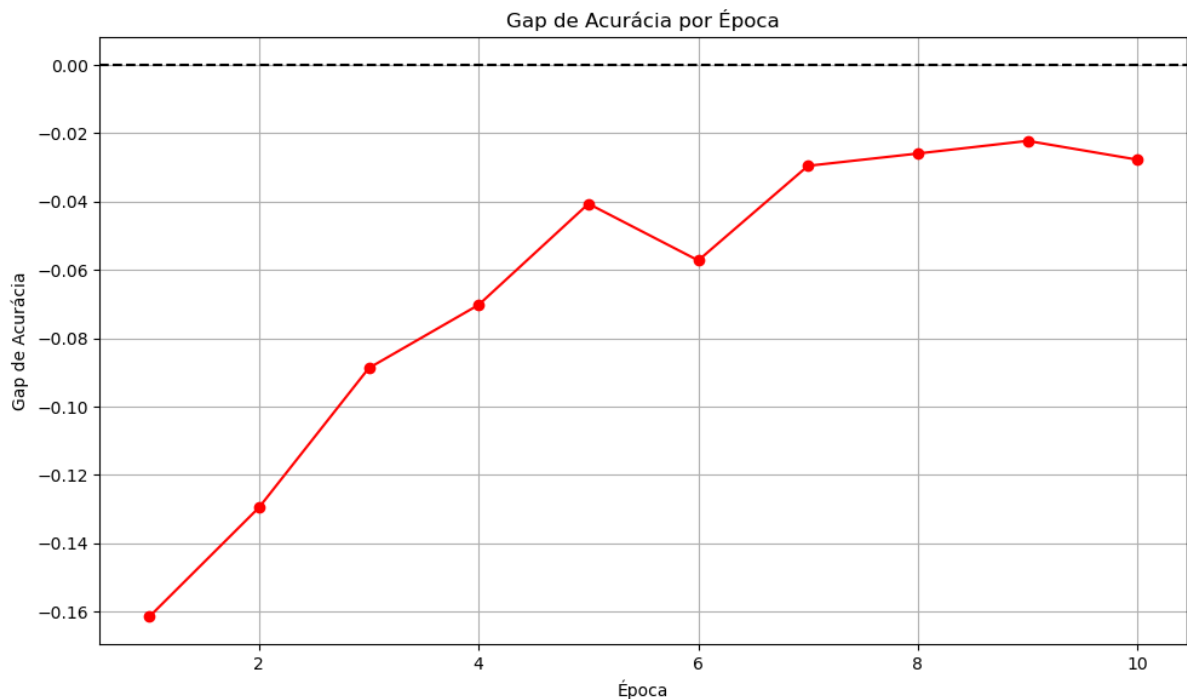


Fonte: Elaborada pelo autor (2025).

O comportamento da acurácia indica um rápido aprendizado do modelo aos padrões discriminativos entre as duas classes de folhas cítricas - saudáveis e infectadas por Huanglongbing – este rápido aprendizado se deve pela aplicação da transferência de aprendizado proporcionado pelo ResNet-50 pré treinado no ImageNet. E a aplicação do fine-tuning contribui, para que o modelo especialize os filtros convolucionais as características morfológicas das folhas cítricas.

A Figura 5 demonstra a diferença entre a acurácia da validação e de treinamento (gap). Durante as primeiras épocas o gap atingiu valores de $-0,16$, o que é indicativo de uma inicial dificuldade do modelo em se ajustar os dados de treinamento. Porém a diferença foi reduzida progressivamente, atingindo valores próximos de zero ($-0,03$) durante as últimas épocas, o que indica que o modelo está equilibrado e bem ajustado.

Figura 5 – Gráfico do diferencia de acurácia (gap) em função da época.



Fonte: Elaborada pelo autor (2025).

Com base nos resultados obtidos e na metodologia escolhida para avaliar o desempenho do modelo pode se inferir que o modelo não apresenta o overfitting, underfitting e instabilidade o que pode indicar que as estratégias de regularização como uso de dropout, normalização em lotes e ajustes dinâmicos de taxa de aprendizado, conforme recomendação de Smith (2017), apresentam se de forma eficaz para o caso de estudo deste estudo.

Além deste ponto, os resultados também evidenciam que mesmo com um conjunto de dados pequenos e com aplicação de pré-processamentos adequados, além do uso de técnicas modernas de aprendizado por transferência tornam viável o desenvolvimento deste modelo para classificação binária de folhas cítricas.

4 CONSIDERAÇÕES FINAIS

Este estudo apresentou o desenvolvimento e a avaliação de um modelo baseado na arquitetura do ResNet-50 visando a classificação de folhas cítricas saudáveis e infectadas por Huanglongbing (HLB). A abordagem foi o uso de aprendizado por transferência (transfer learning) associado a técnicas de pré processamento de Images a fim de obter um modelo robusto apesar de um dataset com dados limitados.

Os resultados obtidos demonstram que o modelo foi capaz de apresentar uma acurácia elevada (99,26%) com um baixo erro de generalização, apesar do dataset utilizado ser reduzido. As curvas de loss e acurácia se demonstraram estáveis ao

longo das épocas o que é um bom indicativo que problemas como overfitting ou underfitting não aconteceram, o que reforça a validade e a eficácia do modelo e das estratégias utilizadas para otimizar o treinamento do modelo como regularização por dropout e ajuste dinâmico da taxa de aprendizado.

Apesar destes resultados promissores este estudo ainda lida com limitações nas quantidades e disponibilidade de imagens para serem utilizadas para treinamento e validação do modelo. Futuros estudos deveriam focar na coleta e aumento de imagens além de testes do modelo com imagens que não fazem parte do dataset utilizado neste estudo.

A aplicação do ResNet-50, originalmente projetado para realizar atividades de alta complexidade mostrou-se eficiente para o contexto deste estudo. Em suma, os resultados que foram obtidos neste estudo de aplicação do ResNet-50 para classificação binária de folhas cítricas, podem vir a contribuir para o futuro desenvolvimento de soluções utilizando poucos dados para o agronegócio.

REFERÊNCIAS

IBGE. Instituto Brasileiro de Geografia e Estatística. Produção Agrícola Municipal, 2023.

BOVÉ, J. M. Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *Journal of Plant Pathology*, v. 88, n. 1, p. 7-37, 2006.

SPITALERI, D. D. et al. Automatic detection of HLB citrus disease using deep learning models. *Expert Systems with Applications*, v. 193, 2022.

Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

Rauf, Hafiz Tayyab; Saleem, Basharat ALi ; Lali, M. Ikram Ullah ; Khan, Muhammad Attique ; Sharif, Muhammad; Bukhari, Syed Ahmad Chan (2019), "A Citrus Fruits and Leaves Dataset for Detection and Classification of Citrus Diseases through Machine Learning", *Mendeley Data*, V2, doi: 10.17632/3f83gxm57.2

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *IJCAI*.

HE, K. et al. Deep Residual Learning for Image Recognition. In: *Proceedings of the IEEE CVPR*, 2016.

SRIVASTAVA, N. et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, v. 15, p. 1929–1958, 2014.

KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A review on development, applications, and challenges. **Computers and Electronics in Agriculture**, v. 147, p. 70-90, 2018.