

Comparative Analysis of Deep Learning Architectures for Classifying Soy Seeds

Vinicius Godoy Marques¹, Lucas Correia Ribas², Caetano Mazzoni Ranieri¹

¹Instituto de Geociências e Ciências Exatas – Universidade Estadual Paulista (UNESP)
Rio Claro – 13506-900 – SP – Brazil

²Instituto de Biociências, Letras e Ciências Exatas – Universidade Estadual Paulista (UNESP)
São José do Rio Preto – 15054-000 – SP – Brazil

{vg.marques, lucas.ribas, cm.ranieri}@unesp.br

Abstract. *Soybean seed quality plays a critical role in agricultural productivity and economic returns. Recent studies have applied deep learning techniques to automate seed classification. However, these approaches either focus on binary classification or do not thoroughly compare multiple architectures for multiclass soybean seed classification. This work proposes a comparative analysis of deep learning models to classify soybean seeds into five categories. The methodology involves training and evaluating different models using the Soybean Seeds dataset, followed by fine-tuning the best-performing model. Experimental results show that InceptionV3 achieved the highest accuracy, improving from 76.68% with transfer learning to 86% after fine-tuning.*

Resumo. *A qualidade das sementes de soja desempenha um papel crucial na produtividade agrícola. Estudos têm aplicado técnicas de aprendizado profundo para automatizar a classificação de sementes. No entanto, essas abordagens focam em classificações binárias ou não comparam múltiplas arquiteturas para a classificação multiclasse. Este trabalho propõe uma comparação de modelos para classificar sementes de soja em cinco categorias. A metodologia envolve a avaliação de diferentes modelos utilizando o conjunto de dados Soybean Seeds, seguido do ajuste fino do modelo com melhor desempenho. Os resultados mostram que o modelo InceptionV3 alcançou a maior acurácia, melhorando de 76,68% com aprendizado por transferência para 86% após ajuste fino.*

1. Introdução

A agricultura de precisão tem se beneficiado significativamente dos avanços em inteligência artificial, particularmente no que diz respeito à classificação automatizada de grãos e sementes [Ayoub Shaikh et al. 2022]. Dentre os diversos cultivos agrícolas, a soja se destaca como uma das *commodities* mais importantes globalmente, sendo fundamental para a alimentação humana, produção de ração animal e biocombustíveis [Siamabele 2021]. A qualidade das sementes de soja está diretamente relacionada ao rendimento das lavouras, à eficiência produtiva e, consequentemente, ao retorno econômico dos agricultores [Vogel et al. 2021]. Tradicionalmente, a classificação dessas sementes é realizada por meio de inspeção visual humana, um processo demorado, subjetivo e propenso a erros, especialmente quando realizado em larga escala [Yun et al. 2024].

Nesse contexto, as técnicas de *Deep Learning*, em especial as Redes Neurais Convolucionais (*CNNs*), destacam-se como alternativas promissoras para automatizar e aprimorar a classificação de sementes [Saleem et al. 2021]. Essas redes são altamente eficazes em tarefas de visão computacional, pois extraem representações hierárquicas das imagens, capturando padrões complexos muitas vezes imperceptíveis ao olho humano [Jogin et al. 2018]. A adoção de arquiteturas avançadas, aliada a estratégias como *Transfer Learning* e *fine tuning*, tem potencial para elevar a acurácia e a robustez dos modelos. Na agricultura, tais soluções contribuem significativamente para a redução de custos operacionais, o aumento da eficiência na triagem de sementes e a minimização de perdas pós-colheita [Upadhyay and Bhargava 2025].

Este artigo propõe o desenvolvimento e a avaliação de modelos baseados em *Deep Learning* para a classificação automática de sementes de soja em cinco categorias distintas: intactas, imaturas, manchadas, partidas e com danos na pele/casca. Utiliza-se o *dataset* “Soybean Seeds” [Lin et al. 2023], composto por 5.513 imagens de alta resolução obtidas em condições controladas. Inicialmente, é implementada uma rede neural convolucional (*CNN*) personalizada para realizar a tarefa de classificação. Em seguida, são avaliadas diferentes abordagens de *Transfer Learning*, utilizando arquiteturas consolidadas como *EfficientNetB0* [Tan and Le 2019], *ResNet50* [He et al. 2015], *VGG16* [Simonyan and Zisserman 2015], *ConvNeXtSmall* [Liu et al. 2022] e *InceptionV3* [Szegedy et al. 2015], a fim de comparar o desempenho entre modelos prontos e personalizados. Os modelos são analisados com base na métrica de acurácia. Dentre os modelos testados, o *InceptionV3* apresentou o melhor desempenho, sendo posteriormente refinado por meio de *fine tuning*, o que resultou em ganhos significativos de desempenho.

A proposta visa oferecer uma solução que possa ser incorporada a sistemas de triagem automática, como classificadores agrícolas, contribuindo com a automação e modernização do setor. A relevância desta pesquisa estende-se além do aspecto tecnológico: um sistema eficiente de classificação pode beneficiar toda a cadeia produtiva da soja — do produtor, que passa a ter maior controle sobre a qualidade das sementes, até a indústria, que pode selecionar matéria-prima com mais precisão. Além disso, a metodologia pode ser adaptada para outros tipos de grãos, ampliando as possibilidades no contexto da agricultura de precisão.

2. Trabalhos Relacionados

A aplicação de redes neurais convolucionais (*CNNs*) para classificação e fenotipagem de sementes, especialmente de soja, tem avançado significativamente. O uso de técnicas de *transfer learning*, aliadas a arquiteturas modernas, tem impulsionado esses avanços ao possibilitar o treinamento eficiente mesmo com conjuntos de dados limitados. Por exemplo, [Yang et al. 2021] desenvolveram um método que utiliza a geração sintética de imagens com *domain randomization*, reduzindo a dependência de grandes volumes de dados anotados manualmente. Paralelamente, avanços em arquiteturas especializadas para a classificação de danos em sementes vêm mostrando resultados promissores. [Huang et al. 2022] propuseram um *pipeline* que combina segmentação via *Mask R-CNN* com uma rede leve, chamada *SNet*, incorporando módulos de recalibração de características para realçar regiões danificadas. De maneira complementar, [Chauhan et al. 2025] abordaram o problema do desbalanceamento de classes em um conjunto de imagens de alta resolução, onde o modelo *EfficientNet-B0* apresentou

desempenho superior.

Estudos sistemáticos que avaliam a eficácia de modelos pré-treinados em classificação de sementes reforçam o papel central do *transfer learning* nessa área. [Zhu et al. 2020] realizaram uma avaliação comparativa entre seis arquiteturas populares aplicadas a 9.600 imagens hiperespectrais de soja. No que tange à generalização dos modelos para aplicações reais, o estudo de [Pereira et al. 2024] trouxe uma importante contribuição ao disponibilizar um *dataset* público com sementes provenientes de três regiões geográficas distintas. Essa diversidade permitiu avaliar a capacidade dos modelos em lidar com variações regionais.

A aplicação de técnicas de *deep learning* para diagnóstico de doenças em soja também tem recebido atenção significativa. [Bevers et al. 2022] montaram um dos maiores conjuntos de dados do campo, com mais de 9.500 imagens, e demonstraram que a arquitetura *DenseNet201* alcançou 96,8% de acurácia para a identificação de oito classes de doenças e deficiências. Importante destacar que a utilização de imagens no contexto da planta inteira, ao invés de imagens isoladas das folhas, melhorou a *performance* dos modelos em cerca de 15%, além de que estratégias de aumento de dados para classes minoritárias foram determinantes para o balanceamento e robustez dos classificadores.

Uma análise comparativa realizada por [Eryigit and Tugrul 2021] avaliou 15 variações de *CNNs* aplicadas a sete espécies diferentes de gramíneas. O *DenseNet201* apresentou a melhor acurácia, apesar de seu maior custo computacional. O estudo também destacou que o uso de pesos pré-treinados aumentou a acurácia em aproximadamente 3% em comparação ao treinamento do zero, e que não há correlação direta entre o número de parâmetros do modelo e seu desempenho.

3. Metodologia

Nesta seção, apresenta-se a metodologia adotada para a comparação de modelos de classificação de imagens de sementes de soja, utilizando redes neurais convolucionais (*CNN*) e técnicas de *Transfer Learning*, além do *fine-tuning* do modelo *InceptionV3*. O desenvolvimento e os experimentos foram realizados no ambiente *Google Colab*, utilizando uma *GPU Nvidia Tesla T4* para acelerar o treinamento dos modelos. O *framework* utilizado foi o *TensorFlow* (versão 2.x) com a *API Keras* para construção, treinamento e avaliação das redes neurais.

Utiliza-se o *dataset* “Soybean Seeds” [Lin et al. 2023], composto por 5.513 imagens de alta resolução obtidas em condições controladas de sementes de soja divididas de forma balanceada em cinco classes distintas. As imagens foram organizadas em três conjuntos principais: treinamento, validação e teste. As pastas contendo as imagens foram carregadas a partir do *Google Drive*, mantendo a estrutura de diretórios compatível com o método *flow_from_directory* do *Keras*, facilitando a alimentação dos dados durante o treinamento.

Para a preparação das imagens, utilizou-se o redimensionamento para tamanhos fixos, sendo 128×128 pixels na comparação inicial entre a *CNN* customizada e as técnicas de *Transfer Learning*, e 256×256 pixels para o *fine-tuning* do melhor modelo, *InceptionV3*. Além disso, aplicou-se a normalização dos valores dos pixels para a faixa de [0,1], por meio da divisão dos valores por 255. No experimento de *fine-tuning*, foi aplicado *Data*

Augmentation no conjunto de treino, incluindo rotações aleatórias de até 20 graus, *zoom* de até 20%, *flips* horizontais e deslocamentos horizontais e verticais de até 10%. Essas técnicas visam aumentar a diversidade das amostras e reduzir o *overfitting*. A arquitetura da *CNN* personalizada utilizada é apresentada na Tabela 1. Trata-se de uma rede composta por três camadas convolucionais seguidas por camadas densas com *dropout* para evitar *overfitting*, com saída para classificar as cinco classes de sementes do *dataset*. As imagens de entrada são processadas em tamanho reduzido para 128×128 pixels e 3 canais (*RGB*).

Tabela 1. Arquitetura da Rede Neural Convolucional Implementada

Camada	Tipo	Parâmetros principais
Entrada	(128, 128, 3)	Imagen RGB redimensionada
Conv2D + Pooling	8 filtros (3×3)	ReLU, padding='same'
Conv2D + Pooling	16 filtros (3×3)	ReLU, padding='same'
Conv2D + Pooling	32 filtros (3×3)	ReLU, padding='same'
Flatten	-	-
Dense + Dropout	32 unidades, 0.3	ReLU
Saída	5 classes	Softmax

Além da *CNN* customizada, foram avaliados diversos modelos pré-treinados através de *Transfer Learning*, utilizando redes já treinadas na base *ImageNet* com suas camadas convolucionais congeladas e adaptando as camadas densas finais para o problema específico. A Tabela 2 apresenta um resumo da arquitetura dos modelos pré-treinados testados.

Tabela 2. Arquitetura dos Modelos com Transfer Learning (Camadas Convolucionais Congeladas)

Camada	Tipo	Parâmetros principais
Entrada	(128, 128, 3)	Imagen RGB redimensionada
Base Pré-treinada	EfficientNetB0, ResNet50, etc.	Pesos da ImageNet, convoluções congeladas
Pooling	GlobalAveragePooling2D	Redução da dimensionalidade
Dense	128 unidades	ReLU
Dropout	0.3	Regularização
Saída	5 classes	Softmax

O treinamento de cada modelo foi conduzido utilizando o otimizador Adam, com taxa de aprendizado inicial de 0.0001, função de perda *categorical crossentropy*, definida por:

$$L = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (1)$$

- C : número de classes;
- y_i : valor real (1 para a classe correta e 0 para as demais);
- \hat{y}_i : probabilidade prevista para a classe i .

A métrica de avaliação utilizada foi a *accuracy*, definida como:

$$\text{Accuracy} = \frac{\sum_{j=1}^N \mathbf{1}(\hat{y}_j = y_j)}{N} \quad (2)$$

- N : número total de amostras;
- $\mathbf{1}(\cdot)$: função indicadora (1 quando a condição é verdadeira, 0 caso contrário);
- \hat{y}_j : classe prevista para a amostra j ;
- y_j : classe real da amostra j .

Para controle do processo, foram utilizados *callbacks* como *EarlyStopping*, que interrompe o treinamento caso a métrica de validação não melhore após oito épocas consecutivas, *ReduceLROnPlateau*, que reduz a taxa de aprendizado pela metade após quatro épocas sem melhora na validação, e *ModelCheckpoint*, utilizado especialmente no *fine-tuning* do *InceptionV3* para salvar o melhor modelo durante o treinamento.

No *fine-tuning* do *InceptionV3*, após o treinamento inicial com a base congelada, todas as camadas a partir da posição 250 foram descongeladas — independentemente do tipo (como camadas convolucionais, de normalização, ativação, *pooling* e concatenação) — para permitir o ajuste fino das partes mais profundas da rede. Essa estratégia é adotada porque as camadas iniciais captam padrões visuais mais genéricos, como bordas e texturas, enquanto as camadas mais profundas representam características mais específicas do problema, sendo mais relevantes para a adaptação ao novo conjunto de dados. A taxa de aprendizado foi reduzida para 0.00001 para evitar grandes alterações nos pesos pré-treinados, e o modelo foi re-treinado para melhorar sua adaptação ao *dataset* específico. Abaixo conforme ilustrado na Figura 1 podemos visualizar o *pipeline* utilizado.

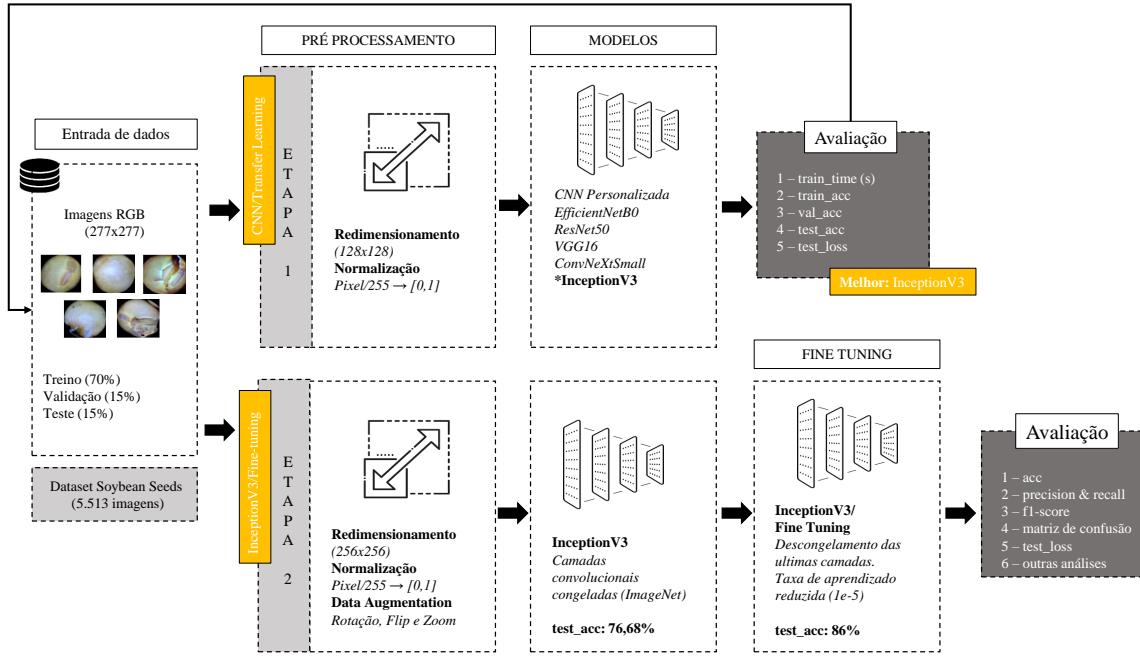


Figura 1. Diagrama completo da metodologia: desde a entrada de dados e pré-processamento até a aplicação do modelo de classificação de sementes de soja. O processo compara uma CNN personalizada com modelos de Transfer Learning, seleciona o InceptionV3 para fine-tuning e avalia os resultados finais.

4. Resultados e Discussão

Nesta seção, são apresentados os resultados obtidos a partir do treinamento e avaliação dos modelos para classificação das sementes de soja. Os experimentos foram organizados em duas etapas: inicialmente, avaliou-se o desempenho de uma rede neural convolucional (CNN) personalizada em comparação com cinco modelos pré-treinados por *Transfer Learning*; posteriormente, realizou-se o *fine-tuning* do modelo *InceptionV3*, com ajustes nas camadas convolucionais e aplicação de técnicas de aumento de dados.

A Tabela 3 sintetiza o desempenho da CNN personalizada e dos modelos transferidos, mostrando o tempo total de treinamento, as acurárias nos conjuntos de treino, validação e teste, além da perda final no conjunto de teste. Os resultados indicam que o *InceptionV3* obteve o melhor desempenho geral, alcançando a maior acurácia no conjunto de teste (76,68%) e a menor *loss* (0,652), demonstrando alta capacidade de generalização. O *VGG16* também apresentou desempenho sólido, com acurácia de teste de 72,24%, enquanto a CNN personalizada, embora tenha consumido o maior tempo de treinamento (mais que o dobro dos demais modelos), apresentou resultado competitivo com 71,03% de acurácia no teste, superando os modelos *EfficientNetB0*, *ResNet50* e *ConvNeXtSmall*.

Vale destacar que modelos como *EfficientNetB0* e *ResNet50*, apesar do menor tempo de treinamento, tiveram desempenho significativamente inferior, indicando que a simples transferência dos pesos pré-treinados não garantiu bons resultados para o problema específico. Já o *ConvNeXtSmall*, mesmo com tempo de treinamento elevado, não conseguiu atingir acurárias superiores a 41%, sugerindo a necessidade de adaptações mais

específicas para esta tarefa. Esses achados reforçam a importância do ajuste fino dos modelos, especialmente para aplicações específicas como a classificação de sementes, onde o *fine-tuning* e técnicas adicionais, como aumento de dados, podem melhorar substancialmente o desempenho final.

Tabela 3. Desempenho da CNN Personalizada e Modelos Transfer Learning

Modelo	Tempo (s)	Acurácia Treino (%)	Acurácia Val (%)	Acurácia Teste (%)	Loss Teste
InceptionV3	896,14	90,64	74,91	76,68	0,652
VGG16	880,48	71,60	73,70	72,24	0,744
CNN Personalizada	1814,81	61,88	69,33	71,03	0,867
ResNet50	340,01	29,82	35,39	37,14	1,574
ConvNeXtSmall	936,52	39,55	39,03	40,75	1,365
EfficientNetB0	188,07	20,75	21,82	21,75	1,610

As curvas de aprendizado dos modelos, ilustradas na Figura 2, mostram a evolução da acurácia nos conjuntos de treinamento e validação ao longo das épocas. Nota-se que o *InceptionV3* apresentou o melhor desempenho, com convergência rápida e estável, mantendo alta acurácia tanto no treinamento quanto na validação. Em contraste, o *EfficientNetB0* demonstrou dificuldades de convergência, exibindo oscilações ou estagnação em suas curvas. Os demais modelos, como *ResNet50*, *VGG16* e *ConvNeXtSmall*, apresentaram comportamentos intermediários, com variações distintas em generalização e *overfitting*. Esses resultados sugerem que a arquitetura do *InceptionV3* é a mais adequada para o problema em questão, enquanto o *EfficientNetB0* pode exigir ajustes adicionais para melhorar sua estabilidade.

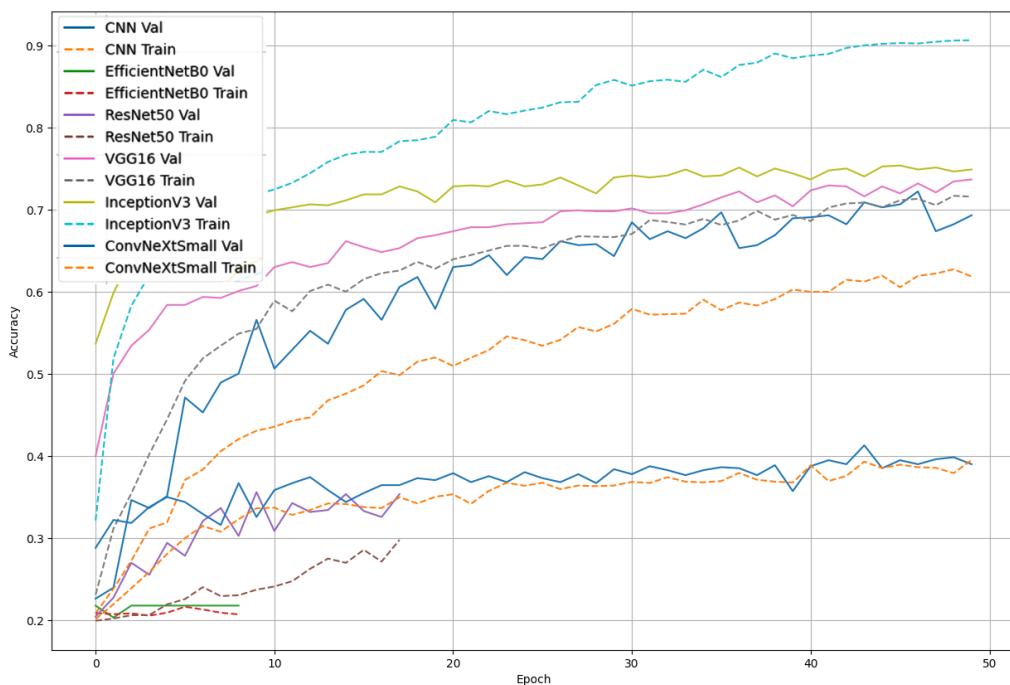


Figura 2. Curvas de acurácia de validação e de treino dos modelos da primeira etapa.

4.1. Fine-Tuning do Modelo InceptionV3

O modelo *InceptionV3*, que demonstrou desempenho superior na etapa inicial de *transfer learning*, foi submetido a um refinamento através de *fine-tuning* para otimizar sua capacidade de classificação. Neste processo, as últimas 20 camadas convolucionais foram descongeladas para ajuste fino dos pesos, enquanto as camadas iniciais permaneceram congeladas para preservar as características genéricas aprendidas no *ImageNet*. O treinamento foi conduzido em duas fases distintas: uma etapa inicial de *transfer learning* (34 épocas, 3860.94 segundos) seguida por *fine-tuning* (50 épocas, 4130.94 segundos), ambas interrompidas pelo critério de *Early Stopping* quando não se observou melhoria contínua. A configuração utilizou taxa de aprendizagem reduzida ($lr = 10^{-5}$) e técnicas de aumento de dados para melhorar a generalização.

A Figura 3 apresenta a evolução das métricas durante todo o processo. Na fase de *transfer learning*, o modelo alcançou 76.68% de acurácia na validação com perda de 0.652. Durante o *fine-tuning*, observou-se uma melhoria progressiva até atingir 84.48% de acurácia e perda de 0.45 na validação. A diferença mínima entre as curvas de treino e validação (inferior a 3%) atesta a eficácia das estratégias de regularização na prevenção de *overfitting*. O tempo médio por época foi de aproximadamente 113 segundos na primeira fase e 83 segundos na segunda, refletindo a complexidade computacional diferenciada de cada etapa.

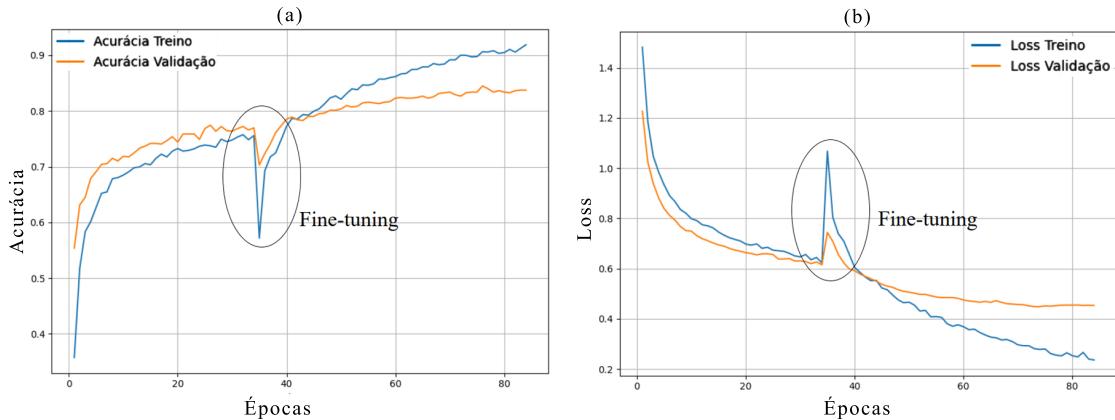


Figura 3. Curvas de aprendizado do InceptionV3 na Etapa 2 durante as fases de transfer learning (34 épocas, 3861s) e fine-tuning (50 épocas, 4131s).
(a) Evolução da acurácia; (b) Evolução da perda. As linhas em azul representam o conjunto de treino e as em laranja, a validação. A demarcação indica o começo do fine-tuning. O treinamento foi interrompido por Early Stopping em ambas as fases.

Na avaliação final com dados de teste, o modelo ajustado alcançou 85.7% de acurácia com perda de 0.415, representando uma melhoria de 9 pontos percentuais em relação ao desempenho inicial. A análise temporal revela que a fase de *transfer learning* demandou 48.2% do tempo total (7991.88 segundos), enquanto o *fine-tuning* consumiu 51.8%, com um custo computacional total justificado pelo ganho de desempenho obtido. A acurácia de teste final superou consistentemente os valores de validação, confirmando a robustez do modelo frente a dados não vistos durante o treinamento. O desempenho detalhado por categoria é apresentado na Tabela 4 após o processo de *fine-tuning*, com métricas específicas para cada categoria de sementes analisadas:

Tabela 4. Desempenho do modelo InceptionV3 com fine-tuning por categoria de semente

Categoria	Precisão	Recall	F1-score	Amostras
Manchada	0.88	0.91	0.89	160
Imatura	0.78	0.83	0.80	170
Intacta	0.85	0.86	0.85	181
Danos na Casca	0.88	0.90	0.89	170
Partida	0.91	0.79	0.84	151
Acurácia		0.86		832
Média Macro	0.86	0.86	0.86	-
Média Pond.	0.86	0.86	0.86	-

Os resultados demonstram que o modelo *InceptionV3* com *fine-tuning* alcançou desempenho equilibrado entre as cinco categorias de sementes analisadas, com acurácia global de 86%. As categorias Manchada (F1-score 0.89) e Danos na Casca (F1-score 0.89) apresentaram os melhores resultados, com valores de *recall* acima ou igual a 90%, indicando que o modelo identifica estes padrões de danos. A categoria Partida obteve a maior precisão (0.91), porém com *recall* moderado (0.79), sugerindo que, embora o modelo seja preciso quando classifica uma semente como partida, eventualmente deixa de identificar alguns casos positivos. A categoria Imatura apresentou o desempenho mais modesto (F1-score 0.80), possivelmente devido à maior variabilidade morfológica nesta classe. Nota-se consistência entre as médias macro e ponderada (ambas 0.86), indicando que o modelo não apresenta viés significativo para classes majoritárias, apesar da leve diferença no tamanho dos grupos (*support* entre 151-181 amostras por classe).

A matriz de confusão é uma representação em forma de tabela que compara as classes previstas pelo modelo com as classes reais, permitindo identificar acertos e erros de classificação. Cada linha da matriz representa as amostras da classe real, enquanto cada coluna representa as amostras previstas. Os valores na diagonal principal indicam o número de classificações corretas para cada classe, enquanto os valores fora da diagonal representam erros, mostrando quais classes foram confundidas entre si.

A Figura 4 apresenta a matriz de confusão obtida na fase de teste do modelo *InceptionV3* após o *fine-tuning*. Observa-se que a maioria das amostras foi corretamente classificada. As classes *Intacta*, *Danos na Casca* e *Manchada* apresentaram os melhores desempenhos, com 155, 153 e 145 acertos, respectivamente. A classe *Partida*, apesar de apresentar um bom número de classificações corretas (119), foi também a que mais gerou confusões com outras classes, especialmente com *Manchada* e *Danos na Casca*, o que pode indicar semelhança visual entre esses danos. A classe *Imatura* teve 141 acertos, mas foi frequentemente confundida com a classe *Intacta* (20 erros), o que sugere que características visuais sutis entre essas classes podem ter dificultado a distinção. Ainda assim, os resultados revelam uma boa capacidade discriminativa do modelo, com poucas confusões significativas entre classes não correlatas.

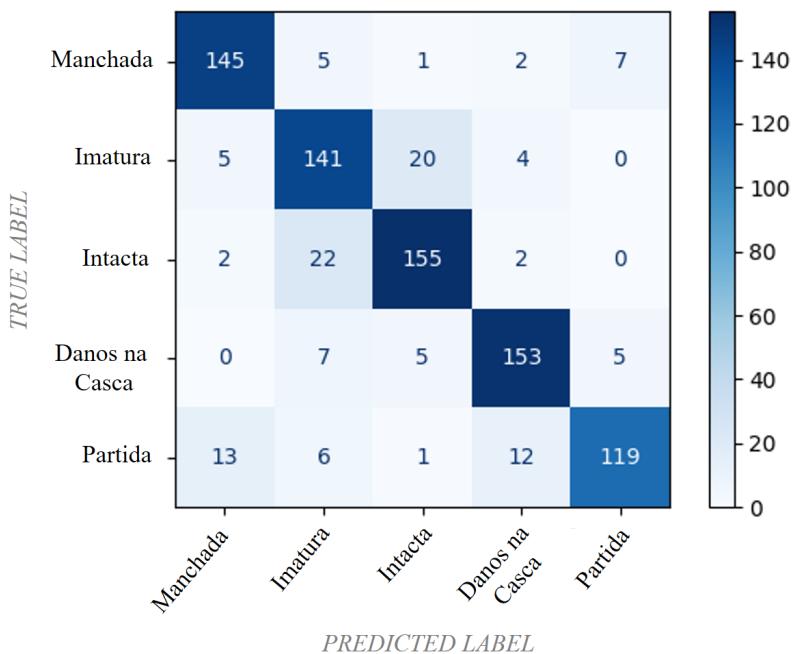


Figura 4. Matriz de confusão na fase de teste para o modelo InceptionV3 após fine-tuning.

A Figura 5 apresenta a distribuição percentual das previsões para cada classe no *dataset* de teste, permitindo observar a taxa de erro relativa em cada categoria.

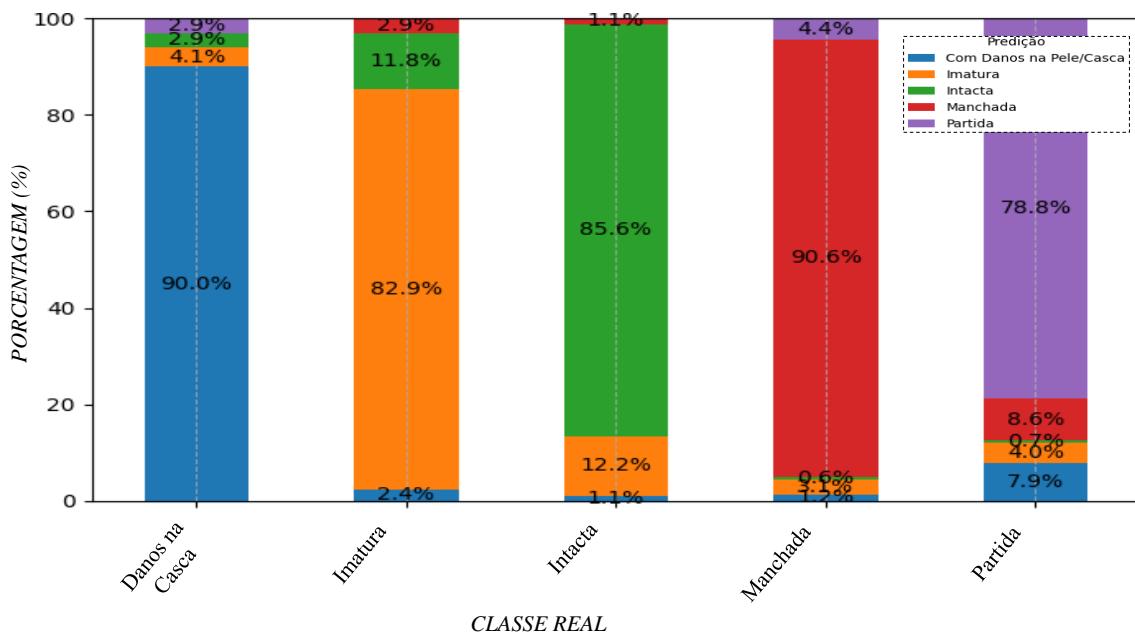


Figura 5. Distribuição percentual das previsões corretas e incorretas por classe real.

A classe Manchada apresentou a maior taxa de acerto proporcional, com 90,6% das instâncias corretamente classificadas, seguida pelas classes Danos na Casca com

90,0% e Intacta com 85,6%. Já a classe Imatura teve 82,9% de acertos, enquanto a classe Partida obteve a menor taxa de acerto proporcional, com 78,8%, indicando maior dificuldade do modelo em identificar corretamente essas amostras.

5. Conclusões

Este trabalho apresentou uma investigação sobre a aplicação de técnicas de *Deep Learning* para classificação automática de sementes de soja, comparando diferentes abordagens de redes neurais convolucionais. Os resultados demonstraram a superioridade da arquitetura *InceptionV3*, que alcançou 76,68% de acurácia na fase inicial de *Transfer Learning*, superando significativamente outros modelos como *VGG16* (72,24%) e *ResNet50* (37,14%). A aplicação de *fine tuning* permitiu melhorar ainda mais o desempenho do modelo, atingindo 86% de acurácia no conjunto de teste, com F1-scores consistentes em todas as classes, variando entre 0,80 para sementes imaturas e 0,89 para sementes manchadas ou com danos na casca.

Embora os resultados sejam promissores, o estudo apresenta algumas limitações importantes. O desempenho em condições não controladas, com variações de iluminação, fundo e ângulo de captura, não foi avaliado. Além disso, o tamanho moderado do *dataset* (5.513 imagens) pode limitar a generalização para outras variedades de soja, e o custo computacional do *fine tuning* pode ser proibitivo para implementações em tempo real em alguns cenários práticos.

Direções para trabalhos futuros incluem (i) a incorporação de técnicas de atenção para melhorar a discriminação de classes desafiantes; (ii) a expansão do *dataset* com imagens de campo e diferentes variedades de soja; (iii) o desenvolvimento de versões otimizadas para implantação em dispositivos móveis; (iv) a exploração de aprendizagem por métrica para reduzir confusões entre classes similares; e (v) a integração com sistemas robóticos para classificação automática em linha de produção.

Referências

- Ayoub Shaikh, T., Rasool, T., and Rasheed Lone, F. (2022). Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*, 198:107119.
- Bevers, N., Sikora, E. J., and Hardy, N. B. (2022). Soybean disease identification using original field images and transfer learning with convolutional neural networks. *Computers and Electronics in Agriculture*, 203:107449.
- Chauhan, I., Kekre, S., Miglani, A., Kankar, P. K., and Ratnaparkhe, M. B. (2025). Cnn-based damage classification of soybean kernels using a high-magnification image dataset. *Journal of Food Measurement and Characterization*, 19(5):3471–3495.
- Eryigit, R. and Tugrul, B. (2021). Performance of various deep-learning networks in the seed classification problem. *Symmetry*, 13(10).
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.
- Huang, Z., Wang, R., Cao, Y., Zheng, S., Teng, Y., Wang, F., Wang, L., and Du, J. (2022). Deep learning based soybean seed classification. *Computers and Electronics in Agriculture*, 202:107393.

- Jogin, M., Mohana, Madhulika, M. S., Divya, G. D., Meghana, R. K., and Apoorva, S. (2018). Feature extraction using convolution neural networks (cnn) and deep learning. In *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pages 2319–2323.
- Lin, W., Fu, Y., Xu, P., Liu, S., Ma, D., Jiang, Z., Zang, S., Yao, H., and Su, Q. (2023). Soybean image dataset for classification. *Data in Brief*, 48:109300.
- Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S. (2022). A convnet for the 2020s.
- Pereira, G. M. L., Foleis, J. H., de Souza Brito, A., and Bertolini, D. (2024). A database for soybean seed classification. In *2024 37th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 1–6.
- Saleem, M. H., Potgieter, J., and Arif, K. M. (2021). Automation in agriculture by machine and deep learning techniques: A review of recent developments. *Precision Agriculture*, 22:2053–2091.
- Siamabele, B. (2021). Soybeans production, driving factors, and climate change perspectives. *Journal for Creativity, Innovation and Social Entrepreneurship (JCISE)*, page 113.
- Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision.
- Tan, M. and Le, Q. V. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *CoRR*, abs/1905.11946.
- Upadhyay, N. and Bhargava, A. (2025). Artificial intelligence in agriculture: applications, approaches, and adversities across pre-harvesting, harvesting, and post-harvesting phases. *Iranian Journal of Computer Science*.
- Vogel, J. T. et al. (2021). Soybean yield formation physiology—a foundation for precision breeding based improvement. *Frontiers in Plant Science*, 12:719706.
- Yang, S., Zheng, L., He, P., Wu, T., Sun, S., and Wang, M. (2021). High-throughput soybean seeds phenotyping with convolutional neural networks and transfer learning. *Plant Methods*, 17(1):50.
- Yun, Y., Li, D., An, X., and Ma, Z. (2024). *Research Progress on Seed Appearance Recognition for Major Crops*, volume 9 of *Smart Agriculture*. Springer, Singapore.
- Zhu, S., Zhang, J., Chao, M., Xu, X., Song, P., Zhang, J., and Huang, Z. (2020). A rapid and highly efficient method for the identification of soybean seed varieties: Hyperspectral images combined with transfer learning. *Molecules*, 25(1).