



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
PRÓ-REITORIA DE PESQUISA, PÓS GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE MATERIAIS E
PROCESSOS INDUSTRIAIS
DISSERTAÇÃO DE MESTRADO**

FERNANDO RODRIGUES RIBEIRO

**CLASSIFICAÇÃO DE DEFEITOS EM GRÃOS DE SOJA COM USO DE VISÃO
COMPUTACIONAL E REDES NEURAIS ARTIFICIAIS**

**TERESINA
2025**

FERNANDO RODRIGUES RIBEIRO

**CLASSIFICAÇÃO DE DEFEITOS EM GRÃOS DE SOJA COM USO DE VISÃO
COMPUTACIONAL E REDES NEURAIS ARTIFICIAIS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Materiais e Processos Industriais do Instituto Federal do Piauí, como parte integrante dos requisitos para a obtenção do título de Mestre em Engenharia de Materiais e Processos Industriais.

Área de concentração: Processamento e Caracterização de Materiais

Orientador: Prof. Dr. Otilio Paulo da Silva Neto

TERESINA

2025

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Ribeiro, Fernando Rodrigues

R484c Classificação de defeitos em grãos de soja com o uso de visão computacional e redes neurais artificiais / Fernando Rodrigues Ribeiro. - 2025.
74 f.: il. color.

Dissertação (Mestrado em Engenharia de Materiais) - Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central, 2025.

Orientador : Prof Dr. Otílio Paulo da Silva Neto.

1. Visão computacional. 2. Redes neurais convolucionais. 3. Classificação de grãos de soja. I. Título.

CDD - 620.1

FERNANDO RODRIGUES RIBEIRO

**CLASSIFICAÇÃO DE DEFEITOS EM GRÃOS DE SOJA COM USO DE VISÃO
COMPUTACIONAL E REDES NEURAIS ARTIFICIAIS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Materiais e Processos Industriais do Instituto Federal do Piauí, como parte integrante dos requisitos para a obtenção do título de Mestre em Engenharia de Materiais e Processos Industriais.

Aprovada em: 27/08/2025.

BANCA EXAMINADORA

Prof. Dr. Otílio Paulo da Silva Neto (Orientador)
Instituto Federal do Piauí (IFPI)

Prof. Dr. Gilvan Moreira da Paz
Instituto Federal do Piauí (IFPI)

Prof. Dr. Rodolpho Carvalho Leite
Instituto Federal do Piauí (IFPI)

Prof^a. Dr^a. Elanne Cristina Oliveira dos Santos
Instituto Federal do Piauí (IFPI)

Dedico este trabalho aos meus pais, ao meu irmão, à minha namorada e a quem esteve ao meu lado com carinho e incentivo durante esta caminhada.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder sabedoria, força e perseverança para concluir mais esta etapa da minha vida.

À minha família, pelo amor, apoio e compreensão incondicionais em todos os momentos, fundamentais para que eu chegasse até aqui.

À minha companheira, pela paciência, incentivo e presença constante durante toda a jornada.

Registro um agradecimento especial ao meu orientador, pela dedicação, paciência, disponibilidade e pelas valiosas orientações que contribuíram decisivamente para o desenvolvimento deste trabalho.

Aos amigos de turma e de trabalho, pelas conversas, pelo companheirismo e pela troca de experiências que tornaram o percurso mais leve e enriquecedor.

À Indústria Dureino, pela disponibilidade e colaboração na coleta das amostras de grãos de soja, sem a qual este estudo não seria possível.

Por fim, agradeço a todos os professores e servidores do IFPI – Campus Teresina Central, que, de forma direta ou indireta, contribuíram para o alcance deste resultado.

RESUMO

A soja é uma das principais commodities agrícolas do Brasil, desempenhando um papel fundamental na economia nacional e global. A classificação da qualidade dos grãos é um processo essencial para a sua comercialização e aplicação industrial, sendo atualmente realizada de forma manual e sujeita a subjetividades e inconsistências. Neste contexto, o presente trabalho propõe o desenvolvimento de um sistema de classificação automática de defeitos em grãos de soja por meio de técnicas de visão computacional e inteligência artificial, com o objetivo de aumentar a precisão e a eficiência do processo. A metodologia empregada consistiu no treinamento e avaliação de diferentes arquiteturas de redes neurais convolucionais (CNNs). Foram testadas redes customizadas, além dos modelos pré-treinados ResNet50 e InceptionV3, com o intuito de comparar seu desempenho em termos de acurácia e aplicabilidade prática. Os resultados demonstraram que as redes pré-treinadas superaram significativamente as redes customizadas. O modelo InceptionV3 destacou-se como a melhor alternativa, atingindo 95,9% de acurácia e apresentando alta capacidade de generalização, mostrando-se uma opção viável para implementação em ambientes industriais. O modelo ResNet50, com 63,1% de acurácia, mostrou-se adequado para aplicações que exigem um equilíbrio entre desempenho e custo computacional. Já as CNNs customizadas apresentaram desempenho insatisfatório, alcançando no máximo 16,6% de acurácia, o que indica a necessidade de aprimoramento da arquitetura e do treinamento. Nesta etapa do processo, conclui-se que a aplicação de redes neurais convolucionais para a classificação de grãos de soja tem grande potencial para modernizar e otimizar o setor agroindustrial. Destaca-se ainda a criação de um dataset com imagens de grãos de soja contendo diversos defeitos, que poderá subsidiar futuras pesquisas na área de visão computacional aplicada à análise de grãos.

Palavras-chave: visão computacional; redes neurais convolucionais; classificação de grãos de soja.

ABSTRACT

Soybean is one of Brazil's main agricultural commodities, playing a fundamental role in both the national and global economy. Grain quality classification is an essential process for its commercialization and industrial application, currently performed manually and subject to subjectivities and inconsistencies. In this context, this study proposes the development of an automatic defect classification system for soybean grains using computer vision and artificial intelligence techniques, aiming to increase the accuracy and efficiency of the process. The methodology employed consisted of training and evaluating different convolutional neural network (CNN) architectures. Custom networks were tested, along with the pre-trained models ResNet50 and InceptionV3, to compare their performance in terms of accuracy and practical applicability. The results demonstrated that pre-trained networks significantly outperformed custom networks. The InceptionV3 model stood out as the best alternative, achieving 95.9% accuracy and demonstrating high generalization capability, making it a viable option for implementation in industrial environments. The ResNet50 model, with 63.1% accuracy, proved suitable for applications requiring a balance between performance and computational cost. Meanwhile, the custom CNNs showed unsatisfactory performance, reaching a maximum accuracy of 16.6%, indicating the need for improvements in architecture and training. At this stage, it is concluded that the application of convolutional neural networks for soybean grain classification has great potential to modernize and optimize the agro-industrial sector. It is also worth highlighting the creation of a dataset containing images of soybean grains with various defects, which may support future research in the field of computer vision applied to grain analysis.

Keywords: computer vision; convolutional neural networks; soybean grain classification.

LISTA DE ILUSTRAÇÕES

Figura 1 - Produção mundial de soja por país (2023/2024), em milhões de toneladas.....	14
Figura 2 - Destinos e usos da soja brasileira produzida na safra de 2022.....	15
Figura 3 - Arquitetura de uma CNN.....	30
Figura 4 - Esquema de Aprendizado de uma CNN.....	30
Figura 5 - Câmara para captação de imagens, fechada.....	42
Figura 6 - Câmara para captação de imagens, fechada: (a) Iluminação desligada, (b) Iluminação ligada.....	43
Figura 7 - Grãos de soja por categorias: (a) ardidos, (b) bons, (c) esverdeados, (d) mofados, (e) picados, e (f) queimados.....	44
Figura 8 - Exemplos de imagens de (a) esverdeado, (b) queimado.....	47
Figura 9 - Matriz de confusão do modelo 3.....	56
Figura 10 - Matriz de confusão do modelo ResNet50.....	57
Figura 11 - Gráficos de acurácia e perda (loss) do modelo ResNet50 durante o treinamento e validação.....	58
Figura 12 - Matriz de confusão do modelo InceptionV3.....	59
Figura 13 - Gráficos de acurácia e perda (loss) do modelo InceptionV3 durante o treinamento e validação.....	60
Figura 14 - Tela inicial do aplicativo com opções para captura ou seleção de imagem para análise.....	64
Figura 15 - Tela com imagem selecionada e opções de classificação e segmentação disponíveis.....	64
Figura 16 - Tela com imagem mostrando contorno e segmentação dos grão.....	65
Figura 17 - a) histórico de análises e b) últimas análises.....	65
Figura 18 - Tela de Análise Individual: Resumo por Classe e Distribuição Quantitativa.....	66

LISTA DE TABELAS

Tabela 1 - Descrição Sumária dos estádios reprodutivos da soja.....	21
Tabela 2 - Descrição dos Grãos Avariados.....	24
Tabela 3 - Principais métricas de avaliação de redes neurais e suas fórmulas.....	37
Tabela 4 - Número de imagens coletadas por categoria de grãos.....	43
Tabela 5 - Distribuição do Dataset nas Etapas de Treinamento, Validação.....	46
Tabela 6 - Resultados de Acurácia Geral e Tempo de Treinamento.....	54
Tabela 7 - Métricas de desempenho por classe do melhor modelo customizado.....	56
Tabela 8 - Métricas de desempenho por classe do modelo ResNet50	58
Tabela 9 - Métricas de desempenho por classe do modelo InceptionV3.....	60
Tabela 10 - Estudos sobre Classificação de Defeitos em Grãos de Soja utilizando Redes Neurais e Visão Computacional.....	61

LISTA DE QUADROS

Quadro 1 - Resumos das CNNs personalizadas utilizadas.....	50
Quadro 2 - Resumos das CNNs pré-treinadas utilizadas.....	51

LISTA DE ABREVIATURAS E SIGLAS

CNN	Rede Neural Convolucional
FAO	<i>Food and Agricultural Organization</i>
IA	Inteligência Artificial
IFPI	Instituto Federal de Educação, Ciência e Tecnologia do Piauí
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
USDA	Departamento de Agricultura dos Estados Unidos
VC	Visão Computacional

SUMÁRIO

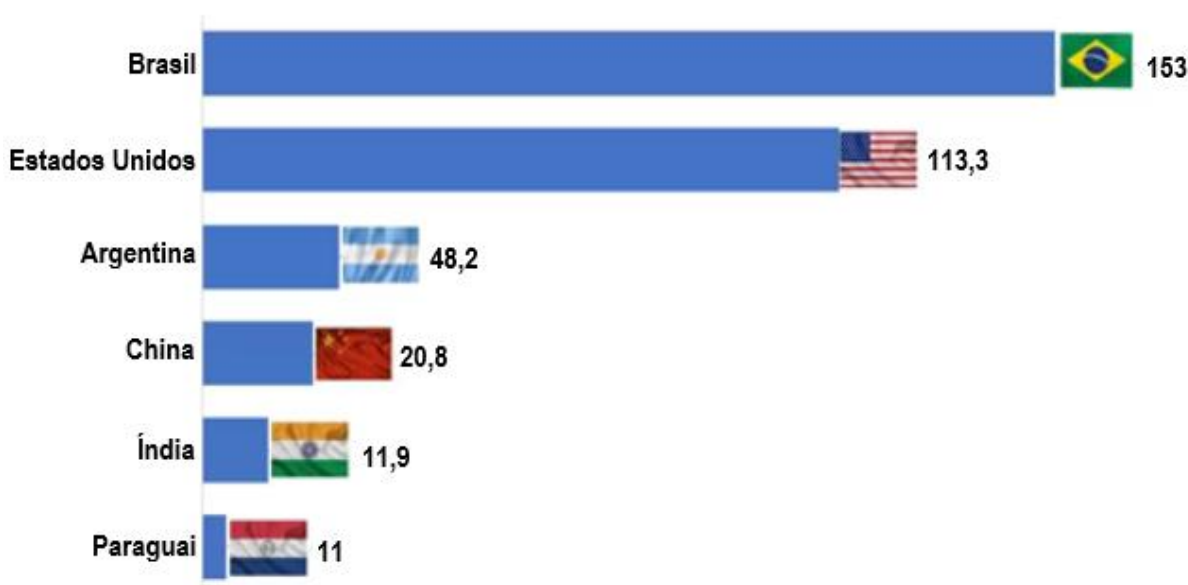
1	INTRODUÇÃO	14
1.1	JUSTIFICATIVA	16
1.2	OBJETIVO	18
1.2.1	Objetivo Geral	18
1.2.2	Objetivos Específicos	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	CULTURA DA SOJA	19
2.2	FENOLOGIA DA CULTURA DA SOJA	20
2.3	DEFEITOS NOS GRÃOS/CLASSIFICAÇÃO DOS DEFEITOS	22
2.4	AVALIAÇÃO DOS GRÃOS	25
2.5	VISÃO COMPUTACIONAL E REDES NEURAI	26
2.6	REDES NEURAI	27
2.7	REDES NEURAI CONVOLUCIONAIS	28
2.7.1	Camadas das Redes Neurais Convolucionais	29
2.7.1.1	<i>Camada Convolucional</i>	31
2.7.1.2	<i>Camada de Unidade Linear Retificada (ReLU)</i>	31
2.7.1.3	<i>Camada de Pooling</i>	31
2.7.1.4	<i>Camada Totalmente Conectada (FC)</i>	32
2.7.1.5	<i>Camada de Normalização por Lote (Batch Normalization)</i>	32
2.7.1.6	<i>Camada de Concatenação (Concatenate)</i>	32
2.7.1.7	<i>Camada de Adição (Add Layer)</i>	33
2.8	ARQUITETURAS DE REDES PRÉ-TREINADAS: INCEPTIONV3 E RESNET50	33
2.8.1	InceptionV3	33
2.8.2	ResNet50	34
2.9	DETECÇÃO DE OBJETOS COM YOLO (YOU ONLY LOOK ONCE)	34
2.10	AVALIAÇÃO DE DESEMPENHO EM MODELOS DE CLASSIFICAÇÃO	35
2.10.1	Validação Cruzada	36
2.10.2	Matriz de Confusão	36
2.10.3	Métricas de Avaliação de Desempenho	36

2.11	APLICAÇÕES DAS REDES NEURAIS CONVOLUCIONAIS.....	38
3	METODOLOGIA.....	41
3.1	COLETA DE DADOS.....	41
3.1.1	Obtenção das Amostras.....	41
3.2	TRATAMENTO DOS DADOS.....	46
3.2.1	Segmentação com YOLO.....	46
3.2.2	Pré-processamento.....	46
3.2.3	Processamento.....	48
3.3	ARQUITETURAS TESTADAS.....	48
3.4	TREINAMENTO DOS MODELOS.....	48
3.5	FERRAMENTAS UTILIZADAS.....	49
3.6	COMPARAÇÃO DE ARQUITETURAS.....	50
3.7	APLICATIVO DE CLASSIFICAÇÃO.....	51
3.7.1	Estrutura e Desenvolvimento.....	51
4	RESULTADOS E DISCUSSÃO.....	54
4.1	AVALIAÇÃO DO DESEMPENHO DOS MODELOS: IMPLICAÇÕES E UTILIZAÇÕES PRÁTICAS.....	55
4.2	COMPARAÇÃO COM TRABALHOS RELACIONADOS.....	61
4.3	TESTE DO APLICATIVO.....	63
4.4	LIMITAÇÕES.....	66
5	CONSIDERAÇÕES FINAIS.....	68
	REFERÊNCIAS.....	70

1 INTRODUÇÃO

A soja (*Glycine max (L.) Merril*) é considerada como uma das principais commodities brasileiras e ocupa destaque nacional e mundial no contexto econômico como importante produto agrícola. Na safra de 2023/2024 a produção mundial de soja foi de 395,91 milhões de toneladas. O Brasil ocupa a posição de maior produtor mundial de soja, produzindo cerca de 147,35 milhões de toneladas nesse mesmo ciclo, ficando a frente dos Estados Unidos, que produziu 113,34 milhões e toneladas (Agroadvance, 2023; Embrapa, 2025). A Figura 1 ilustra a produção mundial de soja por país na temporada 2023/2024, em milhões de toneladas.

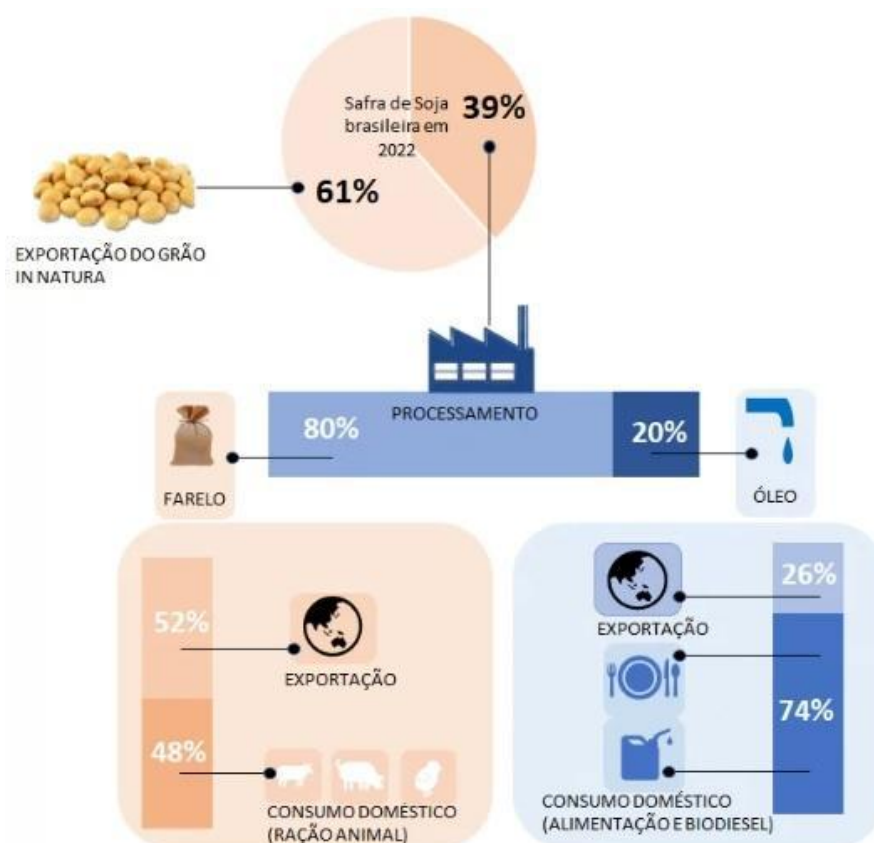
Figura 1 - Produção mundial de soja por país (2023/2024), em milhões de toneladas.



Fonte: USDA (2025).

Sendo uma importante fonte de proteína vegetal, os grãos de soja se apresentam como produtos versáteis em que se destacam por sua ampla aplicação, como uso na alimentação humana, alimentação animal, indústrias químicas, e produção de biodiesel. No Brasil, os grãos têm seu destino e uso mostrado na Figura 2, em que destacam a exportação dos grãos in natura, e a produção de farelo e óleo (Agroadvance, 2023).

Figura 2 - Destinos e usos da soja brasileira produzida na safra de 2022.



Fonte: Agroadvance (2023).

Assim, a existência de um padrão que garanta o controle da qualidade do grão se torna indispensável para a sua avaliação para que seja possível definir sua aplicação de uso e ainda servir como fator de validação técnica para sua comercialização (Lorini, 2018). No Brasil, a classificação da soja é regulamentada pela Instrução Normativa Nº 11, de 15 de maio de 2007 e Instrução Normativa Nº 37 de 27 de julho de 2007, do Ministério da Agricultura, Pecuária e Abastecimento (Brasil, 2007a, 2007b).

Na prática, as indústrias de beneficiamento de grãos realizam a classificação de forma manual, por profissionais especializados. É realizada com a inspeção visual de grão a grão da amostra selecionada. Esse processo se torna repetitivo, demorado e exaustivo, o que pode levar, devido ao cansaço e aliado ao fator subjetivo, a resultados diferentes de classificação.

Os grãos com defeitos apresentam alterações físicas e/ou químicas que afetam e limitam seu uso para processamento. Atualmente a classificação é feita de forma manual, de caráter visual, em que um especialista realiza a avaliação qualitativa e, posteriormente, mediante contabilização dos defeitos na amostra, determina a qualidade da mesma de forma quantitativa.

Com os avanços tecnológicos aplicados à indústria, a chamada Agricultura 4.0 também oferece possibilidades de transformação no setor do agronegócio, integrando inovações como inteligência artificial, internet das coisas (IoT), automação e análise avançada de dados. Esse novo paradigma possibilita um aumento significativo na produtividade, otimiza a utilização de recursos e reduz desperdícios, além de aprimorar a eficiência operacional e a qualidade dos produtos agrícolas. (BIP BRASIL, 2022).

Dessa forma, a adoção de tecnologias inteligentes tem o potencial de contribuir para a modernização dos processos, tornando-os mais precisos, sustentáveis e economicamente viáveis. Assim, a implementação da Inteligência Artificial e pela Visão Computacional (VC), é considerada um fator importante para a crescente adoção da agricultura de precisão. Nesse sentido, torna-se possível a obtenção de muitas informações com riqueza de detalhes a partir de imagens, por meio do processamento em computador, e com isso possibilita a execução de tarefas inteligentes que se aproximam da inteligência humana (Milano, 2010; Agencia EY, 2022).

A proposta de classificação automática dos grãos de soja visa utilizar visão computacional, com processamento de imagens, e técnicas de redes neurais artificiais de forma a aumentar a confiabilidade e reduzir o tempo de classificação e análise.

1.1 JUSTIFICATIVA

A avaliação da qualidade dos grãos de soja, tradicionalmente realizada por meio de inspeção visual manual, apresenta limitações significativas que comprometem a consistência e a eficiência do processo. Entre os principais problemas estão a subjetividade inerente ao avaliador, a variabilidade entre diferentes operadores, a morosidade na execução das análises, além de questões ergonômicas e de infraestrutura laboratorial.

A classificação é regulamentada pela Instrução Normativa nº 11 do MAPA, que delimita critérios para categorização dos grãos em classes leves e graves — subdivididas em ardidos, mofados e queimados — exigindo do avaliador alto nível de conhecimento técnico para distinguir visualmente defeitos muitas vezes sutis. Em situações de alta demanda ou diante de avaliadores menos experientes, o processo torna-se ainda mais suscetível a erros e inconsistências.

Nos últimos anos, diversas abordagens automatizadas utilizando técnicas de processamento de imagens e aprendizado de máquina foram propostas para mitigar tais limitações. No entanto, muitos desses sistemas enfrentam desafios recorrentes, como dependência de condições rígidas de iluminação, dificuldade em generalizar para diferentes variedades de grãos ou ambientes de captura, além de desempenho reduzido diante de defeitos menos visíveis ou combinados.

Modelos mais simples tendem a apresentar baixa acurácia frente à variabilidade natural dos grãos, enquanto soluções mais complexas frequentemente demandam hardware especializado ou grandes volumes de dados anotados — recursos nem sempre disponíveis em ambientes produtivos, especialmente em pequenas e médias estruturas agrícolas.

Diante desse cenário, a proposta deste trabalho é desenvolver um algoritmo otimizado para a classificação automática de grãos de soja que supere tanto as limitações do processo manual quanto os principais entraves observados nas soluções computacionais existentes. Busca-se, assim, construir um sistema capaz de operar com maior robustez em diferentes condições de captura, reduzir o tempo de classificação, padronizar os resultados, melhorar as condições de trabalho do avaliador e ampliar a escalabilidade do processo.

Adicionalmente, será estruturado um banco de dados abrangente que não apenas suporte o treinamento dos modelos, mas que também complemente a aplicação prática da Instrução Normativa nº 11, servindo como base para futuras pesquisas e aplicações industriais.

1.2 OBJETIVO

1.2.1 Objetivo Geral

Realizar uma análise comparativa de diferentes estratégias baseadas em redes neurais convolucionais aplicadas à classificação de defeitos em grãos de soja, desenvolvendo e validando uma aplicação computacional capaz de automatizar o processo de avaliação com base em imagens, utilizando uma base de dados construída especificamente para esse propósito.

1.2.2 Objetivos Específicos

- Coletar amostras de grãos de soja com defeitos (nas classes ardidos, mofados, picados, queimados, verdes e bons) que possam ser classificados, dentro dos parâmetros estabelecidos pelo Ministério da Agricultura e Pecuária (MAPA);
- Estudar e analisar resultados de trabalhos correlatos para desenvolvimento do projeto;
- Formar e compilar um banco de dados para desenvolvimento do projeto e disponibilizá-la para outros trabalhos que possam ser desenvolvidos;
- Desenvolver um sistema de visão computacional com implemento de redes neurais artificiais que seja capaz de classificar os grãos por tipo de defeitos ou não, utilizando técnicas de inteligência artificial e processamento de imagem;
- Avaliar a acurácia encontrada com os algoritmos das redes neurais convolucionais;

2 FUNDAMENTAÇÃO TEÓRICA

Embora existam vários trabalhos que abordam a classificação de grãos em suas várias espécies para observação de avarias, ainda existe uma lacuna de estudos que busquem realizar a classificação dos defeitos dos grãos de soja por meio de ferramentas computacionais e de inteligência artificial na literatura. Para embasar este estudo, neste capítulo serão apresentados conceitos e aspectos relacionados ao contexto da pesquisa que servirão de arcabouço para compreensão do tema e aplicação da solução proposta.

2.1 CULTURA DA SOJA

A soja (*Glycine Max* L.) é uma cultura que pertence à família da *Fabaceae* e que é cultivada praticamente em todas as regiões do Brasil. Na região Norte ela se concentra nos estados do Tocantins, Roraima, Rondônia e o sul do Pará. Já nos estados da região Nordeste, é produzida em alta escala nas unidades federativas do Piauí, Maranhão, Alagoas e Bahia. Nas regiões Centro-Oeste, Sudeste e Sul, o cultivo está presente em quase todos os estados (Embrapa, 2005).

Com o avanço da ciência e da tecnologia, a produtividade da soja chega a patamares elevados. Com o fomento das organizações públicas e privadas, o Brasil detém alta tecnologia na produção de grãos que traz um produto de alta qualidade e que tenha competitividade no mercado internacional. O Brasil lidera na produção mundial da cultura, logo à frente dos Estados Unidos da América e da Argentina (Agroadvance, 2023).

O Brasil ocupa a primeira colocação do *ranking* com 156 milhões de toneladas na safra 2023/2024, perfazendo a soja como carro chefe do cenário do agronegócio brasileiro (Embrapa, 2025). A evolução do agronegócio brasileiro permitiu que as propriedades se tornassem empreendimentos financeiros, com tecnologia aplicada no uso eficiente da terra que possibilitou a soja estar entre as *commodities* internacionais. A alta produtividade da cultura teve como um dos principais fatores o uso de sementes providas de melhoramento genético (Seixas et al., 2020).

A soja é uma planta que possui vários subprodutos, dentre os quais o principal é o grão. O grão de soja é empregado em diversos processos na agroindústria,

servindo como alimento (humanos e animais), na indústria de cosméticos, lubrificantes industriais, entre outros. O grão da cultura é o principal produto no mercado internacional e necessita que esteja dentro do padrão internacional de comercialização (Silva et al., 2022).

A produção de soja (grão) no Brasil está centrada nos estados de Mato Grosso, Paraná, Goiás, Rio Grande do Sul e Mato Grosso do Sul, os quais representam aproximadamente 73% da produção nacional. O potencial agrícola da soja tem como uma das condições a extensa área de plantio, que ultrapassa 40 milhões de hectares. O crescimento das áreas produtivas tem como foco a área denominada MATOPIBA, compreendendo os estados do Maranhão, Tocantins, Piauí e Bahia. A substituição de atividades menos atrativas pela cultura da soja cria um espaço para áreas agricultáveis (MAPA, 2023).

2.2 FENOLOGIA DA CULTURA DA SOJA

A soja é uma planta de hábito herbáceo, anual, ereta possuindo o sistema radicular pivotante e raízes secundárias que podem se associar à bactéria *Rhizobium japonicum* que contribui para a fixação biológica do nitrogênio. Como pertence ao estrato herbáceo, sua altura varia de 30 a 50 centímetros, com números de ramificações distintos e sua floração é do tipo racemo curto, dispondo de 10 flores em cada. As vagens podem carregar de 2 a 5 grãos, e o tempo de colheita (precoce) em torno de 75 dias (Sediyama et al., 2009).

A frutificação da cultura é uma das etapas mais esperadas do processo produtivo. É nessa fase que os grãos são formados e posteriormente disponibilizados à venda no mercado. No entanto, o desenvolvimento depende de alguns fatores como a disponibilidade de nutrientes, o manejo adequado, a isenção de pragas e doenças, um material genético de alta produtividade, condições edafoclimáticas entre outros (Câmara, 1998).

Para uma produção de grãos satisfatória é necessário que a planta esteja em condições adequadas para a realização da fotossíntese. A fotossíntese é o principal processo metabólico no qual a planta produz seu próprio alimento através da luz solar. Nesse metabolismo, a energia produzida é usada em outros processos metabólicos secundários que são essenciais ao desenvolvimento da planta, como é na etapa da

frutificação. Todas as cultivares de soja apresentam os estádios desenvolvimento representados pela letra V (fase vegetativa) e R (fase reprodutiva) (Fehr; Caviness, 1977). Os estádios reprodutivos da soja abordam oito fases, agrupadas em: florescimento (R1 e R2), desenvolvimento da vagem (R3 e R4), desenvolvimento do grão (R5 e R6) e maturação da planta (R7 e R8). A etapa do florescimento é caracterizada pelo aparecimento do primeiro botão floral e tem finalização com a floração completa. O início da fase de desenvolvimento das vagens é marcado pela formação da vagem de 5 mm, tornando-se um parâmetro da quantidade de vagens cada planta irá ter. O desenvolvimento dos grãos se dá ao final da fase R4 e iniciação da R5 que é descrito pela presença de pelo menos uma vagem com um grão de 3 mm presente num dos quatro nós superiores, sendo nessa fase o enchimento das vagens e redistribuição de nutrientes (Neumaier et al., 2000).

A última fase consiste na maturação dos grãos a campo. A vagem completa traduz a maturação fisiológica da planta que ainda ocorre acúmulo de matéria seca e a coloração verde vai enfraquecendo sua tonalidade. É nesta etapa que 95% das vagens estão preenchidas e aptas ao processo de colheita. A tabela 1 apresenta uma descrição sucinta das fases de desenvolvimento da soja (Thomas e Costa, 2010).

Tabela 1 – Descrição Sumária dos estádios reprodutivos da soja

Estádio	Denominação	Descrição
R1	Início do Florescimento	Uma flor aberta em qualquer nó da haste principal;
R2	Florescimento Pleno	Uma flor aberta num dos 2 últimos nós da haste principal com folha completamente desenvolvida;
R3	Início da formação da vagem	Vagem com 5 mm de comprimento num dos 4 últimos nós da haste principal, com folha completamente desenvolvida;
R4	Vagem completamente desenvolvida	Vagem com 2 cm de comprimento num dos 4 últimos nós da haste principal, com folha completamente desenvolvida;
R5	Início do enchimento do grão	Grão com 3 mm de comprimento em vagem num dos 4 últimos nós da haste principal, com folha completamente desenvolvida;
R6	Grão verde ou vagem cheia	Uma vagem contendo grãos verdes preenchendo as cavidades da vagem de um dos 4 últimos nós da haste principal, com folha completamente desenvolvida;
R7	Início da maturação	Uma vagem normal na haste principal com coloração de madura
R8	Maturação plena	95% das vagens com coloração de madura

OBS: Os últimos nós referem-se aos superiores.

Fonte: Thomas e Costa (2010).

2.3 DEFEITOS NOS GRÃOS/CLASSIFICAÇÃO DOS DEFEITOS

A soja possui um mercado já consolidado, que abrangem seus derivados, incluindo, principalmente, o farelo, destinado à nutrição animal, óleo para alimentação humana, além de centenas de produtos que possuem ingredientes a base de soja em sua formulação.

Na alimentação humana, as bebidas à base de soja (BBS) têm ganhado destaque e registram um crescimento expressivo no mercado. Além de atenderem aos novos conceitos de alimentação saudável, são uma excelente alternativa para consumidores com intolerância à lactose. No âmbito da produção de biodiesel, o óleo de soja tem se mostrado o principal impulsionador do sucesso do Programa Nacional de Produção e Uso de Biodiesel (PNPB) (Hirakuri et al., 2018).

Este cenário revela que a qualidade do produto, aliada à eficiência dos processos logísticos, é fundamental para aumentar a competitividade e promover a sustentabilidade da cadeia produtiva de grãos em um mercado cada vez mais exigente (Hirakuri et al., 2018).

Neste contexto, a qualidade dos grãos de soja torna-se um fator determinante para sua destinação final, seja para consumo humano, produção de biodiesel ou outros usos. No Brasil, a classificação dos grãos de soja é regulamentada pelas Instruções Normativas Nº 11 e Nº 37 de 2007, ambas do Ministério da Agricultura, Pecuária e Abastecimento (MAPA).

Essas normativas visam garantir que os grãos comercializados atendam aos padrões de qualidade exigidos pelo mercado, possibilitando o reconhecimento dos produtos de melhor qualidade e a conformidade com as exigências de segurança alimentar. Para tanto, as normativas estabelecem padrões de qualidade baseados em defeitos observáveis, além de limites específicos para sua comercialização (Brasil, 2007a, 2007b). Esses regulamentos garantem que os grãos adquiridos atendam às exigências do mercado, proporcionando uma padronização que facilita a escolha de fornecedores e o controle de qualidade.

A avaliação da qualidade dos grãos de soja é multifacetada, abrangendo desde propriedades físico-químicas até compostos bioativos e características intrínsecas, como a presença de antioxidantes e compostos lipídicos. Estudos indicam que os grãos de soja contêm carotenoides, tocoferóis e tocotrienóis, compostos conhecidos por seus benefícios à saúde, como a prevenção de danos celulares, doenças

cardiovasculares, envelhecimento acelerado e até mesmo câncer e diabetes. No entanto, esses benefícios só podem ser preservados se o grão mantiver sua integridade física e química ao longo de todo o processo de colheita e armazenamento (Ramos, 2019).

Um dos maiores desafios na manutenção da qualidade dos grãos de soja é o manejo adequado durante a pós-colheita, especialmente nas etapas de secagem e armazenamento. Os grãos são geralmente colhidos com aproximadamente 22% de umidade, tornando a secagem uma etapa essencial para evitar a deterioração, que pode comprometer tanto a qualidade quanto a segurança alimentar (Paraginski et al., 2015).

Fatores como a temperatura ambiente, umidade relativa do ar e o teor de impurezas presentes no grão afetam diretamente sua qualidade durante o armazenamento. Estudos sugerem que, uma vez armazenados, a qualidade dos grãos não pode ser melhorada, apenas mantida. Assim, o manejo correto das condições de secagem e armazenamento é crucial para evitar a proliferação de defeitos (Ramos, 2019).

A classificação da soja, segundo as diretrizes do MAPA, ocorre com base na identificação de defeitos nos grãos colhidos, que, por sua vez, determinam descontos na comercialização quando os limites regulamentares são excedidos. Assim, os principais defeitos dos grãos de soja podem ser agrupados nas seguintes categorias: grãos avariados; amassados; partidos e quebrados; esverdeados; com manchas no tegumento; matérias estranhas e impurezas (Hirakuri et al., 2018).

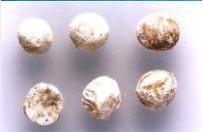
Os defeitos dos grãos também podem ser classificados como graves e ou leves. Esses defeitos, resultantes de condições adversas tanto no campo quanto durante o manejo pós-colheita, impactam diretamente no valor comercial e na adequação dos grãos para diferentes finalidades. A classificação de graves ou leves inclui defeitos de diferentes categorias. Os defeitos graves são: os grãos queimados, ardidos e mofados. Já os leves são: os grãos fermentados, esverdeados, germinados, danificados, imaturos, chochos, amassados, partidos e quebrados (Brasil, 2007a, 2007b; Hirakuri et al., 2018).

As normativas brasileiras estabelecem que a soja pode ser desclassificada e ter sua comercialização proibida se apresentar defeitos graves superiores a 12% para a soja destinada à alimentação humana, ou superiores a 40% para outros usos. Além

disso, características como odor impróprio, presença de insetos ou sementes tóxicas também inviabilizam o uso do produto (Brasil, 2007a, 2007b; Hirakuri et al., 2018).

Este trabalho se concentrará nos grãos avariados. Esta categoria inclui grãos que apresentam danos visíveis e que afetam sua qualidade. Os principais tipos de grãos avariados são: ardidos, mofados, fermentados, danificados, imaturos, chochos, germinados e queimados. A tabela 2 apresenta a descrição de cada um dos defeitos e sua tolerância na comercialização.

Tabela 2: Descrição dos Grãos Avariados

Defeito	Classificação	Descrição	Grão*
Ardido	Grave	Visivelmente fermentados, de coloração marrom escura, com impacto no cotilédono;	
Choco	Leve	Com atrofia interna e enrugados, com formato irregular;	
Danificado	Leve	Afetados por manchas, deformações, perfurações ou ataques de doenças ou insetos em qualquer estágio de desenvolvimento;	
Fermentado	Leve	Alterados em razão da fermentação, com mudança de cor no cotilédono, distinta daquela dos grãos ardidos;	
Germinado	Leve	Evidenciam a emissão da radícula, indicando o início do processo de germinação;	
Imaturo	Leve	Com formato oblongo, intensamente verdes e enrugados, indicando que não atingiram o desenvolvimento fisiológico completo;	
Mofado	Grave	Apresentam fungos (mofo ou bolor) visíveis a olho nu;	
Queimado	Grave	Carbonizados.	

Fonte: MAPA (2008).

Com base nas características elencadas na Tabela 2, os grãos são identificados e de acordo com a tolerância de comercialização, que define se a soja pertence ao Grupo I, soja destinada ao consumo in natura; ou Grupo II, soja destinada a outros usos. Para tanto, é preciso utilizar um mecanismo de avaliação dos grãos eficiente na

diferenciação entre os grãos de qualidade adequada e os defeituosos. Estes mecanismos de avaliação serão discutidos na próxima seção.

2.4 AVALIAÇÃO DOS GRÃOS

A avaliação de grãos de soja tradicionalmente utiliza métodos visuais e manuais que permitem classificar os grãos de acordo com padrões estabelecidos para o consumo e processamento. Este processo é fundamental para garantir a qualidade e a uniformidade dos lotes destinados ao mercado, eliminando grãos considerados fora do padrão devido a defeitos como impurezas, tamanho inadequado, presença de doenças e danos físicos, que comprometem o valor e a aplicação do produto (Silva et al., 2019).

No método tradicional, a inspeção visual é o recurso mais comum, sendo realizado com o auxílio de peneiras e balanças de precisão, juntamente com a análise subjetiva dos operadores. Este tipo de avaliação exige que os profissionais estejam capacitados para identificar diferentes tipos de defeitos, entre eles os danos mecânicos, quebras, manchas, grãos fermentados e ardidos, e alterações de coloração, além de outras deformações que indicam perda de qualidade (Freitas, Almeida, 2020). Além disso, essa técnica pode envolver a separação manual dos grãos, uma prática que, embora eficaz, apresenta limitações em relação à objetividade e à precisão, principalmente em grandes escalas de produção (Mendes, Carvalho, 2021).

Para a indústria agrícola, a avaliação tradicional dos grãos de soja é considerada um processo essencial, mas tem sido cada vez mais questionada quanto à sua eficiência e reprodutibilidade. A subjetividade intrínseca ao processo pode levar a variações nos resultados de classificação, que dependem diretamente da experiência e da percepção do operador. Dessa forma, as limitações do método tradicional impulsionam a busca por alternativas que garantam maior uniformidade e precisão na separação dos grãos fora do padrão, especialmente em volumes elevados e demandas por uma análise rápida (Souza et al., 2023).

2.5 VISÃO COMPUTACIONAL

A visão computacional é uma área em crescente desenvolvimento na ciência da computação, com raízes que remontam a mais de meio século, quando surgiram áreas fundamentais como a inteligência artificial (IA) e o processamento digital de imagens. Muitos autores ainda consideram a visão computacional uma subárea da IA ou uma extensão do processamento de imagens; no entanto, ela se consolida cada vez mais como uma área independente, com interface direta com ambas (Szeliski, 2022).

Seu foco central está na extração de informações significativas a partir de imagens digitais, as quais podem ser capturadas por diversos dispositivos, desde câmeras convencionais e smartphones até aparelhos mais avançados, como microscópios e tomógrafos. A disponibilidade de imagens geradas por esses dispositivos tornou a visão computacional uma ferramenta valiosa para automação e análise visual em diversos campos (Szeliski, 2022).

O objetivo da visão computacional é desenvolver algoritmos que permitam aos computadores analisar imagens e extrair informações de maneira autônoma ou semi autônoma. Esse campo se assemelha ao sistema visual humano, embora, em muitos casos, não dependa de uma limitação direta dos processos biológicos da visão humana (Barelli, 2018).

Os engenheiros da área frequentemente desenvolvem métodos inspirados em outras abordagens técnicas, utilizando desde técnicas de filtragem e reconhecimento de padrões até métodos de análise estrutural de imagem. Ainda que as capacidades humanas de identificação visual sejam extremamente complexas e difíceis de replicar, os avanços na visão computacional já permitem a solução de diversos problemas específicos, como o reconhecimento de objetos e a análise de imagens médicas, áreas nas quais o campo tem obtido grande sucesso (Ansari, 2020).

A visão computacional se beneficia também do processamento digital de imagens, que serve como base para o tratamento inicial das imagens, melhorando aspectos como contraste e nitidez antes de realizar análises mais complexas. A distinção entre as áreas é que, enquanto o processamento digital de imagens tende a retornar imagens aprimoradas, a visão computacional transforma as imagens em dados analíticos que podem ser usados em processos de tomada de decisão. Essa transformação permite a geração de informações descritivas, como a contagem de

objetos ou a identificação de padrões específicos em uma imagem, que podem ser aplicados em sistemas de segurança e robótica, por exemplo (Barelli, 2018).

As técnicas de reconhecimento de padrões e aprendizado de máquina complementam a visão computacional, oferecendo abordagens que facilitam a classificação e identificação automática de elementos nas imagens. Em especial, as redes neurais têm se mostrado ferramentas eficazes para esse reconhecimento, pois podem "aprender" a identificar características complexas através de um grande volume de exemplos (Szeliski, 2022).

Assim, redes neurais profundas (deep learning), especificamente projetadas para a análise de imagens, constituem atualmente uma das principais tecnologias para avançar a visão computacional, permitindo a criação de sistemas que simulam, de forma limitada, a percepção humana em várias aplicações visuais.

2.6 REDES NEURAI

As redes neurais artificiais (RNAs) são estruturas computacionais inspiradas na arquitetura e funcionamento dos sistemas neurais biológicos, especialmente no cérebro humano. Essas redes são formadas por "neurônios" artificiais organizados em camadas e interconectados para processar e transmitir informações de maneira hierárquica e paralela. Cada neurônio em uma rede neural recebe dados como entrada, realiza uma operação matemática específica sobre esses dados e emite um sinal para os neurônios das camadas subsequentes (Haykin, 2001).

Os neurônios em uma rede neural artificial são unidades matemáticas inspiradas nos neurônios biológicos do cérebro humano. Eles são os blocos fundamentais que recebem informações, processam e transmitem sinais para outros neurônios na rede — tudo isso com o objetivo de aprender padrões e tomar decisões (Haykin, 2001).

Esse fluxo de informação possibilita que a rede ajuste seus parâmetros (ou "pesos") para aprender a mapear corretamente a entrada para a saída desejada, um processo central para aplicações que exigem reconhecimento de padrões, classificação de dados e tomadas de decisão baseadas em padrões complexos. O aprendizado se dá por meio de um processo iterativo, no qual a rede ajusta seus

pesos, geralmente minimizando uma função de erro para aperfeiçoar sua precisão (Haykin, 2001).

Dentro da ampla categoria das redes neurais artificiais, destacam-se alguns tipos com arquiteturas e finalidades distintas, sendo os principais: redes neurais convencionais (feedforward), redes neurais recorrentes (RNNs) e redes neurais convolucionais (CNNs). As redes neurais convencionais, ou de múltiplas camadas, têm uma estrutura de alimentação direta, na qual a informação se desloca em um único sentido, da camada de entrada para a camada de saída, sem ciclos ou recorrências. Essas redes são amplamente empregadas em problemas de classificação básica e regressão, porém apresentam limitações em problemas que requerem análise de dependências temporais ou relações sequenciais (Silva; Spatti; Flauzino, 2016).

As redes neurais recorrentes (RNNs), por outro lado, são caracterizadas por suas conexões cíclicas, que permitem a "memória" de estados anteriores ao longo do tempo. Essa característica possibilita que as RNNs mantenham informações temporais, sendo adequadas para tarefas que envolvem dados sequenciais, como o processamento de linguagem natural e séries temporais. Sua habilidade de processar sequências é essencial em problemas onde a ordem e a relação entre eventos passados são relevantes (Kovács, 2023).

Por fim, as redes neurais convolucionais (CNNs) foram desenvolvidas com foco específico na análise de dados espaciais, sobretudo em imagens. As CNNs são compostas de camadas convolucionais, que aplicam filtros sobre a entrada para identificar características locais, como bordas, texturas e formas, sem perder a informação de posicionamento espacial desses elementos na imagem (Furtado, 2019).

Essas características tornam as CNNs particularmente úteis para o reconhecimento e classificação de imagens, pois, ao aplicar múltiplos filtros, a rede consegue extrair padrões hierárquicos cada vez mais complexos, avançando de simples bordas para objetos e formas mais detalhadas em camadas mais profundas.

2.7 REDES NEURAIIS CONVOLUCIONAIS

As Redes Neurais Convolucionais (CNNs, do inglês Convolutional Neural Networks), uma forma especializada de Redes Neurais Artificiais (ANNs), têm sido

amplamente utilizadas para processar dados de imagem digital, devido à sua capacidade de aprender padrões hierárquicos complexos em dados visuais (Bezdan; Džakula, 2019).

Inspiradas pelo sistema visual humano, as CNNs empregam uma arquitetura onde cada neurônio é associado a um campo receptivo específico, permitindo que essas redes processem dados de maneira incremental. Cada camada foca em pequenas seções da imagem e, gradativamente, forma uma interpretação coesa através das conexões em camadas subsequentes (Mayank, 2020). A técnica operacional central das CNNs é a convolução, uma operação matemática linear que substitui a multiplicação matricial tradicional, transformando as imagens através de uma sequência de camadas interconectadas.

2.7.1 Camadas das Redes Neurais Convolucionais

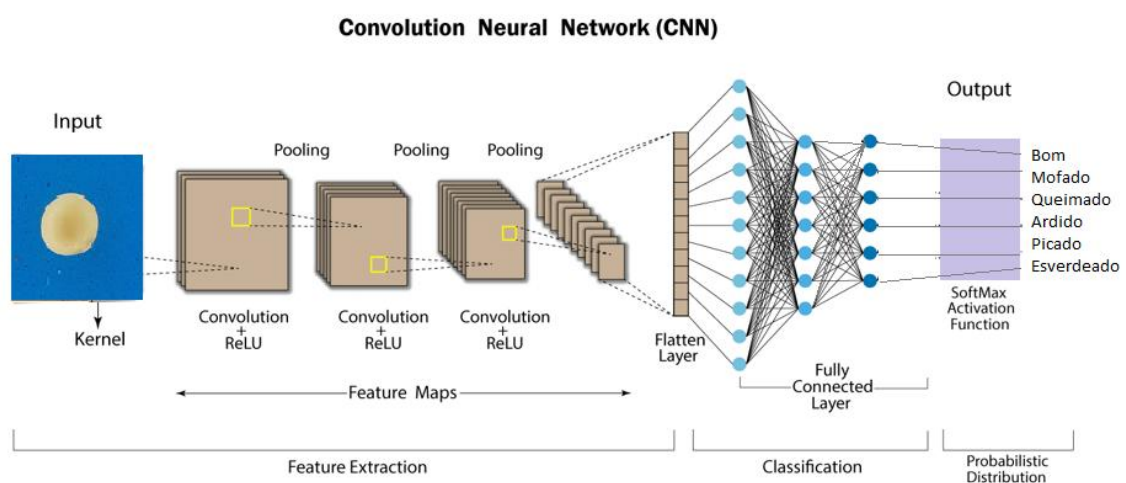
As CNNs são compostas por várias camadas-chave, cada uma responsável por uma função específica que contribui para o reconhecimento preciso de padrões e a extração de características a partir de dados visuais. A estrutura dessas redes neurais foi projetada para imitar o funcionamento do córtex visual humano, que processa informações visuais de maneira hierárquica e progressiva. Assim, as CNNs operam em camadas, onde cada uma delas aplica transformações e refinamentos distintos aos dados de entrada, permitindo que a rede construa gradualmente uma representação mais complexa e abstrata da imagem original (Ayeni, 2022).

A arquitetura em camadas possibilita à rede detectar desde padrões simples, como bordas e texturas, nas camadas iniciais, até formas mais complexas e características de alto nível nas camadas posteriores. Esse processamento hierárquico é um dos aspectos que tornam as CNNs altamente eficazes em tarefas de reconhecimento de imagem e classificação, pois permite que as redes extraiam e compilem informações visuais de forma semelhante à percepção humana (Ayeni, 2022).

Para que esse processo ocorra, cada tipo de camada na CNN desempenha uma função especializada: as camadas convolucionais identificam padrões locais por meio de filtros, as camadas de ativação introduzem não-linearidade, as camadas de *pooling* reduzem a dimensionalidade e as camadas totalmente conectadas consolidam as características extraídas para a classificação final. Essas camadas

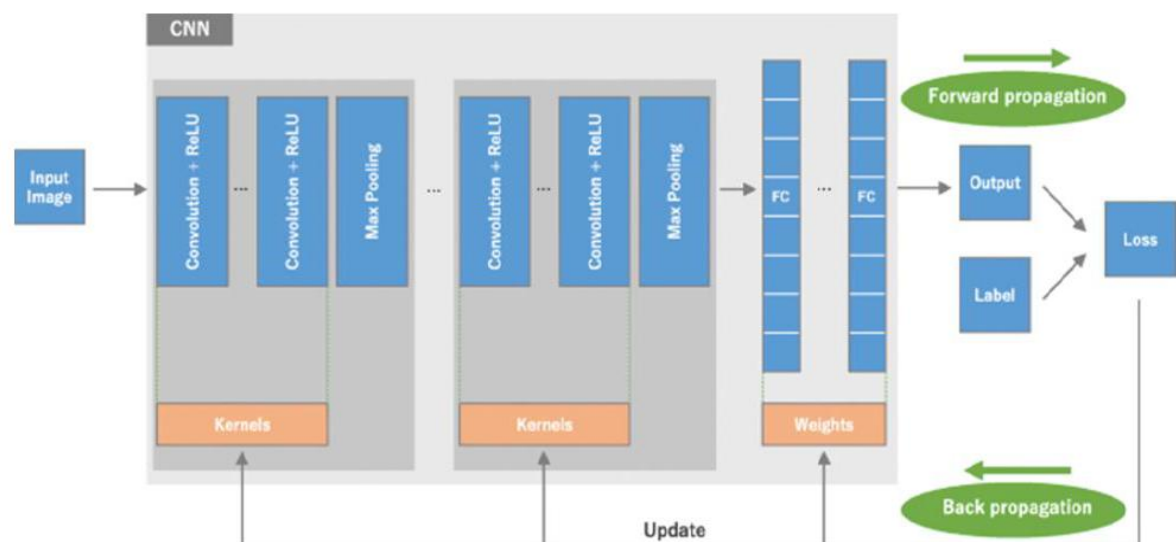
trabalham de forma integrada para aprimorar a capacidade da rede em aprender padrões visuais complexos e realizar previsões com alta precisão em uma variedade de tarefas, desde a classificação de imagens até a segmentação e detecção de objetos. A Figura 3 ilustra a arquitetura de uma CNN enquanto a Figura 4 esquematiza o processo de treinamento de uma CNN. As camadas que compõem uma rede neural deste tipo, suas funções, mecanismos operacionais e a forma como contribuem para o aprendizado e o reconhecimento de padrões visuais são apresentadas abaixo.

Figura 3 - Arquitetura de uma CNN



Fonte: Adaptado de Nazeer et al. (2024).

Figura 4 - Esquema de Aprendizado de uma CNN



Fonte: Yamashita et al. (2018).

2.7.1.1 Camada Convolucional

Esta camada constitui o núcleo das CNNs, utilizando um filtro ou kernel que percorre pequenas regiões localizadas da imagem, conhecidas como campos receptivos. Esse processo gera um mapa de características, onde os valores de saída representam padrões identificados nos dados de entrada. A convolução permite que as CNNs detectem estruturas específicas, como bordas ou texturas, nas camadas inferiores, avançando para padrões mais complexos nas camadas superiores. As camadas convolucionais são particularmente intensivas em termos de computação, pois realizam operações de produto escalar que capturam hierarquias espaciais essenciais para o processamento preciso de imagens (Mayank, 2020; Bezdan, Džakula, 2019).

2.7.1.2 Camada de Unidade Linear Retificada (ReLU)

Esta camada de ativação introduz não linearidade, aplicando uma função que retorna o valor de entrada diretamente se positivo, ou zero, caso contrário. A função ReLU é preferida em relação às funções tradicionais, como sigmoid ou tanh, que são suscetíveis à saturação de gradiente (Ohri, 2021).

"O gradiente é um vetor que indica a direção e intensidade da maior taxa de variação de uma função. Em redes neurais, essa função geralmente é a função de erro ou perda (como o erro quadrático médio ou a entropia cruzada), que mede o quanto longe a rede está da resposta correta" (Ohri, 2021).

Ao manter os gradientes acima de zero, a ReLU melhora a eficiência do aprendizado e permite que a rede adapte-se e atualize seus pesos durante o treinamento, abordando o problema do gradiente desaparecendo, comumente encontrado em redes profundas (Ohri, 2021).

2.7.1.3 Camada de Pooling

O *pooling*, frequentemente realizado através do pooling máximo, reduz a amostragem dos mapas de características ao resumir informações dentro de regiões localizadas, reduzindo efetivamente as dimensões espaciais. Esta camada não só

minimiza a carga computacional, mas também introduz invariância à translação, permitindo que a rede reconheça características independentemente de pequenas mudanças de posição dentro da imagem. As camadas de *pooling* ajudam a combater o sobreajuste ao consolidar informações essenciais, permitindo que as CNNs generalizem melhor em dados diversos (Pedamkar, 2021).

2.7.1.4 Camada Totalmente Conectada (FC)

Nas camadas totalmente conectadas, cada neurônio é conectado a todos os neurônios da camada anterior, consolidando informações dos mapas de características em um vetor unidimensional. Esta transformação facilita a classificação, interpretando características de alto nível em toda a imagem. A camada FC é responsável por gerar a saída final, representando tipicamente as probabilidades de cada classe, possibilitando aplicações como identificação de objetos, categorização de imagens e detecção de anomalias.

2.7.1.5 Camada de Normalização por Lote (Batch Normalization)

A camada de Batch Normalization é geralmente aplicada logo após as camadas convolucionais ou densas e antes da ativação. Sua função é normalizar as ativações dentro de um mini-batch, reduzindo a covariância interna das ativações - um problema conhecido com *internal covariate shift*. Essa normalização melhora a estabilidade do treinamento, acelera a convergência e possibilita o uso de taxas de aprendizado maiores, além de exercer um leve efeito regularizador (Ioffe; Szegedy, 2015).

2.7.1.6 Camada de Concatenação (Concatenate)

A camada de Concatenação é utilizada para combinar diferentes mapas de características produzidos por operações paralelas, sem a necessidade de somá-los. Essa técnica é essencial em redes como a InceptionV3, nas quais diferentes tamanhos de filtros (por exemplo, 1x1, 3x3, 5x5) são aplicados simultaneamente à mesma entrada. A concatenação dessas saídas permite que a rede analise a

informação em múltiplas escalas e profundidades, ampliando sua capacidade de extração de padrões visuais complexos (Szegedy et al., 2016).

2.7.1.7 Camada de Adição (Add Layer)

A camada de Adição é uma característica central em arquiteturas com conexões residuais como a ResNet50. Diferente da concatenação, ela realiza a soma elementar de duas ou mais entradas. Essa operação é usada para criar atalhos (skip connections) entre camadas distantes, facilitando o fluxo do gradiente durante o treinamento e reduzindo significativamente o risco de desaparecimento de gradientes em redes profundas. Esse mecanismo permite que a rede aprenda funções residuais mais facilmente, melhorando a performance sem aumentar a complexidade computacional (He et al., 2016).

Durante o treinamento de redes neurais, usamos o algoritmo de *backpropagation* para ajustar os pesos com base nos gradientes da função de erro. Em redes muito profundas, os gradientes podem se tornar muito pequenos à medida que são propagados para trás pelas camadas. Isso faz com que os pesos nas camadas iniciais mal sejam atualizados, o que prejudica o aprendizado, e a estratégia de *backpropagation* inibe este problema.

2.8 ARQUITETURAS DE REDES PRÉ-TREINADAS: INCEPTIONV3 E RESNET50

O avanço das Redes Neurais Convolucionais possibilitou o desenvolvimento de arquiteturas mais profundas e sofisticadas, capazes de extrair padrões visuais complexos com alta acurácia. No contexto industrial, onde o desempenho computacional e a precisão são fatores críticos, redes pré-treinadas como a InceptionV3 e ResNet50 se destacam como soluções viáveis e robustas para tarefas de classificação e inspeção visual.

2.8.1 InceptionV3

A InceptionV3 é uma evolução das arquiteturas GoogLeNet, desenvolvidas pelo Google, e introduz o conceito de módulos Inception, que aplicam diferentes tamanhos de filtros (como 1x1, 3x3, 5x5) simultaneamente à mesma entrada, permitindo que a

rede capture características em múltiplas escalas. Esses blocos são organizados de forma eficiente, utilizando convoluções fatoradas e camadas de *pooling* para reduzir o custo computacional, mantendo uma alta capacidade de extração e padrões visuais (Szegedy et al., 2016).

Uma das principais vantagens da InceptionV3 é seu equilíbrio entre profundidade e eficiência, sendo altamente indicada para aplicações industriais em que há grande variação nas características visuais dos objetos, como no caso dos grãos de soja. Além disso, sua arquitetura incorpora camadas do tipo *Batch Normalization* e *Concatenate*, que contribuem para a estabilidade do treinamento e para a combinação de diferentes representações de uma mesma entrada.

2.8.2 ResNet50

A ResNet50 é uma rede profunda com 50 camadas convolucionais, baseada no conceito de conexões residuais. Sua principal inovação é a introdução de atalhos (skip connections), que permitem que a saída de uma camada seja somada à uma saída de uma camada posterior. Essa técnica, viabilizada pelas camadas Add, facilita o fluxo de gradientes durante o treinamento e permite que redes muito profundas sejam treinadas sem perda significativa de desempenho (He et al., 2016).

O uso de conexões residuais tornou a ResNet50 altamente estável e escalável, com excelente capacidade de generalização, mesmo em cenários com ruído ou variação nos dados de entrada. Isso a torna adequada para ambientes industriais que exigem precisão consistente, mesmo diante de mudanças nas condições de captura de imagem ou variações sutis nos objetos analisados. Apesar disso, o InceptionV3 é mais aplicável a esse trabalho do que a ResNet50.

2.9 DETECÇÃO DE OBJETOS COM YOLO (YOU ONLY LOOK ONCE)

Com o avanço das aplicações de visão computacional em ambientes industriais, técnicas de detecção de objetos tem se tornado cada vez mais relevantes para identificar, segmentar e localizar elementos de interesse em imagens. Dentre os

algoritmos mais eficazes e amplamente utilizados destaca-se o YOLO (You Only Look Once), um modelo de detecção de objetos em tempo real que possui elevada velocidade e precisão.

O YOLO é baseado em uma rede neural convolucional que realiza, em uma única etapa, tanto a localização dos objetos na imagem quanto a classificação dos mesmos. Ao dividir a imagem em uma grade e prever caixa delimitadoras (bounding boxes) e probabilidades de classe para cada região, o YOLO oferece um alto desempenho em aplicações que exigem resposta rápida e precisa, como inspeções visuais e automatizadas na indústria (Redmon et al., 2016).

Uma das principais vantagens do YOLO em relação a outros métodos de detecção, é sua abordagem unificada de detecção, o que reduz a complexidade do pipeline e permite a implementação em dispositivos com recursos computacionais limitados, como câmeras inteligentes e sistemas embarcados. Essa característica torna o YOLO uma alternativa promissora para pré-processamento de imagens em linhas de produção, especialmente em sistemas de triagem automatizada de grãos.

Na aplicação desenvolvida neste trabalho, o YOLO foi empregado para segmentar individualmente os grãos de soja a partir de imagens compostas, fornecendo imagens recortadas que foram posteriormente utilizadas na etapa de classificação por redes neurais convolucionais.

2.10 AVALIAÇÃO DE DESEMPENHO EM MODELOS DE CLASSIFICAÇÃO

Para a análise dos resultados encontrados no Desenvolvimento de classificação dos grãos de soja, torna-se importante compreender os métodos de avaliação de desempenho utilizados em modelos de redes neurais. A eficácia dos modelos é possível verificar por meio da interpretação dessas ferramentas. Assim, esses conceitos são apresentados a seguir de forma a fundamentar a interpretação dos resultados.

2.10.1 Validação Cruzada

A validação cruzada é uma técnica estatística essencial utilizada para estimar a capacidade de generalização dos modelos de aprendizado de máquina, sobretudo em cenários com conjuntos de dados limitados. Este método consiste em dividir o conjunto de dados em múltiplas partições (ou "folds") e, iterativamente, treinar o modelo em algumas dessas partições e validá-lo nas restantes. A forma mais comum, denominada validação cruzada k-fold, segmenta os dados em k subconjuntos do mesmo tamanho, assegurando que cada subconjunto seja utilizado como validação exatamente uma vez (Kohavi, 1995). Ao mitigar o risco de overfitting e proporcionar uma estimativa mais precisa do desempenho do modelo, a validação cruzada é especialmente útil para tarefas complexas como a classificação de defeitos em grãos de soja com redes neurais convolucionais.

2.10.2 Matriz de Confusão

A matriz de confusão é uma representação tabular que descreve o desempenho de um modelo de classificação, organizando as previsões em categorias de verdadeiros positivos (TP), falsos positivos (FP), verdadeiros negativos (TN) e falsos negativos (FN) (Park et al., 2018). Cada célula da matriz reflete a contagem das instâncias em que as previsões coincidem ou divergem das classes reais. Este instrumento é crucial para calcular métricas complementares, como precisão, sensibilidade e especificidade, proporcionando uma avaliação mais abrangente do desempenho dos classificadores, especialmente em aplicações agrícolas de inspeção de qualidade de sementes.

2.10.3 Métricas de Avaliação de Desempenho

Com os dados da matriz de confusão, diversas métricas podem ser derivadas para mensurar o desempenho de modelos de classificação. Entre as principais métricas destacam-se:

- Acurácia: representa a proporção de previsões corretas em relação ao total de amostras avaliadas (Sokolova; Lapalme, 2009);

- Precisão (Precision): indica a proporção de predições positivas que foram corretas (Powers, 2011);
- Sensibilidade (Recall): mede a capacidade do modelo de identificar corretamente as instâncias positivas (Powers, 2011);
- Especificidade (Specificity): avalia a capacidade do modelo de identificar corretamente as instâncias negativas (Sokolova; Lapalme, 2009);
- F1-Score: calcula a média harmônica entre precisão e sensibilidade, sendo especialmente relevante em bases de dados desbalanceadas.

Além dessas métricas, outro parâmetro utilizado para análises é a Função de perdas (Loss), que serve para avaliar o erro da previsão em relação ao valor real, em que erros maiores indicam pior desempenho (Goodfellow; Bengio; Courville, 2016). A Tabela 3 apresenta um resumo das principais métricas utilizadas, suas fórmulas.

Tabela 3: Principais métricas de avaliação de redes neurais e suas fórmulas

Métrica	Fórmula
Acurácia	$Acurácia = (VP + VN) / (VP + VN + FP + FN)$
Precisão	$Precisão = VP / (VP + FP)$
Sensibilidade (Recall)	$Sensibilidade = VP / (VP + FN)$
Especificidade	$Especificidade = VN / (VN + FP)$
F1-Score	$F1 = 2 \times (Precisão \times Sensibilidade) / (Precisão + Sensibilidade)$
Função de Perda (Loss)	$L(\hat{y}_i, y_i) = \sum_{i=1}^k y_i \log(\hat{y}_i)$
<i>Legenda: VP = Verdadeiro Positivo; VN = Verdadeiro Negativo; FP = Falso Positivo; FN = Falso Negativo; $\hat{y}_i$ = probabilidade predita para a classe i; y_i = valor real (1 se a classe correta for i, 0 caso contrário); k = número de classes</i>	

Fonte: Adaptado de (Goodfellow; Bengio; Courville, 2016) e (Powers, 2011). *Deep Learning*. MIT Press (2016); POWERS, David M. W. *Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation*. Journal of Machine Learning Technologies, 2011.

2.11 APLICAÇÕES DAS REDES NEURAIIS CONVOLUCIONAIS

As Redes Neurais Convolucionais (CNNs) têm desempenhado um papel fundamental na resolução de problemas complexos em diversas áreas, destacando-se especialmente em aplicações que exigem alto desempenho em aprendizado profundo. Estas redes têm capacidade robusta para identificar padrões, interpretar características visuais complexas e processar dados em grande escala, o que as torna adequadas para uma variedade de domínios. Dentre as aplicações mais significativas das CNNs, estão a análise de imagens médicas, reconhecimento facial, descoberta de medicamentos e processamento de linguagem natural (PLN).

A aplicação das CNNs na análise de imagens médicas é uma das mais promissoras, devido à capacidade das redes em extrair automaticamente representações de dados complexos com alta precisão. CNNs são amplamente utilizadas para identificar tumores, fraturas e outras anomalias em imagens de raios-X e ressonância magnética (MRI). Por exemplo, imagens de órgãos como pulmões podem ser analisadas para detectar tumores, enquanto radiografias ósseas são examinadas para identificar fraturas.

As CNNs compararam imagens de pacientes com bancos de dados que contêm características similares, possibilitando diagnósticos mais precisos, conforme afirmado por Gandharv (2022). Além disso, ao serem treinadas em conjuntos de dados amplos, as CNNs são capazes de distinguir sinais de malignidade em imagens de tomografia (CT) e mamografia, auxiliando no diagnóstico precoce de câncer e aumentando a taxa de sucesso de tratamentos preventivos (Sharma, 2021).

O uso de CNNs na análise de imagens médicas também se expande para a ciência de dados em saúde, onde a análise preditiva é aplicada para avaliar o risco de doenças, como câncer de mama, COVID-19 e outros. Pesquisadores, como Li et al. (2014), Yadav e Jadhav (2019) e Afshar et al. (2019), investigaram o potencial das CNNs na classificação e predição de gravidade de doenças, oferecendo insights críticos para a intervenção médica.

As CNNs são amplamente aplicadas em sistemas de reconhecimento facial, uma área de particular importância no contexto da identificação e segurança. Esses sistemas lidam com imagens complexas, nas quais é necessário capturar e identificar características faciais únicas. Plataformas de mídia social, sistemas de vigilância e

tecnologias de identificação digital utilizam CNNs para autenticação biométrica. O reconhecimento facial por CNNs também pode ser estendido para a classificação de outras entidades visuais, como animais e objetos específicos, oferecendo robustez em cenários de vigilância e proteção. Segundo Sharma (2021), CNNs são particularmente eficazes para tais aplicações devido à sua capacidade de processar e diferenciar padrões sutis em imagens de alta complexidade.

A descoberta de medicamentos é um processo longo e dispendioso que pode ser significativamente otimizado com a aplicação de CNNs. A complexidade da pipeline de desenvolvimento de medicamentos e o vasto conjunto de dados envolvidos tornam as CNNs uma ferramenta essencial para análise preditiva e extração de padrões em bioinformática (Vamathevan et al., 2019).

CNNs auxiliam na análise de dados genômicos, permitindo a identificação de interações entre fármacos e genes, uma tarefa fundamental na farmacogenômica. De acordo com Vaz e Balaji (2021), a aplicação das CNNs no desenvolvimento de medicamentos ajuda a reduzir o tempo necessário para identificar e otimizar novos compostos químicos, promovendo avanços no tratamento de doenças emergentes e na resposta a pandemias.

Ao utilizar CNNs, as empresas de biotecnologia podem simplificar as fases críticas do desenvolvimento de fármacos, levando a uma produção mais rápida e menos onerosa. Essas redes otimizam a análise de bancos de dados complexos, facilitando a descoberta de substâncias químicas promissoras para novos tratamentos.

No processamento de linguagem natural, CNNs têm sido aplicadas com sucesso em tarefas como análise de sentimentos, categorização de tópicos e detecção de spam. A capacidade das CNNs de identificar padrões em dados não estruturados, como texto, torna-as adequadas para resolver problemas complexos no PLN, mesmo que esse campo seja mais naturalmente associado a redes neurais recorrentes (RNNs). Wang e Gang (2018) argumentam que a CNN tem se destacado devido ao seu design rápido e eficiente para análise de linguagem, que permite a compreensão de contextos e significados em grandes volumes de texto.

As CNNs também têm sido usadas no reconhecimento de fala, transformando áudio em imagens de espectrogramas e analisando esses dados para identificar palavras ou padrões fonéticos. Essa aplicação é especialmente útil em contextos de

tradução automática e assistentes de voz, onde a análise rápida e precisa de padrões acústicos é fundamental (Shamsaldin et al., 2019).

Diante do exposto, pode-se afirmar que as Redes Neurais Convolucionais (CNNs) representam um avanço significativo e inovador na análise de dados visuais e não visuais. A estrutura em camadas dessas redes permite um processamento e refinamento de informações com um nível de precisão incomparável, transformando setores tão diversos quanto medicina, segurança, descoberta de fármacos e processamento de linguagem natural (PLN). Esses avanços são fundamentais para melhorar a capacidade de análise e a eficácia na tomada de decisão, demonstrando o potencial contínuo das CNNs para atender e superar os desafios tecnológicos atuais.

Entre as aplicações mais recentes e promissoras, destaca-se o uso das CNNs na agroindústria, em particular para a identificação de defeitos e avarias em grãos de cereais. Combinando algoritmos de CNN e processamento de imagem, é possível detectar imperfeições nos grãos, como fungos, deformidades e danos estruturais, promovendo uma classificação mais rápida e precisa. Essa aplicação não apenas otimiza a qualidade do produto final, mas também auxilia na redução de perdas, contribuindo para uma cadeia produtiva mais eficiente e sustentável na agricultura moderna.

3 METODOLOGIA

Nesse estudo foram utilizados os frameworks TensorFlow e Keras para a construção e teste de diferentes arquiteturas CNN. A metodologia adotada segue uma abordagem sistemática dividida em etapas que incluem coleta de dados, pré-processamento, processamento, desenvolvimento das arquiteturas, treinamento e testes e implementação de um aplicativo para a classificação. Além disso, técnicas de aumento de dados (Data augmentation) e validação cruzada foram utilizadas para garantir a escalabilidade e robustez dos modelos, bem como para melhorar a generalização dos resultados.

3.1 COLETA DE DADOS

3.1.1 Obtenção das Amostras

Foram utilizados grãos de soja disponibilizados e coletados na unidade de recebimento da indústria Dureino localizada na cidade de Teresina-Piauí.

Para aquisição de imagens dos grãos, foi desenvolvida uma câmara fechada com iluminação artificial, com o objetivo de uniformizar a captura das imagens, eliminando sombras e variações externas de iluminação. Tal controle é fundamental para garantir a padronização das condições de aquisição, promovendo maior confiabilidade e acurácia nos processos de segmentação e extração de características (Gonzalez e Woods, 2010).

A câmara, ilustrada na Figura 5, foi construída em madeira, apresentando dimensões de 400 mm de comprimento, 300 mm de largura e 280 mm de profundidade. Na parte superior da estrutura, foi inserida uma abertura central de 100 mm, destinada ao posicionamento da câmera, possibilitando a obtenção de imagens perpendiculares ao plano onde os grãos são posicionados. A estrutura interna conta com iluminação artificial difusa, disposta de forma a minimizar a ocorrência de sombras ou excesso de luz, condições que podem comprometer a qualidade da imagem e, por consequência, os resultados da análise computacional (Santos et al., 2021).

As imagens foram capturadas com o auxílio de um aparelho celular, equipado com câmera de 48 megapixels e resolução de 8000 x 6000 pixels. Essa configuração

permitiu a obtenção de imagens com elevado nível de detalhe, condição essencial para a diferenciação visual entre grãos íntegros e grãos com defeitos, com base em suas características morfológicas e cromáticas.

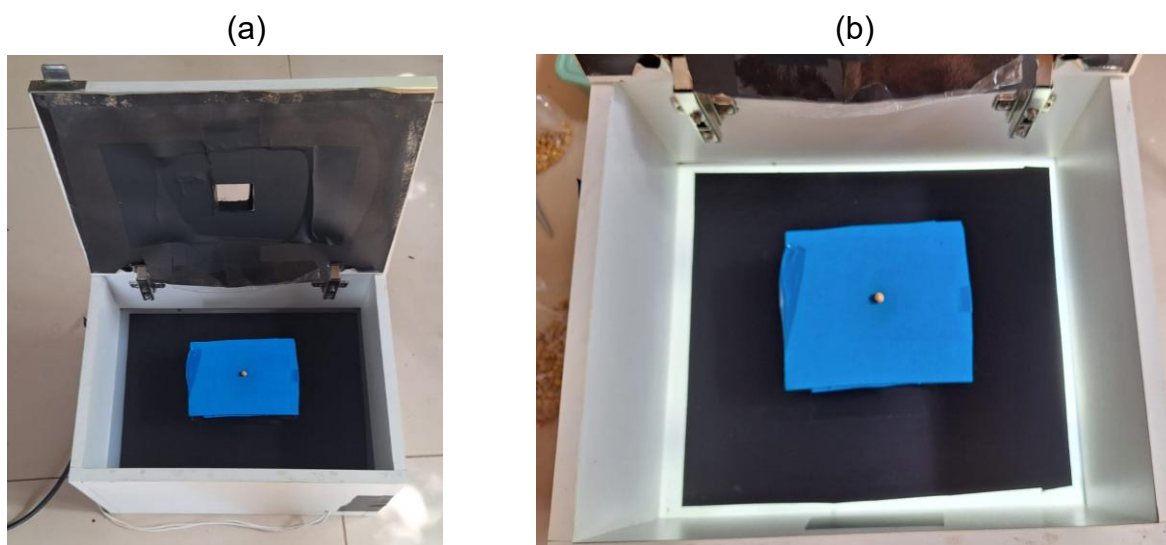
Figura 5 – Câmara para captação de imagens, fechada.



Fonte: Elaborado pelo autor (2025).

Durante a aquisição das imagens, os grãos foram dispostos sobre um fundo azul, conforme ilustra a Figura 6. A escolha dessa cor deve-se à sua ampla utilização em classificações visuais manuais e, sobretudo, ao fato de não apresentar semelhança tonal com os grãos avaliados, o que facilita o processo de segmentação e identificação de anomalias visuais (Silva et al., 2018). Tal escolha visa realçar os contornos e facilitar a extração de características relevantes para a classificação automática.

Figura 6 - Câmara para captação de imagens, fechada: (a) Iluminação desligada, (b) Iluminação ligada



Fonte: Elaborado pelo autor (2025).

A construção da câmara com controle de iluminação permitiu evitar perdas de detalhes visuais e inconsistências entre diferentes imagens, assegurando condições uniformes de captação, aspecto crítico para o treinamento e validação de modelos de classificação baseados em algoritmos de aprendizado de máquina (Goodfellow; Bengio; Courville, 2016). O *dataset* contém um total de 7328 imagens de grãos de soja, classificadas em seis categorias distintas, conforme apresentado na Tabela 4.

Tabela 4 - Número de imagens coletadas por categoria de grãos

Classe	Número de Imagens
Ardidos	1356
Bons	1385
Esverdeados	1021
Mofados	1186
Picados	1203
Queimados	1177

Fonte: Elaborado pelo autor (2025).

A Figura 7 apresenta os exemplos de imagem utilizados na classificação dos grãos de soja nas 6 classes consideradas: ardidos; bons; esverdeados; mofados; picados; queimados

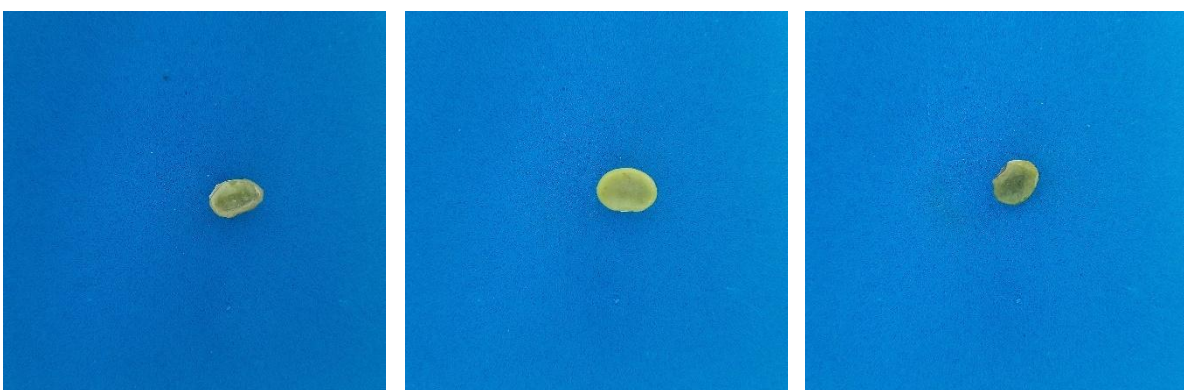
Figura 7 - Grãos de soja por categorias: (a) ardidos, (b) bons, (c) esverdeados, (d) mofados, (e) picados, e (f) queimados.



(a)



(b)



(c)



(d)



(e)



(f)

Fonte: Elaborado pelo autor (2025).

As imagens foram capturadas em condições controladas de iluminação e ângulos uniformes, cada imagem possui resolução de 2992 X 2992 pixels. Todas as imagens foram rotuladas manualmente. O *dataset* foi dividido em 80% para treinamento e 20% para testes.

Ainda, utilizando o YOLOv8n.pt, segmentou-se as imagens coletadas por meio de Augmentation, utilizando técnicas de auto-orientação, resize para

600X600 pixels e aumento de dados com flipagem horizontal e vertical com rotação entre -15° e 15° e ajuste de brilho entre -15% e 15% , que gerou 17.957 exemplos.

Tabela 5 - Distribuição do Dataset nas Etapas de Treinamento, Validação e Teste

Fase	Treinamento(%)	Validação (%)	Teste (%)
Detecção	17957 (88)	1723 (8)	859 (4)
Classificação	5.863 (80)	—	1465 (20)

Fonte: Elaborado pelo autor (2025).

3.2 TRATAMENTO DOS DADOS

3.2.1 Segmentação com YOLO

Para garantir que apenas as regiões relevantes fossem analisadas, foi empregada uma etapa de segmentação baseada no modelo YOLOv8n.pt, uma versão leve e eficiente do algoritmo You Only Look Once (YOLO). Essa técnica foi aplicada para detectar os grãos de soja nas imagens, gerar caixas delimitadoras (*bounding boxes*) ao redor de cada grão e extrair automaticamente as regiões de interesse.

A segmentação permitiu reduzir a interferência do fundo — como sombras e variações de cor — e assegurou que apenas os grãos isolados fossem utilizados na etapa de classificação. Essa abordagem tornou o processo mais limpo, padronizado e adequado ao uso em ambientes industriais.

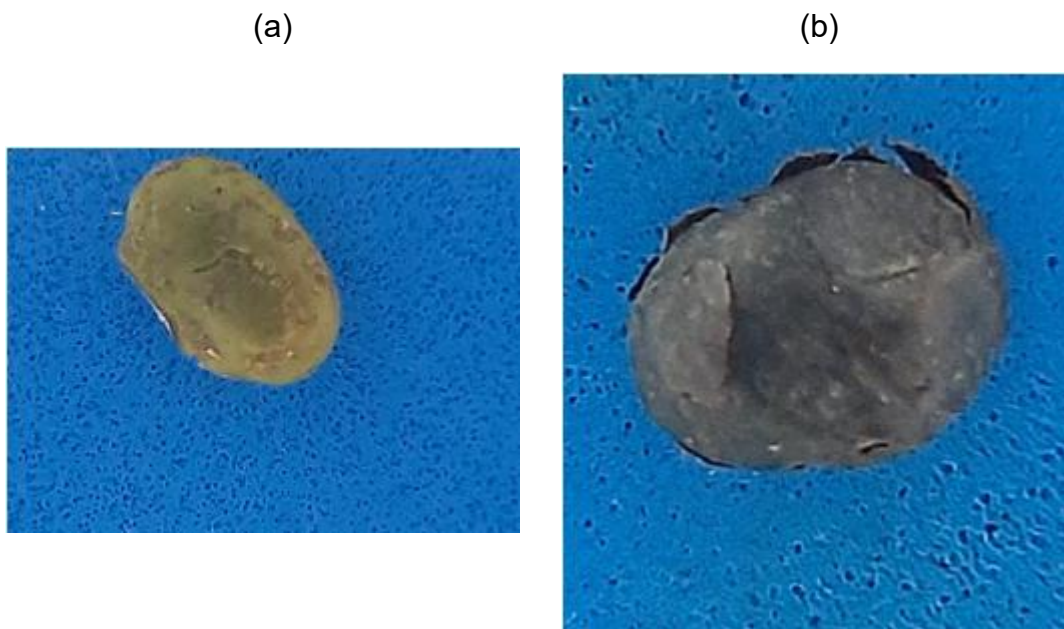
3.2.2 Pré-processamento

O pré-processamento das imagens foi composto pelas etapas: recorte; normalização; aumento de dados; redimensionamento e conversão para tons de cinza, conforme descritas abaixo:

(a) Recorte das Imagens: As imagens recebidas continham grãos que ocupavam menos de 4.1% da área total, portanto, foi necessário recortá-las para garantir que os grãos fossem mais proeminentes. O recorte variou entre dimensões de 500x550 a

900x850 pixels, para auxiliar o processo de data augmentation posterior. A Figura 8 apresenta exemplos de imagens de grãos de soja cortados.

Figura 8 - Exemplos de imagens de (a) esverdeado, (b) queimado



Fonte: de Mesquita Vieira, João Pedro; Ribeiro, Fernando Rodrigues; Paulo, Otilio (2025), "soybeans-not-noted-cut", Mendeley Data, V1, doi: 10.17632/xnx9bfh6b.1

(b) Normalização: Os valores de pixels das imagens foram normalizados para o intervalo $[0,1]$ dividindo-se cada valor por 255. Isso foi feito para reduzir a variância nas entradas, facilitando o processo de aprendizado do modelo.

(c) Aumento de dados (Data Augmentation): Para melhorar a robustez do modelo, aplicamos técnicas de aumento de dados, como rotações de até 90° , inversões horizontais, deslocamentos verticais e horizontais de até 10% e zoom máximo entre 5% e 20%.

(d) Redimensionamento das Imagens: Para garantir uniformidade, todas as imagens foram redimensionadas para 600x600 pixels utilizando a função target size do Keras.

(e) Conversão para tons de Cinza: As imagens foram convertidas para tons de cinza, diminuindo a complexidade computacional ao reduzir o número de canais de cor de 3 (RGB) para 1. Isso também ajudou na segmentação dos grãos.

3.2.3 Processamento

As imagens pré-processadas foram tratadas por camadas convolucionais para extrair padrões visuais relevantes como bordas, texturas e formas na sequência:

- i. **Convolução:** As camadas convolucionais aplicam filtros às imagens, gerando mapas de ativação que destacam características importantes, como bordas e contornos dos grãos.
- ii. **Pooling:** O MaxPooling foi utilizado para reduzir a dimensionalidade dos mapas de ativação, mantendo as informações mais relevantes. Isso reduz a carga computacional, permitindo uma análise mais eficiente dos dados.
- iii. **Função de ativação:** Foi aplicada a função de ativação ReLU (Rectified Linear Unit), que introduz não-linearidades ao modelo, possibilitando que a rede aprenda relações complexas entre as características visuais extraídas.

3.3 ARQUITETURAS TESTADAS

Foram testadas cinco arquiteturas principais de redes neurais convolucionais (CNNs). Dentre elas, três foram desenvolvidas do zero, permitindo a variação de parâmetros como zoom aplicável, tamanho dos kernels, pool size, número de camadas convolucionais e de *pooling*, além do número de neurônios nas camadas densas ocultas. Essas personalizações foram projetadas para avaliar como as alterações em cada um desses parâmetros impactam o desempenho geral na classificação dos grãos. Além das arquiteturas personalizadas, também foram incluídos dois modelos pré-treinados: ResNet e Inception, que serviram como referências para comparar a eficácia das abordagens customizadas.

3.4 TREINAMENTO DOS MODELOS

O treinamento das CNNs criadas do zero, foram realizados utilizando o otimizador Adam com uma taxa de aprendizado de 0,0001 para garantir um ajuste adequado do modelo sem *overfitting*, procedimento que acontece quando um modelo aprende

demaís os detalhes e ruídos do conjunto de dados de treinamento, ao ponto de memorizar os exemplos em vez de aprender padrões gerais. Isso faz com que ele tenha um desempenho excelente nos dados de treino, mas fracasse ao lidar com dados novos.

Cada arquitetura foi treinada por uma média de 80 épocas, com um batch size de 16 imagens por iteração. A função de perda selecionada foi a Categorical Crossentropy, apropriada para problemas de classificação multiclasse. Durante o treinamento, foram monitoradas métricas como acurácia e perda, tanto no conjunto de treinamento quanto no de validação, para avaliar o desempenho e ajustar os hiperparâmetros.

Já para os modelos pré-treinados, foram utilizadas entre 10 e 20 épocas para cada, com batch size de 32 imagens por iteração, com a mesma função de perda e métricas observadas nos modelos de CNNs personalizadas.

Métricas de precisão, acurácia, recall e F1-score foram calculadas apenas na fase de classificação dos modelos, a fim de avaliar o desempenho final da rede. A acurácia foi utilizada para medir o percentual de classificações corretas, enquanto precisão e recall avaliaram o desempenho em cada uma das classes individuais. O F1-score, uma média ponderada entre precisão e recall, foi aplicado para observar a eficiência do modelo nas classes com menor representatividade.

3.5 FERRAMENTAS UTILIZADAS

As ferramentas e bibliotecas empregadas nesse estudo foram escolhidas para facilitar o desenvolvimento, treinamento e avaliação das arquiteturas de CNNs, além de fornecer visualizações e manipulações eficientes dos dados. As principais ferramentas utilizadas foram:

- i. **Tensorflow e Keras** tanto para construção e treinamento e ajuste fino das redes neurais quanto para facilitar o processamento eficiente de imagens dentro do pipeline de treinamento.
- ii. **OpenCV** para operações de pré-processamento especialmente recortes.
- iii. **Matplotlib e Seaborn** para geração de gráficos de curvas de aprendizado e matrizes de confusão, auxiliando na análise dos resultados e comparações das diferentes arquiteturas.

- iv. **Google Colab**, utilizando o ambiente **Python 3** e **T4 GPU** como acelerador de hardware, permitindo o treinamento mais rápido e eficiente, otimizando tempo de processamento das operações convolucionais.

3.6 COMPARAÇÃO DE ARQUITETURAS

Foram desenvolvidas três arquiteturas personalizadas de CNNs que diferem na profundidade das camadas convolucionais, no número de neurônios nas camadas densas e no uso de *dropout* para regularização. O Dropout é uma técnica de regularização usada em redes neurais para evitar o famoso problema de overfitting — quando a rede aprende tão bem os dados de treinamento que perde a capacidade de generalizar para dados novos.

O Quadro 1 compila os parâmetros de cada uma das CNNs. Além das arquiteturas personalizadas, foram testados dois tipos populares de redes neurais convolucionais pré-treinadas conhecidas pela eficiência em classificação de imagens, estas estão agrupadas no Quadro 2.

Quadro 1: Resumos das CNNs personalizadas utilizadas

Parâmetros	CNNs		
	Modelo 1 - Inicial	Modelo 2	Modelo 3
Camadas Convolucionais	3 (32, 64 e 128 kernels)	3 (32, 64 e 128 kernels)	3 (32, 64 e 128 kernels)
Camadas de Pooling	3 (3 x 3)	3 (2 x 2)	3 (2 x 2)
Dropout	0	6 (20% cada uma)	6 (20% cada uma)
Camadas Densas	4 (64, 32, 32, 16 kernels)	4 (128, 64, 32 e 32 neurônios)	4 (128, 64, 32 e 32 neurônios)

Fonte: Elaborado pelo autor (2025).

Quadro 2: Resumos das CNNs pré-treinadas utilizadas

Parâmetros	CNNs	
	ResNet50	InceptionV3
Camadas Convolucionais	53 (variando de 64 à 2048 kernels)	94 (variando de de 32 à 448 kernels)
Camadas de Batch Normalization	53	94
Camadas de Pooling	2 (MaxPooling e AveragePooling2D)	4 (MaxPooling, variando de 64 a 768 feature maps)
Camadas Densas	2 (1024 e 6 neurônios, sendo a última como a camada final para classificação)	2 (1024 e 6 neurônios, sendo a última como a camada final para classificação)
Camadas Concatenate	---	15 (camadas que combinam diferentes tensores para produzir de 256 à 2048 feature maps)
Camadas Add Layers	16	---

Fonte: Elaborado pelo autor (2025).

3.7 APLICATIVO DE CLASSIFICAÇÃO

3.7.1 Estrutura e Desenvolvimento

O aplicativo destinado à classificação de grãos de soja foi desenvolvido com base em uma arquitetura descentralizada, em que adotou-se uma arquitetura modular com divisão clara entre as camadas de frontend e backend, o que contribui significativamente para a criação de um sistema mais robusto, escalável e de fácil manutenção. Essa separação permite que a interface com o usuário e os processos de lógica e dados evoluam de forma independente, otimizando tanto a organização do código quanto a implementação de melhorias, integrações com outras tecnologias e eventuais expansões da aplicação.

No que diz respeito à interface frontend, optou-se pelo uso do framework Flutter, reconhecido por sua rapidez no desenvolvimento e por possibilitar a criação de interfaces responsivas e de alta performance para múltiplas plataformas a partir de uma única base de código. Essa escolha foi estratégica, considerando a necessidade

de um aplicativo leve e eficiente, capaz de operar em dispositivos com recursos limitados — como aqueles comumente utilizados em campo — além de facilitar futuras expansões para iOS, web e desktop.

A estrutura do aplicativo foi construída com base no padrão MaterialApp e faz uso da biblioteca FlutterModular, que facilita a organização do projeto por meio de injeção de dependências e modularização de rotas. Essa abordagem contribui para uma separação clara das responsabilidades dentro do sistema, permitindo que funcionalidades como captura de imagem, visualização de resultados e pré-visualização fiquem organizadas em módulos independentes. Essa modularidade favorece a manutenção do código e torna o sistema mais escalável, especialmente à medida que novas funcionalidades forem incorporadas.

No desenvolvimento com Flutter, foram empregadas bibliotecas específicas para atender às necessidades do projeto, entre elas:

- **image_picker**: responsável pela captura de imagens, tanto pela câmera quanto pela galeria;
- **fl_chart**: utilizada para a exibição de gráficos de maneira clara e interativa;
- **provider**: para o gerenciamento de estado de forma reativa e eficiente;
- **http**: encarregada da comunicação segura com a API.

A camada de backend foi construída utilizando o FastAPI, uma ferramenta moderna e de alta performance para desenvolvimento de APIs RESTful com Python. O uso do FastAPI permite roteamento eficiente, integração nativa com OpenAPI (Swagger) e facilidade na criação de endpoints assíncronos.

A comunicação com o frontend ocorre por meio de requisições HTTP seguras, utilizando o formato multipart/form-data para envio de imagens. Após o recebimento, um pipeline de visão computacional é executado no servidor: algoritmos de segmentação, como o Watershed, são aplicados para isolar individualmente cada grão de soja presente na imagem.

Cada grão segmentado passa por uma etapa de normalização e é então enviado a um modelo de aprendizado profundo previamente treinado, que realiza a classificação automatizada com base em critérios técnicos definidos. Esse modelo encapsulado na camada Model é responsável pela inferência, enquanto o Controller trata o pré-processamento e a formatação da resposta para o frontend. Essa abordagem segue a arquitetura MVC (Model-View-Controller) adaptada ao contexto

de serviços web, garantindo uma separação lógica entre as responsabilidades do sistema.

A resposta enviada ao aplicativo inclui o resultado da classificação de cada grão, permitindo exibição estruturada, agregações em dashboards e geração de relatórios. Esse fluxo automatizado assegura eficiência operacional, escalabilidade e alta confiabilidade, mesmo sob demandas elevadas

4 RESULTADOS E DISCUSSÃO

Os testes de validação de acurácia foram realizados de forma padronizada, garantindo consistência e comparabilidade nas avaliações dos modelos. Os resultados obtidos para avaliação de todas as arquiteturas testadas estão agrupados na Tabela 6.

Tabela 6 - Resultados de Acurácia Geral e Tempo de Treinamento

Modelo	Acurácia Geral (%)	Tempo de Treinamento (min/época)
Modelo 1	6.1%	3.3
Modelo 2	16.4%	3.6
Modelo 3	16.6%	3.9
ResNet50	63.1%	10.2
InceptionV3	95.9%	9.9

Fonte: Elaborado pelo autor (2025).

Os modelos de CNN personalizados e as redes pré-treinadas compartilham camadas comuns de convolução, *pooling* e funções de ativação. Cada uma delas é responsável por extrair características visuais (como bordas e texturas) dos grãos de soja. Nos modelos personalizados, as camadas de convolução são seguidas por camadas de pooling para reduzir a dimensionalidade dos dados. As funções de ativação ReLU (Unidade Linear Retificada) introduzem não-linearidade, o que permite que a rede aprenda padrões mais complexos.

A ResNet50 é amplamente utilizada devido à sua capacidade de preservar o gradiente durante o treinamento, graças às suas camadas residuais (*Add layers*), que evitam o problema de desaparecimento de gradiente (*gradient vanishing*). Isso permite que a rede seja mais profunda, com mais de 50 camadas, resultando em uma melhor extração de características e, portanto, maior precisão, especialmente em classes visuais complexas como as encontradas nos grãos de soja.

Com o desempenho de 63.1% de acurácia, mostrou-se eficiente e, apesar de ser menos precisa que a InceptionV3, mostrou capacidade de generalização e robustez em redes mais profundas tornando-a também uma boa escolha para implementações em casos com maior variabilidade de dados.

O modelo InceptionV3 se destaca pela sua abordagem de "blocos Inception" que permitem que a rede use múltiplos tamanhos de filtro convolucionais ao mesmo tempo, capturando características visuais em diferentes escalas. Isso o torna altamente eficiente na classificação de imagens com uma variedade de texturas e padrões.

Com desempenho de 95.9% de acurácia em poucas épocas treinadas, demonstra a sua capacidade de capturar variações sutis de diferentes categorias de grãos, tornando uma opção ideal para cenários que exigem alta precisão.

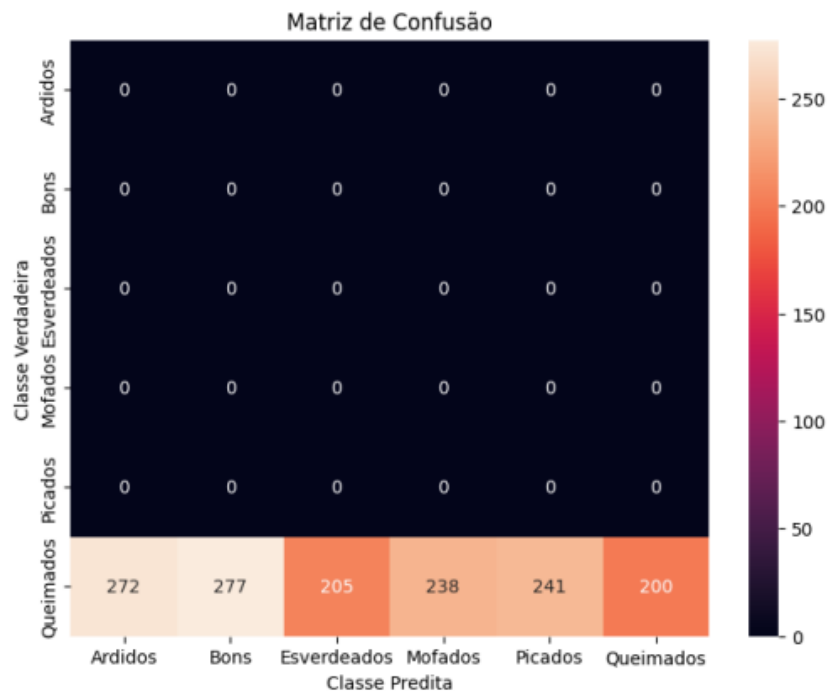
4.1 AVALIAÇÃO DO DESEMPENHO DOS MODELOS: IMPLICAÇÕES E UTILIZAÇÕES PRÁTICAS

A aplicação de redes neurais convolucionais para classificação de grãos de soja revelou-se promissora, especialmente no contexto industrial, onde há crescente demanda por agilidade, padronização e confiabilidade nos processos de controle de qualidade. A substituição da avaliação visual manual — sujeita a variações humanas — por um sistema automatizado baseado em visão computacional permite maior uniformidade nos resultados, além de significativa redução no tempo de análise.

As redes pré-treinadas se destacaram, apresentando desempenho consideravelmente superior em relação às redes customizadas. A escolha entre essas abordagens deve considerar fatores como precisão necessária, custo computacional e infraestrutura disponível.

As CNNs customizadas demonstraram limitações significativas quanto à acurácia, atingindo no máximo 16,6% (modelo 3). No entanto, devido à sua leveza e menor exigência computacional, ainda podem ser úteis em sistemas de pré-triagem ou prototipagem, onde o objetivo é apenas uma separação preliminar de grãos com defeitos mais evidentes. Também se mostram viáveis em dispositivos embarcados de baixo custo, onde há restrições quanto a memória e processamento. A Figura 9 apresenta a matriz de confusão do modelo 3.

Figura 9 - Matriz de confusão do modelo 3



Fonte: Elaborado pelo autor (2025).

Tabela 7 - Métricas de desempenho por classe do melhor modelo customizado

Classe	Precision	Recall	F1-Score	Support
Ardidos	0.00	0.00	0.00	272
Bons	0.00	0.00	0.00	277
Esverdeados	0.00	0.00	0.00	205
Mofados	0.00	0.00	0.00	238
Picados	0.00	0.00	0.00	241
Queimados	0.14	1.00	0.24	200
Accuracy			0.16	1433
Macro Avg	0.02	0.17	0.04	1433
Weighted Avg	0.02	0.14	0.03	1433

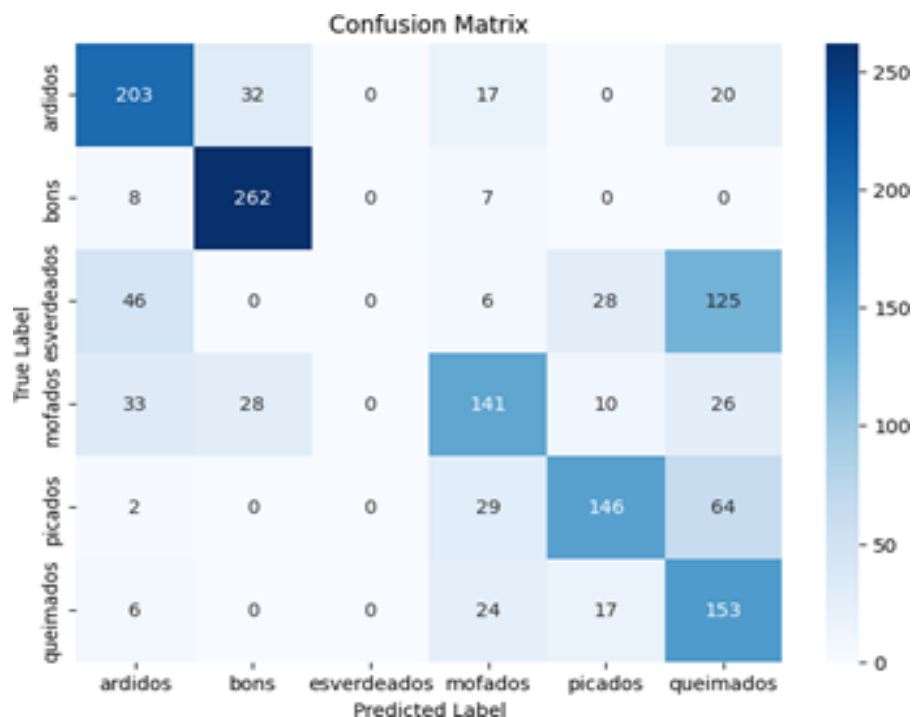
Fonte: Elaborado pelo autor (2025).

Pelos resultados observados, percebe-se que o modelo não foi capaz de classificar corretamente as classes "ardidos", "bons", "esverdeados", "mofados" e "picados", uma vez que nas métricas de precisão, recall e F1-Score, foram encontrados valores nulos. Para a classe "queimados", embora o recall tenha

alcançado 1.00, indicando que todos os exemplos reais foram identificados, a precisão foi bastante baixa (0,14), sugerindo que a maioria das previsões dessa classe foram incorretas, ou seja, falsos positivos. O F1-Score de 0,24 reflete esse desequilíbrio entre precisão e recall. De forma geral, o modelo apresentou uma acurácia de apenas 16%, o que revela um desempenho insatisfatório na tarefa de classificação.

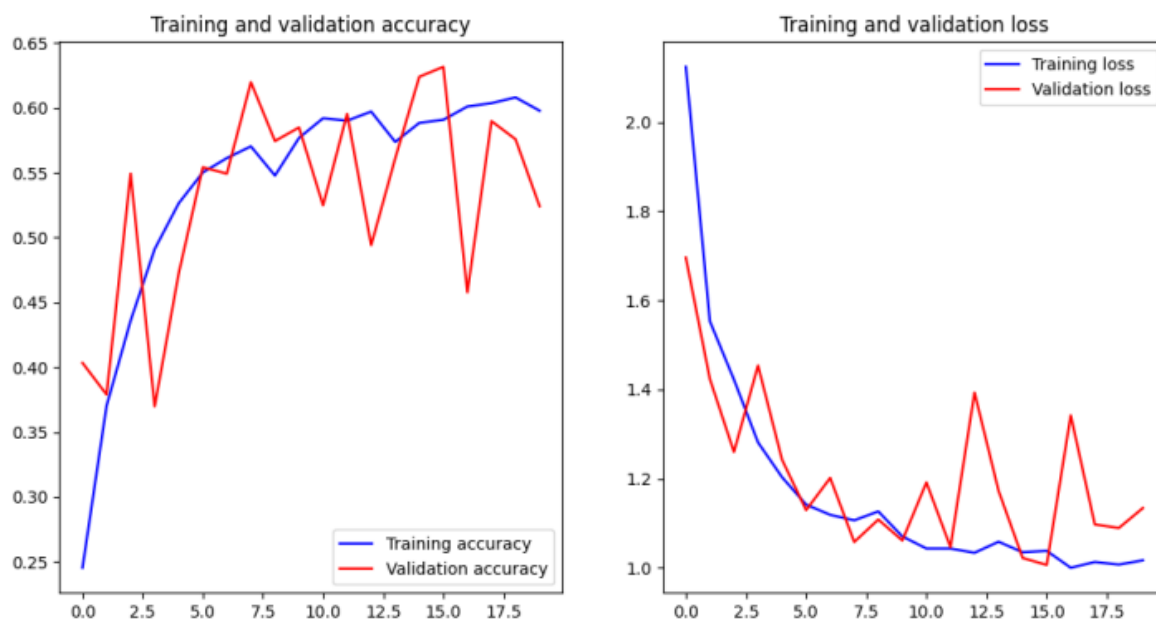
O modelo pré-treinado ResNet50 obteve uma acurácia de 63,1%, sendo eficaz na extração de características visuais mais complexas. Essa rede é recomendada para ambientes industriais com infraestrutura computacional robusta, como workstations com GPU ou servidores remotos, onde grandes volumes de grãos são processados continuamente. Sua estrutura com camadas residuais contribui para uma maior capacidade de generalização, sendo adequada para aplicações em que se busca equilíbrio entre desempenho e custo operacional. A Figura 10 apresenta a matrix de confusão modelo ResNet50.

Figura 10 - Matriz de confusão do modelo ResNet50



Fonte: Elaborado pelo autor (2025).

Figura 11 - Gráficos de acurácia e perda (loss) do modelo ResNet50 durante o treinamento e validação.



Fonte: Elaborado pelo autor (2025).

Tabela 8 - Métricas de desempenho por classe do modelo ResNet50

Classe	Precision	Recall	F1-Score	Support
Ardidos	0.68	0.75	0.71	272
Bons	0.81	0.95	0.87	277
Esverdeados	0.00	0.00	0.00	205
Mofados	0.63	0.59	0.61	238
Picados	0.73	0.61	0.66	241
Queimados	0.39	0.77	0.52	200
Accuracy			0.63	1433
Macro Avg	0.54	0.61	0.56	1433
Weighted Avg	0.57	0.63	0.59	1433

Fonte: Elaborado pelo autor (2025).

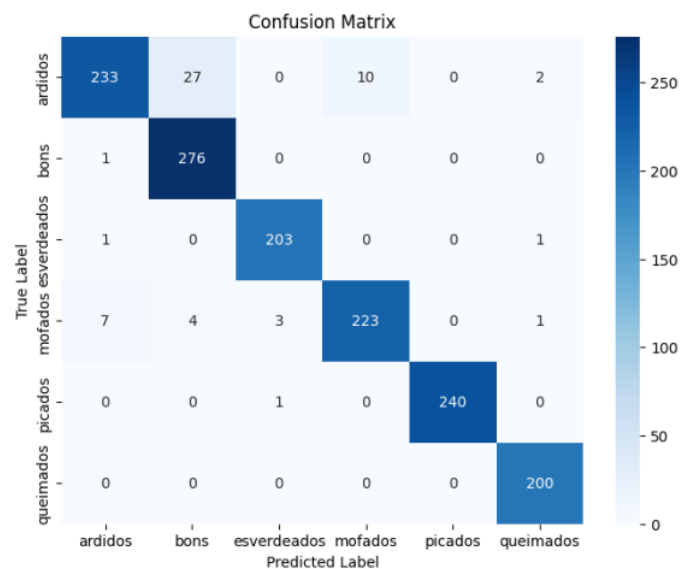
Com os valores apresentados, o modelo revelou um bom desempenho na maioria das classes, mas para a classe “esverdeados” não conseguiu classificar corretamente nenhum exemplo, a qual foi confundida com as classes "ardidos", "mofados", "picados" e "queimados". Isso sugere que o modelo ainda precisa de ajustes para melhorar a capacidade de diferenciar essas categorias. Os melhores resultados foram obtidos nas classes "bons" (F1-Score de 0,87) e "ardidos" (F1-Score

de 0,71). Já para a classe "queimados", o modelo teve um recall alto (0,77), ou seja, identificou a maioria dos exemplos reais, mas a precisão foi baixa (0,39), indicando muitos falsos positivos. A acurácia geral foi de 63%, com médias de F1-Score de 0,56 (macro) e 0,59 (ponderada), o que mostra que o desempenho geral foi prejudicado pelas dificuldades em algumas classes e que ainda há margem para melhorias.

A arquitetura InceptionV3 obteve o melhor desempenho, com acurácia de 95,9%, evidenciando seu alto potencial para aplicações de classificação de alta precisão em ambientes industriais. Sua capacidade de capturar variações visuais em múltiplas escalas permite uma análise mais sensível e confiável dos grãos, mesmo diante de variações sutis. Isso a torna especialmente indicada para linhas de produção automatizadas, como em fábricas de processamento de grãos, onde a padronização da classificação é essencial para garantir a qualidade do produto final.

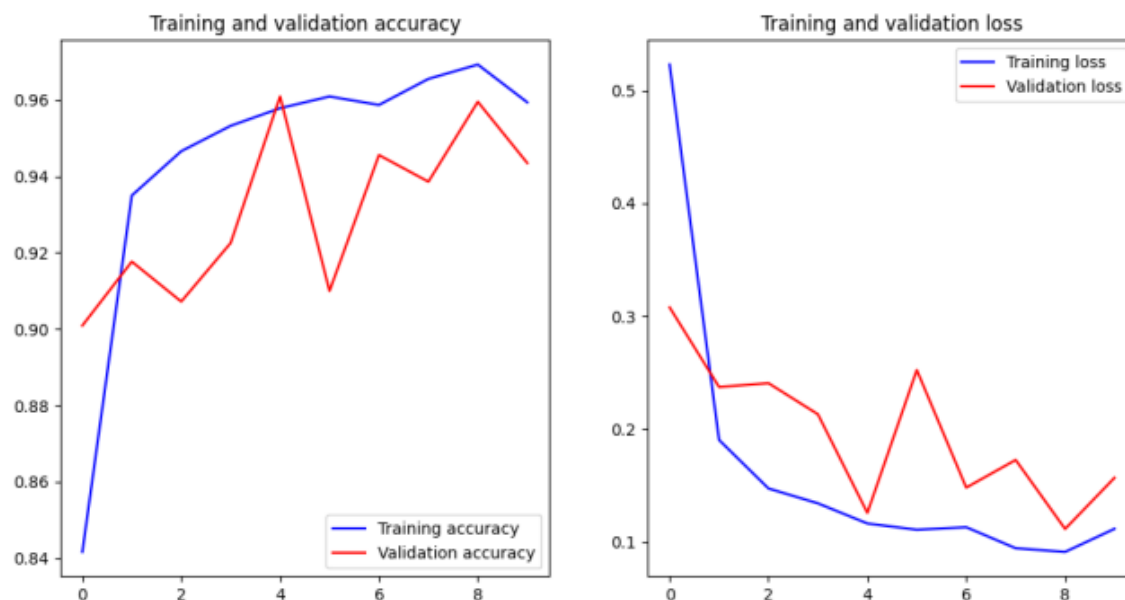
Além disso, sua eficiência na utilização de recursos computacionais permite que seja implementada tanto em ambientes em nuvem, quanto em dispositivos móveis mais avançados, como drones agrícolas ou sistemas autônomos de inspeção, que podem realizar a classificação em tempo real, diretamente no campo ou na linha de produção. A Figura 12 apresenta a matrix confusão do modelo InceptionV3.

Figura 12 - Matriz de confusão do modelo InceptionV3



Fonte: Elaborado pelo autor (2025).

Figura 13 - Gráficos de acurácia e perda (loss) do modelo InceptionV3 durante o treinamento e validação.



Fonte: Elaborado pelo autor (2025).

Tabela 9 - Métricas de desempenho por classe do modelo InceptionV3

Classe	Precision	Recall	F1-Score	Support
Ardidos	0.96	0.86	0.91	272
Bons	0.90	1.00	0.95	277
Esverdeados	0.98	0.99	0.99	205
Mofados	0.96	0.94	0.95	238
Picados	1.00	1.00	1.00	241
Queimados	0.98	1.00	0.99	200
Accuracy			0.96	1433
Macro Avg	0.96	0.96	0.96	1433
Weighted Avg	0.96	0.96	0.96	1433

Fonte: Elaborado pelo autor (2025).

Com base nos resultados obtidos do modelo, verifica-se um excelente desempenho na classificação das classes de defeitos em grãos de soja. Para as métricas de precisão, recall e F1-Score foram encontrados valores altos para todas as classes, destacando-se as classes "Picados", "Esverdeados" e "Queimados", que obtiveram F1-Score de 1.00 ou próximo disso. A acurácia geral do modelo foi de 0.96,

o que indica que ele classificou corretamente a grande maioria dos exemplos, demonstrando um desempenho equilibrado e confiável.

4.2 COMPARAÇÃO COM TRABALHOS RELACIONADOS

Apesar de diversos estudos abordarem a classificação de grãos em suas diferentes variedades para a obtenção de variadas características, ainda é possível observar uma limitação de pesquisas focadas na identificação de defeitos em grãos de soja utilizando técnicas de visão computacional e redes neurais.

Com o objetivo de avaliar o desempenho da abordagem proposta, realizou-se uma comparação com estudos previamente publicados que abordam a classificação de defeitos em grãos de soja por meio de redes neurais convolucionais.

A Tabela 10 apresenta um resumo comparativo das abordagens e resultados descritos em diversos estudos, destacando as principais características, arquiteturas, quantidades de imagens analisadas, acurácias, sensibilidade e F1-Score.

Tabela 10 - Estudos sobre Classificação de Defeitos em Grãos de Soja utilizando Redes Neurais e Visão Computacional

Autor(es) e Ano	Características Analisadas	Arquitetura / Método	Qtd. de Imagens de Grãos	Acurácia (%)	Sensibilidade	F1-Score
Machado (2021)	Defeitos em grãos de soja (9 classes)	ResNet34, MobileNet, VGG19	900 (21 por classe)	99,55%(com augmentation)	Não informado	Não informado
Zhao et al. (2021)	Defeitos em sementes com visão 360°(6 classes)	MobileNetV2 aprimorada	1.260 (12960 processadas)	97,84%	0,99	0,99
Huang et al. (2022)	Defeitos em sementes de soja (5 classes)	Mask R-CNN + SNet	336	96,2%	0,94	0,955
Lin et al. (2023)	Defeitos em sementes de soja (4 classes)	SoyNet	3100	95,63%	Acima de 0,95	Acima de 0,96
Saito et al. (2024)	Defeitos externos de sementes (4 classes)	CNN multi-entrada (cores e fluorescência)	800g	93,9%	Não informado	Não informado
Sable et al. (2024)	Defeitos em sementes de soja (8 classes)	SSDINet	1000	98,64%	0,9764	0,9866

Autor(es) e Ano	Características Analisadas	Arquitetura / Método	Qtd. de Imagens de Grãos	Acurácia (%)	Sensibilidade	F1-Score
<i>Este trabalho</i>	<i>Defeitos em grãos de soja (6 classes)</i>	<i>InceptionV3</i>	<i>7328</i>	<i>95,9%</i>	<i>0,96</i>	<i>0,96</i>

Fonte: Elaborado pelo autor (2025).

Machado (2021) propôs a utilização das arquiteturas ResNet34, MobileNet e VGG19 para classificar nove categorias distintas de defeitos em grãos, utilizando um conjunto reduzido de 900 imagens (100 imagens por classe), obtendo-se acurácia de 99,55% mediante o uso de técnicas de data augmentation. Embora o índice de acurácia seja elevado, o tamanho limitado da base compromete a generalização do modelo.

Estudos como os de Zhao et al. (2021) e Huang et al. (2022) também alcançaram ótimos desempenhos, com acurácia de 97,84% e 96,2%, respectivamente. No entanto, utilizaram conjuntos de dados consideravelmente menores que o deste trabalho — 1.260 imagens (Zhao) e apenas 336 imagens (Huang) — o que, embora relevantes, representa desafios distintos em termos de robustez e escalabilidade.

Já Lin et al. (2023) propôs a arquitetura SoyNet, com 3.100 imagens e acurácia de 95,63%, números mais próximos à realidade enfrentada neste estudo, no entanto utilizou apenas para 4 classes: soja normal, danificada, anormal e não classificável.

O trabalho de Saito et al. (2024) apresenta um enfoque diferenciado ao utilizar uma CNN multi-entrada, incorporando imagens de cor e fluorescência. Apesar de não especificarem o tamanho da base de dados, apenas o peso da amostra (800g), os autores relataram uma acurácia de 93,9%, embora os indicadores de sensibilidade e F1-score não tenham sido informados. destacando os benefícios da combinação de diferentes modalidades de imagem.

Sable et al. (2024) o estudo analisou características visuais de sementes de soja com defeitos bons e 7 tipos de defeitos, com 1.000 imagens . Utilizou a rede SSDINet, o modelo apresentou acurácia de 98,64%, com sensibilidade de 97,64% e F1-Score de 98,66.

Este trabalho se destaca por utilizar uma base substancialmente maior (7.328 imagens) com seis classes no total, cinco correspondentes a diferentes tipos de defeitos e uma representando grãos bons. As imagens foram capturadas em caixa de

iluminação controlada, o que assegura padronização nas condições de entrada e contribui para a consistência na classificação, especialmente em defeitos visualmente sutis.

O modelo treinado com a arquitetura InceptionV3 atingiu 95,9% de acurácia, com sensibilidade e F1-score de 0,96, demonstrando desempenho robusto mesmo em um cenário com maior volume de dados e diversidade de classes.

Em síntese, os resultados obtidos neste estudo são altamente competitivos em relação à literatura, conciliando volume de dados, rigor técnico na aquisição das imagens e desempenho equilibrado nos principais indicadores. O modelo proposto apresenta potencial de aplicação prática consistente para sistemas automatizados de classificação de grãos, contribuindo de forma significativa para a automação e confiabilidade na avaliação da qualidade da soja.

Ainda destaca-se como um dos grandes diferenciais deste estudo a criação e disponibilização pública de sua base de imagens, algo inédito entre os trabalhos comparados.

4.3 TESTE DO APLICATIVO

Com o desenvolvimento do aplicativo concluído, foram conduzidos testes para verificar sua funcionalidade, usabilidade e desempenho geral em condições reais de uso. A seguir, são descritas as principais etapas de interação do usuário com a aplicação, acompanhadas das observações obtidas durante os testes práticos.

A primeira etapa do processo ocorre na tela inicial do aplicativo, ilustrada na Figura 14. Esta é a interface principal (home page), na qual o usuário pode optar por capturar uma nova imagem utilizando a câmera do dispositivo ou selecionar uma imagem já existente a partir da galeria. Para isso, estão disponíveis dois botões: “Câmera”, que aciona a função de captura em tempo real, e “Galeria”, que permite navegar pelos arquivos armazenados no dispositivo.

Figura 14 - Tela inicial do aplicativo com opções para captura ou seleção de imagem para análise



Fonte: Elaborado pelo autor (2025).

Após a seleção da imagem, seja por captura com a câmera ou importação da galeria, o aplicativo exibe a visualização da foto escolhida na interface principal. A partir dessa etapa, o usuário pode acionar o botão "Classificar Imagem", que envia a imagem para o backend, onde é realizado o processamento e a classificação automática dos grãos com base no modelo treinado. Também estão disponíveis as opções para segmentação da imagem e acesso ao histórico de análises anteriores.

Figura 15 - Tela com imagem selecionada e opções de classificação e segmentação disponíveis



Fonte: Elaborado pelo autor (2025).

Ao clicar em "Segmentar imagem" aparecerá essas imagens que são das regiões que a segmentação encontrou e da máscara de segmentação que ele utilizou sobre a imagem original, conforme ilustrado na Figura 16.

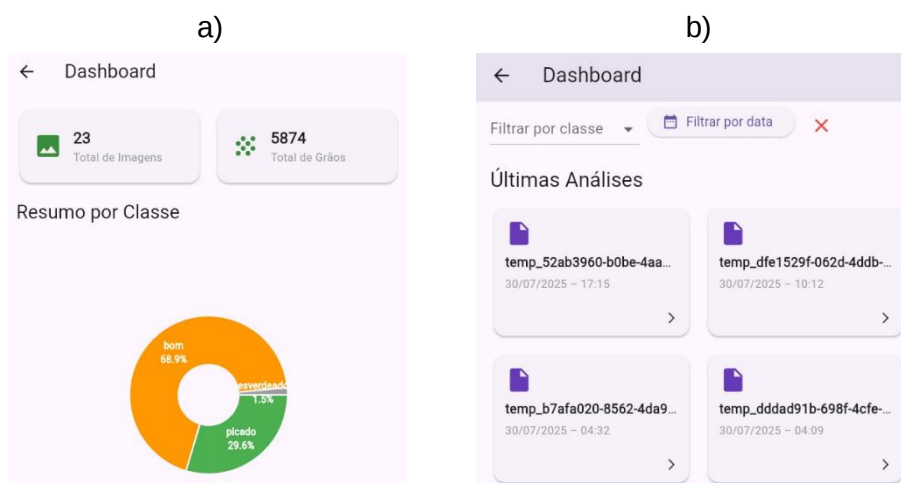
Figura 16 - Tela com imagem mostrando contorno e segmentação dos grãos



Fonte: Elaborado pelo autor (2025).

Ao clicar em histórico de análises aparecem a quantidade total de imagens e o número total de grãos analisados e a porcentagem de cada classe analisada e também abaixo mostra as últimas análises onde podemos clicar e verificar cada uma individualmente conforme a Figura 17.

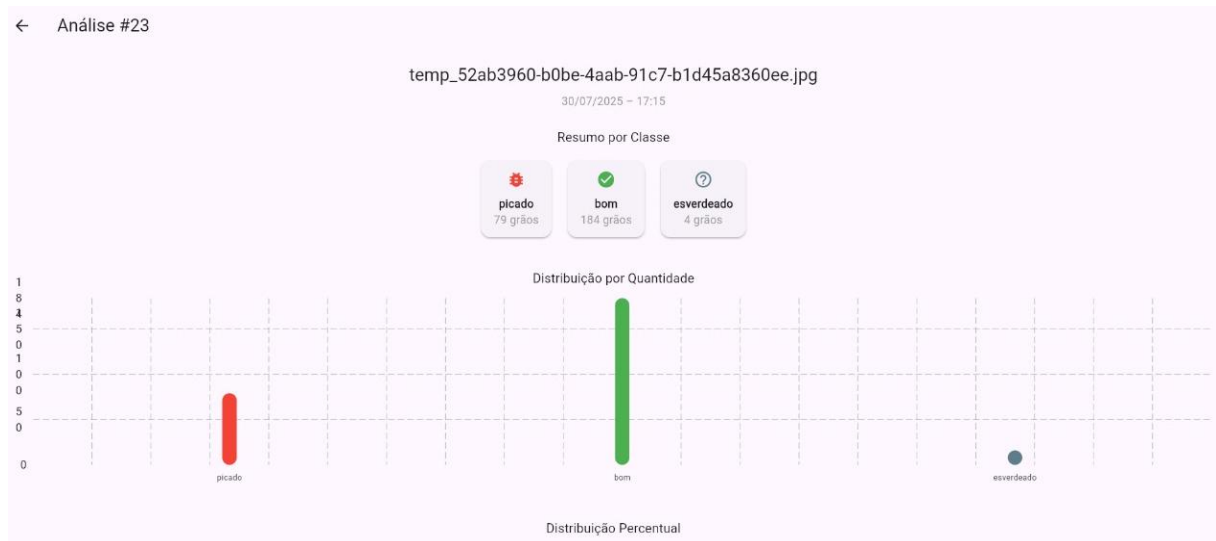
Figura 17 - a) histórico de análises e b) últimas análises



Fonte: Elaborado pelo autor (2025).

Ao clicar em algum desses pontos individuais da Fig 17 (b), surge a tela de análise individual de cada imagem, onde mostra o resumo por classe e um gráfico exemplificando, conforme imagem abaixo.

Figura 18 - Tela de Análise Individual: Resumo por Classe e Distribuição Quantitativa



Fonte: Elaborado pelo autor (2025).

4.4 LIMITAÇÕES

Apesar dos resultados encorajadores, algumas limitações foram observadas e devem ser consideradas para futuras implementações. As CNNs customizadas sofreram com baixa acurácia e indícios de overfitting, principalmente pela ausência de técnicas mais avançadas de regularização e camadas de normalização. Isso restringe sua aplicabilidade a tarefas de menor complexidade ou à operação em ambientes com severas restrições de hardware, onde o objetivo é detectar padrões muito evidentes.

O modelo ResNet50, embora eficiente, requer maior poder computacional e tempo de treinamento, o que pode limitar sua aplicação em sistemas embarcados ou em soluções que exijam processamento local em tempo real, como classificadores de linha independentes.

Já a InceptionV3, mesmo com alto desempenho, apresenta complexidade elevada. Seu uso em dispositivos de baixo consumo energético ou ambientes com recursos computacionais limitados pode não ser viável sem adaptações ou compressões no modelo. Além disso, o tempo de treinamento inicial e o custo

computacional ainda representam desafios, principalmente para empresas que desejam implementar soluções rápidas e de baixo custo.

5 CONSIDERAÇÕES FINAIS

Os resultados obtidos até o momento reforçam a viabilidade e a importância da utilização de redes neurais convolucionais para a classificação automática de grãos de soja, evidenciando os avanços e desafios inerentes à aplicação da inteligência artificial nesse contexto. A análise comparativa entre diferentes modelos demonstrou que as redes pré-treinadas superam significativamente as redes customizadas em termos de acurácia e capacidade de generalização. Em particular, o modelo InceptionV3 destacou-se como a alternativa mais promissora, atingindo 95,9% de acurácia e evidenciando sua aptidão para tarefas de classificação de alta precisão em contextos industriais.

As CNNs customizadas, apesar de oferecerem controle sobre a arquitetura e menor custo computacional, apresentaram um desempenho insatisfatório, com a maior acurácia atingindo apenas 16,6%. Esse resultado indica a necessidade de ajustes adicionais e a implementação de técnicas mais avançadas para evitar problemas como overfitting e baixa capacidade de generalização. Já o modelo ResNet50, embora tenha alcançado uma acurácia intermediária de 63,1%, mostrou-se uma opção viável para aplicações que exigem um equilíbrio entre desempenho e custo computacional, sobretudo em ambientes de processamento remoto ou em linhas de produção com infraestrutura computacional adequada.

Apesar dos avanços apresentados, este trabalho enfrentou limitações tanto de ordem computacional quanto estrutural. Um dos principais desafios identificados foi a necessidade de um banco de dados mais amplo e diversificado. A inclusão de imagens capturadas em diferentes condições ambientais e etapas da cadeia produtiva poderia aumentar a acurácia e a capacidade de generalização dos modelos, tornando-os mais eficazes em cenários reais.

Outro ponto relevante foi a criação de um dataset, formado por um extenso conjunto de dados com imagens de grãos de soja contendo diversos tipos de defeitos. Esse dataset representa uma importante contribuição para a área, pois poderá servir como fonte para futuras pesquisas e a novos estudos que objetivem à identificação e análise de defeitos em grãos com o uso de visão computacional e redes neurais.

O aplicativo desenvolvido se mostrou uma ferramenta eficaz para apoiar a análise de grãos de soja, unindo praticidade de uso com recursos avançados de processamento de imagem. Sua interface simples, aliada à precisão dos resultados,

torna o sistema útil tanto para atividades técnicas quanto para aplicações mais amplas no setor agrícola. Ao automatizar etapas antes feitas manualmente, o aplicativo contribui para tornar a avaliação de grãos mais ágil, objetiva e padronizada, destacando-se como uma solução promissora para quem busca maior confiabilidade no controle de qualidade.

Esta pesquisa representa um avanço significativo no desenvolvimento de um sistema automatizado para a classificação de grãos de soja, contribuindo para a modernização do setor agroindustrial e para a implementação de práticas alinhadas aos princípios da agricultura 4.0. No entanto, ajustes adicionais ainda são necessários para garantir um modelo plenamente funcional e aplicável em larga escala. As próximas etapas do estudo serão fundamentais para consolidar a tecnologia e possibilitar sua futura implementação no setor produtivo.

REFERÊNCIAS

- AFSHAR, P.; PLATANIOTIS, K. N.; MOHAMMADI, A. Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. **In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)**, 2019. **Proceedings [...]**. Piscataway: IEEE, 2019. p. 1368–1372. DOI: 10.1109/ICASSP.2019.8683759.
- AGÊNCIA EY. **A visão computacional na transformação para a agropecuária 4.0**. 2022. Disponível em: <https://revistalide.com.br/a-visao-computacional-na-transformacao-para-a-agropecuaria-40>. Acesso em: 20 mar. 2025.
- AGROADVANCE. **Subprodutos da soja**: conheça os destinos e usos da soja brasileira. 2023. Disponível em: <https://agroadvance.com.br/blog-subprodutos-da-soja/>. Acesso em: 18 mar. 2025.
- AYENI, J. A. **Convolutional Neural Network (CNN)**: the architecture and applications. 2022.
- BARELLI, F. **Introdução à visão computacional**: uma abordagem prática com Python e OpenCV. São Paulo: Casa do Código, 2018.
- BEZDAN, T.; DŽAKULA, N. B. Convolutional neural network layers and architectures. **In: SINTEZA 2019 – INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION TECHNOLOGY AND DATA RELATED RESEARCH**, 2019, Belgrado. **Anais [...]**. Belgrado: Singidunum University, 2019. p. 445–451. DOI: 10.15308/Sinteza-2019-445-451.
- BIP BRASIL. **A quarta revolução industrial e a agricultura 4.0**: insights e exemplos de aplicações da tecnologia. 2022. Disponível em: <https://bipbrasil.com.br/a-quarta-revolucao-industrial-e-a-agricultura-4-0-insights-e-exemplos-de-aplicacoes-da-tecnologia/>. Acesso em: 20 mar. 2025.
- BRASIL. **Instrução Normativa n.º 11, de 2007**. Brasília, DF: Ministério da Agricultura, Pecuária e Abastecimento, 2007a.
- BRASIL. **Instrução Normativa n.º 37, de 2007**. Brasília, DF: Ministério da Agricultura, Pecuária e Abastecimento, 2007b.
- BRASIL. Ministério da Agricultura e Pecuária. **Projeções do agronegócio: 2023/2024 a 2033/2034**. Brasília: MAPA, 2023. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/todas-publicacoes-de-politica-agricola/projecoes-do-agronegocio/projecoes-do-agronegocio-2022-2023-a-2032-2033.pdf>. Acesso em: 27 mar. 2024.
- EMBRAPA. **Manual de segurança e qualidade para a cultura da soja**. Brasília, DF: Embrapa Transferência de Tecnologia: CampoPAS, 2005.

EMBRAPA. **Soja em números (safra 2023/24)**. Disponível em: <https://www.embrapa.br/soja/cultivos/soja1/dados-economicos>. Acesso em: 15 mar. 2025.

FEHR, W. R.; CAVINESS, C. E. **Stages of soybean development**. Ames: Iowa State University of Science and Technology, 1977. 11 p. (Special Report, 80).

FURTADO, M. I. V. **Redes neurais artificiais: uma abordagem para sala de aula** [recurso eletrônico]. Ponta Grossa: Atena Editora, 2019.

GANDHARV, M. **Convolutional Neural Network (CNN): The architecture and applications**. 2022.

GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. 3. ed. Upper Saddle River: Pearson Education, 2010.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Cambridge: MIT Press, 2016.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.

HE, K. et al. **Deep residual learning for image recognition**. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

HIRAKURI, M. H. et al. **Análise de aspectos econômicos sobre a qualidade de grãos de soja no Brasil**. Londrina: Embrapa Soja, 2018. (Circular técnica, 145).

HUANG, Z. et al. **Deep learning based soybean seed classification**. Computers and Electronics in Agriculture, [S.l.], v. 202, p. 107393, nov. 2022. DOI: 10.1016/j.compag.2022.107393.

IOFFE, S.; SZEGEDY, C. **Batch normalization: Accelerating deep network training by reducing internal covariate shift**. In: Proceedings of the 32nd International Conference on Machine Learning (ICML), 2015.

KOHAVI, R. **A study of cross-validation and bootstrap for accuracy estimation and model selection**. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995.

KOVÁCS, Z. L.. **Redes neurais artificiais: fundamentos e aplicações**. 4. ed. São Paulo: LF Editorial, 2023.

LI, Q. et al. Medical image classification with convolutional neural network. In: INTERNATIONAL CONFERENCE ON CONTROL AUTOMATION ROBOTICS AND VISION (ICARCV), 13., 2014, Singapura. **Proceedings [...]**. [S. l.]: IEEE, 2014. p. 844–848. DOI: 10.1109/ICARCV.2014.7064414.

LIN, W. et al. Online classification of soybean seeds based on deep learning. **Engineering Applications of Artificial Intelligence**, [S.I.], v. 123, p. 106434, ago. 2023. DOI: 10.1016/j.engappai.2023.106434.

LORINI, I. (Ed.). **Qualidade de sementes e grãos comerciais de soja no Brasil – safra 2016/17**. Londrina: Embrapa Soja, 2018.

MACHADO, F. C. **Classificação de grãos de soja utilizando redes neurais convolucionais e aumento de dados**. 2021. Dissertação (Mestrado em Engenharia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2021.

MAYANK, M. **Convolutional neural networks, explained**. 2020. Disponível em: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>. Acesso em: 01 nov. 2024.

MILANO, D.; HONORATO, L. B. **Visão computacional**. Disponível em: https://www.academia.edu/9621896/VIS%C3%83O_COMPUTACIONAL_Palavras_Chaves. Acesso em: 19 mar. 2025.

NAZEER, R. et al. Detection of cotton leaf curl disease's susceptibility scale level based on deep learning. **Journal of Cloud Computing**, [S. I.], v. 13, p. 50, 2024. <https://doi.org/10.1186/s13677-023-00582-9>

NEUMAIER, N. et al. Estádios de desenvolvimento da cultura de soja. In: BONATO, Elio R. (org.). **Estresses em soja**. Londrina: Embrapa Soja, 2000. p. 19–44.

OHRI, A. Performance Analysis of Sigmoid and ReLU Activation Functions in Deep Neural Networks. In: WAOO, Akhilesh A.; SONI, Brijesh K. (ed.). **Intelligent Systems**. Singapore: Springer, 2021. p. 39–52. (Algorithms for Intelligent Systems). DOI: 10.1007/978-981-16-2248-9_5.

PARAGINSKI, R. T. et al. Qualidade de grãos de milho armazenados em diferentes temperaturas. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 19, n. 4, p. 358–363, 2015. DOI: 10.1590/1807-1929/agriambi.v19n4p358-363.

PARK, D. S. et al. Multiclass classification of crop diseases using RGB and multispectral images. **Computers and Electronics in Agriculture**, v. 153, p. 318–325, 2018.

PEDAMKAR, P. **Convolutional Neural Network (CNN): The architecture and applications**. 2021.

POWERS, D. M. W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. **Journal of Machine Learning Technologies**, v. 2, n. 1, p. 37–63, 2011.

RAMOS, A. H. **Implicações dos defeitos fermentado, ardido e queimado em grãos de soja no rendimento e na qualidade lipídica e proteica**. 2019. 82 f.

Dissertação (Mestrado em Agronomia) – Universidade Federal de Pelotas, Pelotas, 2019.

REDMON, J. et al. **You only look once**: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

SABLE, A. et al. Quantifying soybean defects: a computational approach to seed classification using deep learning techniques. **Agronomy**, Basel, v. 14, n. 6, p. 1098, 2024. DOI: 10.3390/agronomy14061098.

SAITO, T. et al. Soybean seed quality evaluation using multi-input convolutional neural networks based on color and fluorescent images. **Biosystems Engineering**, v. 230, p. 48–59, 2024. DOI: 10.1016/j.biosystemseng.2023.10.006.

SANTOS, F. R.; OLIVEIRA, M. S.; CUNHA, F. F. Aplicações da visão computacional na agricultura de precisão. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 25, n. 1, p. 45–52, 2021. DOI: <https://doi.org/10.1590/1807-1929/agriambi.v25n1p45-52>. Acesso em: 10 dez. 2024.

SEDIYAMA, T. **Tecnologia da produção de soja**. Londrina: Mecenas, 2009.

SEIXAS, C. D. S. et al. (ed.). **Tecnologias de produção de soja**. Londrina: Embrapa Soja, 2020. 347 p. (Sistemas de Produção / Embrapa Soja, n. 17).

SHAMSALDIN, R.; ALI, H.; MOHAMMED, A. Speech Emotion Recognition Using Deep 1D & 2D CNN LSTM Networks. **Biomedical Signal Processing and Control**, v. 47, p. 312–323, 2019. DOI: 10.1016/j.bspc.2018.08.035.

SHARMA, P. Convolutional Neural Networks for Breast Cancer Detection in Mammography: A Survey. **Computer Methods and Programs in Biomedicine**, v. 200, p. 105843, 2021. DOI: 10.1016/j.cmpb.2021.105843.

SILVA, A. P.; ROCHA, D. M.; OLIVEIRA, L. S. Classificação de grãos utilizando técnicas de processamento de imagens e aprendizado de máquina. **Revista Engenharia na Agricultura**, Viçosa, v. 26, n. 2, p. 140–148, 2018. DOI: <https://doi.org/10.13083/reveng.v26i2.865>. Acesso em: 10 dez. 2024.

SILVA, F. et al. (org.). **Soja: do plantio à colheita**. 2. ed. São Paulo: Oficina de Textos, 2022.

SILVA, S. P. da et al. **Redes neurais artificiais**: para engenharia e ciências aplicadas. 2. ed. São Paulo: LTC, 2016.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, v. 45, n. 4, p. 427–437, 2009.

SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

SZELISKI, R. **Computer vision**: algorithms and applications. 2. ed. Cham: Springer, 2022. (Texts in Computer Science).

THOMAS, A. L.; COSTA, J. A. **Soja**: manejo para alta produtividade de grãos. Porto Alegre: Evangraf, 2010.

USDA – U.S. Department of Agriculture. **Production – Soybeans**. Disponível em: <https://www.fas.usda.gov/data/production/commodity/2222000>. Acesso em: 18 mar. 2025.

VAMATHEVAN, J. et al. Applications of machine learning in drug discovery and development. **Nature Reviews Drug Discovery**, v. 18, n. 6, p. 463–477, 2019. DOI: 10.1038/s41573-019-0024-5.

VAZ, J.; BALAJI, S. Convolutional neural networks (CNNs): concepts and applications in pharmacogenomics. **Molecular Diversity**, v. 25, n. 3, p. 1569–1584, 2021. DOI: 10.1007/s11030-021-10225-3.

WANG, W.; GANG, J. Application of convolutional neural network in natural language processing. In: INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS AND COMPUTER AIDED EDUCATION (ICISCAE), 2018, Changchun. **Proceedings [...]**. [S. l.]: IEEE, 2018. p. 64–70. DOI: 10.1109/ICISCAE.2018.8666928.

YADAV, S. S.; JADHAV, S. M. Deep convolutional neural network based medical image classification for disease diagnosis. **Journal of Big Data**, [S.l.], v. 6, n. 1, p. 1–18, 2019

YAMASHITA, R.; NISHIO, M.; DO, R. K. G.; TOGASHI, K. Convolutional neural networks: an overview and application in radiology. **Insights into Imaging**, [S.l.], v. 9, n. 4, p. 611–629, 2018. DOI: <https://doi.org/10.1007/s13244-018-0639-9>.

ZHAO, J. et al. Development of a real-time classification system for recognizing surface defects on soybean seeds based on deep learning and rotating imaging. **Computers and Electronics in Agriculture**, v. 187, p. 106259, 2021. DOI: 10.1016/j.compag.2021.106259.