



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS TERESINA CENTRAL
TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

JOÃO PEDRO DE MESQUITA VIEIRA

**CLASSIFICAÇÃO DE GRÃOS DE SOJA UTILIZANDO REDES NEURAIAS
CONVOLUCIONAIS: uma análise comparativa entre modelos customizados e
pré-treinados**

TERESINA, PIAUÍ
2025

JOÃO PEDRO DE MESQUITA VIEIRA

CLASSIFICAÇÃO DE GRÃOS DE SOJA UTILIZANDO REDES NEURAIIS
CONVOLUCIONAIS: uma análise comparativa entre modelos customizados e
pré-treinados

Trabalho de Conclusão de Curso (monografia) apresentado como exigência parcial para obtenção do diploma do Curso de Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central.

Orientador: Prof. Dr. Otílio Paulo da Silva Neto

TERESINA, PIAUÍ
2025

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

Vieira, João Pedro de Mesquita

V657c Classificação de grãos de soja utilizando redes neurais convolucionais :
uma análise comparativa entre modelos customizados e pré-treinados / João
Pedro de Mesquita Vieira. - 2025.
56 f.: il. color.

Trabalho de Conclusão de Curso (Tecnologia em Análise e
Desenvolvimento de Sistemas) - Instituto Federal de Educação, Ciência e
Tecnologia do Piauí, Campus Teresina Central, 2025.

Orientador : Prof Me. Otílio Paulo da Silva Neto.

1. Redes neurais. 2. YOLO - Métodos de detecção de objetos. 3. Soja -
Classificação de grãos. I.Título.

CDD - 004.21

Elaborado por Sindya Santos Melo CRB 3/1085

JOÃO PEDRO DE MESQUITA VIEIRA

CLASSIFICAÇÃO DE GRÃOS DE SOJA UTILIZANDO REDES NEURAIIS
CONVOLUCIONAIS: uma análise comparativa entre modelos customizados e
pré-treinados

Trabalho de Conclusão de Curso (monografia) apresentado como exigência parcial para obtenção do diploma do Curso de Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central.

Aprovada em 17 de Fevereiro de 2025.

BANCA EXAMINADORA:

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

Prof. M^e. Duany Dreyton Bezerra Sousa
Instituto Federal do Piauí (IFPI)

Prof. M^e. Rogério da Silva Batista
Instituto Federal do Piauí (IFPI)

Dedico este trabalho a Deus, à minha mãe, à meu pai, ao meu irmão, ao meu primo-irmão, à família de círculo próximo, à minha família, àqueles que considero família e a mim. E aos justos que caminham sobre esta terra, que por acaso encontrem estas palavras: que jamais percam a vontade de avançar, de resistir e de prosperar. O caminho é árduo. A vida é dura. O mundo, muitas vezes, é injusto. Mas não vendam suas almas, não traiam seus princípios, não sacrifiquem quem vocês são. Permaneçam íntegros, vivam com dignidade e honrem cada passo dado. A vida é breve, mas que nela haja sentido e verdade.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por ter me dado força, resiliência e sabedoria ao longo dessa caminhada. Sem Sua presença, muitos momentos teriam sido ainda mais desafiadores. À minha família primária que, ao seu modo, me ofereceram apoio incondicional nessa e em outras trajetórias, pelo amor, e pelas palavras de encorajamento que me impulsionaram nos momentos de dúvida. À minha família e aos que considero família, meu sincero reconhecimento pelo suporte e pela amizade.

Por fim, agradeço a mim, pela persistência, pela capacidade de reaprender e por acreditar que cada esforço valeria a pena. Este trabalho é fruto não apenas de conhecimento e dedicação, mas também de superação e luta contra demônios internos.

*Felizes são aqueles que não apenas esquecem seus erros, mas que os compreendem
e os transformam em aprendizado, pois é na falha que se constrói a verdadeira
sabedoria.*

(Inspirado em Nietzsche)

RESUMO

Este estudo investiga a aplicação de redes neurais convolucionais (CNNs) na classificação de grãos de soja, com ênfase na distinção entre grãos saudáveis, defeituosos e danificados. Considerando a importância da soja como uma das culturas agrícolas mais relevantes, amplamente utilizada na alimentação humana e animal, bem como em diversos setores industriais, garantir a qualidade dos grãos é essencial para manter a eficiência e a competitividade das cadeias produtivas. Diversas arquiteturas de CNN foram analisadas, incluindo modelos personalizados e redes pré-treinadas, como a ResNet50 e InceptionV3. Para o pré-processamento das imagens, foi utilizada segmentação com o modelo YOLO, seguida por técnicas de recorte, normalização e aumento de dados para aumentar a robustez dos modelos. Os resultados experimentais mostraram que os modelos pré-treinados apresentaram desempenho significativamente superior aos modelos personalizados, com o modelo InceptionV3 atingindo uma acurácia de 95,9%. Em contrapartida, os modelos personalizados apresentaram uma acurácia máxima de apenas 16,6%. Esses achados destacam a eficácia dos modelos pré-treinados, especialmente em tarefas de classificação de imagens com variações sutis, como é o caso da identificação de diferentes tipos de grãos. Com base nos resultados obtidos, este estudo proporciona uma base sólida para o desenvolvimento futuro de sistemas automatizados de classificação de grãos em tempo real, visando otimizar a eficiência dos processos de inspeção e melhorar a qualidade do monitoramento agrícola.

Palavras-chaves: classificação de grãos de soja; redes neurais convolucionais; aprendizado profundo; YOLO.

ABSTRACT

This study investigates the application of convolutional neural networks (CNNs) in soybean grain classification, with an emphasis on distinguishing between healthy, defective, and damaged grains. Considering the importance of soybeans as one of the most relevant agricultural crops, widely used in human and animal nutrition as well as in various industrial sectors, ensuring grain quality is essential to maintaining the efficiency and competitiveness of production chains. Several CNN architectures were analyzed, including customized models and pre-trained networks such as ResNet50 and InceptionV3. For image pre-processing, segmentation was performed using the YOLO model, followed by cropping, normalization, and data augmentation techniques to enhance the robustness of the models. Experimental results showed that pre-trained models performed significantly better than customized models, with InceptionV3 achieving an accuracy of 95.9%. In contrast, customized models achieved a maximum accuracy of only 16.6%. These findings highlight the effectiveness of pre-trained models, especially in image classification tasks involving subtle variations, such as identifying different types of grains. Based on the results obtained, this study provides a solid foundation for the future development of automated real-time grain classification systems, aiming to optimize inspection processes and improve the quality of agricultural monitoring.

Keywords: soybean grain classification; convolutional neural networks; deep learning; YOLO.

LISTA DE ILUSTRAÇÕES

Figura 1 – Comparação entre imagens de grãos de soja	22
Figura 2 – Comparação entre imagens colorida e em escala de cinza de grãos de soja	24
Figura 3 – Exemplo simplificado do processo de convolução	26
Figura 4 – Exemplo de feature maps extraídos do bloco 1 do modelo VGG16 .	27
Figura 5 – Exemplo de operação de max pooling	28
Figura 6 – Gráfico da função ReLU	28
Figura 7 – Diagrama de uma camada densa	30
Figura 8 – Ilustração do funcionamento do Dropout	31
Figura 9 – Exemplos de imagens de grãos de soja por categorias	37
Figura 10 – Exemplos de imagens de grãos de soja cortados	38
Figura 11 – Métricas do modelo YOLO: (a) Total Loss, (b) mAP0.5	43
Figura 12 – Métricas do modelo YOLO: (a) Recall, (b) Precision	44
Figura 13 – Matriz de confusão do modelo YOLO	45
Figura 14 – Gráficos de treinamento do melhor modelo personalizado	46
Figura 15 – Matriz de confusão do melhor modelo customizado	47
Figura 16 – Gráficos de acurácia e perda (loss) do modelo ResNet50 durante o treinamento e validação	49
Figura 17 – Matriz de confusão do modelo ResNet50	50
Figura 18 – Gráficos de acurácia e perda (loss) do modelo InceptionV3 durante o treinamento e validação	51
Figura 19 – Matriz de confusão do modelo InceptionV3	52

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
CNNs	Convolutional Neural Networks
YOLO	You Only Look Once
RGB	Canais de três cores Red, Green, Blue
RELU	Rectified Linear Unit
ADAM	Adaptative Moment Estimation
mAP	Mean Average Precision

LISTA DE SÍMBOLOS

%	Porcentagem
---	-------------

SUMÁRIO

1	INTRODUÇÃO	14
1.1	IMPORTÂNCIA DA SOJA E O PROCESSO DE CLASSIFICAÇÃO	14
1.2	APLICAÇÕES DE REDES NEURAIS CONVOLUCIONAIS (CNN) NA AGRICULTURA	15
1.3	OBJETIVOS	16
1.3.1	Objetivo Geral	16
1.3.2	Objetivos Específicos	16
1.4	ESTRUTURA DA MONOGRAFIA	16
2	REVISÃO DE LITERATURA	18
2.1	CNN NA CLASSIFICAÇÃO DE GRÃOS	18
2.2	YOLO E SEGMENTAÇÃO DE IMAGENS NA AGRICULTURA	19
3	FUNDAMENTAÇÃO TEÓRICA	20
3.1	CLASSIFICAÇÃO DOS GRÃOS DE SOJA: CRITÉRIOS E PADRÕES	20
3.2	SEGMENTAÇÃO E PRÉ-PROCESSAMENTO DE IMAGENS NA VISÃO COMPUTACIONAL	21
3.2.1	Segmentação com YOLO	21
3.2.2	Recorte e Redimensionamento de Imagens	22
3.2.3	Conversão para Escala de Cinza	23
3.2.4	Normalização das Imagens	24
3.2.5	Aumento de Dados (Data Augmentation)	25
3.3	ARQUITETURA E FUNCIONAMENTO DAS REDES NEURAIS CONVOLUCIONAIS (CNN)	25
3.3.1	Camadas Convolucionais	26
3.3.2	Camadas de Pooling	27
3.3.3	Funções de Ativação	28
3.3.3.1	ReLU	28
3.3.3.2	Softmax	29
3.3.4	Camadas Densas	29
3.3.5	Camadas de Dropout	30
3.3.6	Fluxo Simplificado de Processamento da CNN	31
3.4	ARQUITETURAS DE REDES NEURAIS PARA CLASSIFICAÇÃO DE IMAGENS	31
3.4.1	Arquiteturas Personalizadas	32
3.4.2	Arquiteturas Pré-Treinadas	32

3.4.2.1	ResNet50	32
3.4.2.2	InceptionV3	32
3.5	TREINAMENTO DE MODELOS DE REDES NEURAIIS	33
3.5.1	Otimizador Adam	33
3.5.2	Taxa de Aprendizado (Learning Rate)	33
3.5.3	Função de Perda: Categorical Crossentropy	34
3.5.4	Épocas e Batch Size	34
3.5.5	Métricas de avaliação	34
4	METODOLOGIA	36
4.1	COLETA E ORGANIZAÇÃO DO DATASET	36
4.2	PRÉ-PROCESSAMENTO DAS IMAGENS	37
4.2.1	Segmentação e Recorte	37
4.2.2	Conversão, Normalização e Aumento de Dados	38
4.3	PROCESSAMENTO DAS REDES NEURAIIS CONVOLUCIONAIS (CNN)	38
4.4	ARQUITETURAS TESTADAS	39
4.4.1	Arquiteturas Personalizadas	39
4.4.2	Arquiteturas Pré-treinadas	39
4.5	TREINAMENTO DOS MODELOS	40
4.5.1	Configuração do treinamento das CNNs Personalizadas	40
4.5.2	Configuração do treinamento das CNNs Pré-treinadas	40
4.5.3	Monitoramento durante o treinamento	41
5	RESULTADOS E DISCUSSÕES	42
5.1	COMPARAÇÃO DO DESEMPENHO DAS ARQUITETURAS	42
5.2	IMPLICAÇÕES E UTILIZAÇÕES PRÁTICAS	42
5.3	LIMITAÇÕES	53
6	CONCLUSÃO	54
7	TRABALHOS FUTUROS	55
	REFERÊNCIAS	56

1 INTRODUÇÃO

A agricultura desempenha um papel fundamental na economia global, sendo responsável pela produção de alimentos, matérias-primas e pelo desenvolvimento de tecnologias para otimização do cultivo e da colheita. Com o crescimento populacional e a necessidade de aumentar a eficiência produtiva, novas abordagens tecnológicas vêm sendo incorporadas ao setor agrícola, buscando minimizar perdas e garantir maior qualidade nos produtos finais.

Nesse contexto de eficiência produtiva, a soja (*Glycine max* (L.) Merrill) se destaca como uma das culturas mais relevantes, contribuindo significativamente, pela sua versatilidade que permite o aproveitamento do grão não apenas na alimentação humana e animal, na forma de farelo, óleo e farinha, mas também na produção de biodiesel, plásticos biodegradáveis, cosméticos (PETERSEN, 2022) (BOSCHIERO, 2023). A variedade de aplicações, evidencia sua importância econômica e industrial que, por sua vez, exigem processos produtivos cada vez mais eficientes e padronizados, especialmente em etapas críticas como a classificação de grãos.

Nos últimos anos, a Inteligência Artificial (IA) tem ganhado destaque na automação de processos agrícolas, permitindo análises mais precisas e tomadas de decisão mais eficientes. Uma das áreas que mais se beneficiam dessas inovações é a classificação de grãos e sementes, etapa essencial para garantir a qualidade dos produtos comercializados (EMBRAPA, 2001). Tradicionalmente, a inspeção de grãos de soja, o objeto desse estudo, é realizada manualmente, tornando-se um processo demorado, subjetivo e mais propenso a erros humanos. Métodos automatizados, especialmente aqueles baseados em visão computacional e aprendizado profundo (**deep learning**), surgem como soluções promissoras para aprimorar esse processo.

1.1 IMPORTÂNCIA DA SOJA E O PROCESSO DE CLASSIFICAÇÃO

A soja, é uma das principais culturas agrícolas do Brasil, não apenas pelo seu valor nutricional, mas também por sua ampla aplicação industrial. Rica em proteínas, óleos insaturados e compostos bioativos, ela se destaca na alimentação humana e animal, na produção de biodiesel, cosméticos, plásticos biodegradáveis e produtos farmacêuticos, contribuindo de forma significativa para o agronegócio nacional e para a balança comercial do país (FRANÇA-NETO; HENNING; KRYZANOWSKI, 2011) (PETERSEN, 2022). Essa diversidade de usos exige cadeias produtivas eficientes, onde o controle de qualidade exerce um papel central.

No Brasil, a classificação de grãos de soja ainda é, majoritariamente manual, com base em inspeções visuais normatizadas pela Instrução Normativa N° 11/2007 do

Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Apesar de padronizada, essa abordagem está sujeita à interpretação dos classificadores e à limitação de escalabilidade em grandes volumes de produção (MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO, 2007).

Diante desse cenário, a adoção de métodos automatizados baseados em visão computacional, especialmente aqueles suportados por redes neurais convolucionais (CNNs), despontam como alternativa promissora para aprimorar a eficiência e a padronização do processo classificatório, conforme discutido a seguir.

1.2 APLICAÇÕES DE REDES NEURAI CONVOLUCIONAIS (CNN) NA AGRICULTURA

Dentre as técnicas de aprendizado de máquina, as Redes Neurais Convolucionais (CNNs) se destacam por sua capacidade de analisar imagens e identificar padrões complexos, tornando-se ferramentas poderosas para diversas aplicações na agricultura. Estudos recentes mostram que CNNs podem ser aplicadas em tarefas como detecção de doenças em plantações (BEVERS; SIKORA; HARDY, 2022), monitoramento de pragas (RIBEIRO *et al.*, 2022) e segmentação de frutas e sementes (ZHANG; ZANG; LI, 2023).

No contexto da classificação de grãos, o uso de CNNs tem se mostrado particularmente eficaz, permitindo identificar defeitos em sementes e classificá-las de acordo com critérios de qualidade pré-estabelecidos. Trabalhos como o de (YANG *et al.*, 2024) demonstram que a utilização de CNNs combinadas com imagens hiperespectrais pode aumentar a precisão na identificação de sementes defeituosas. Já (ZHANG; ZANG; LI, 2023) mostram que técnicas de balanceamento de dados e redes profundas podem reduzir vieses de classificação, garantindo maior confiabilidade nos resultados.

Além disso, avanço em modelos pré-treinados como **ResNet50** e **InceptionV3** vêm permitindo que CNNs alcancem performances superiores em tarefas de classificação (KAMILARIS; BOLDÚ, 2018). Esses modelos são treinados previamente em grandes bases de dados, como o ImageNet, permitindo que redes neurais utilizem pesos já otimizados para reconhecer padrões em novas imagens, mesmo em domínios diferentes do seu treinamento original. Isso reduz significativamente o tempo de treinamento e melhora a generalização, tornando os modelos pré-treinados uma opção mais rapidamente viável para aplicações agrícolas.

Outro avanço importante para a classificação de grãos é o uso de modelos de segmentação de imagens, como o **YOLO** (You Only Look Once). O YOLO é um dos métodos mais eficientes para detecção de objetos, permitindo identificar e segmentar grãos em imagens de forma rápida e precisa (REDMON *et al.*, 2016). Essa técnica possibilita isolar os grãos do fundo da imagem, reduzindo ruídos e otimizando o desempenho das redes neurais classificatórias. Estudos como os de (ALMEIDA *et al.*, 2024) mostram que o uso do YOLO pode melhorar significativamente a acurácia da

classificação ao garantir que apenas os elementos de interesse sejam analisados pelo modelo.

Diante desse cenário, a combinação de segmentação com YOLO e classificação com CNNs representa uma solução promissora para a inspeção automatizada de grãos, reduzindo custos operacionais e aumentando a confiabilidade do processo.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Avaliar a eficácia das redes neurais convolucionais na classificação de grãos de soja, comparando o desempenho entre modelos customizados e modelos pré-treinados, e propor um pipeline automatizado que integre técnicas de segmentação e classificação para aplicação prática na inspeção de qualidade.

1.3.2 Objetivos Específicos

- Desenvolver e treinar modelos personalizados de redes neurais convolucionais para a classificação de grãos de soja
- Analisar o impacto do uso da segmentação por YOLO no desempenho dos modelos de classificação
- Comparar as abordagens quanto à acurácia, custo computacional e viabilidade prática
- Criar e organizar um dataset anotado de grãos de soja para treinamento e validação dos modelos.
- Estruturar um sistema automatizado com foco em eficiência computacional e confiabilidade.

1.4 ESTRUTURA DA MONOGRAFIA

A organização desta monografia está estruturada da seguinte maneira:

- No **Capítulo 2**, é realizada uma revisão de literatura abordando a aplicação da inteligência artificial na agricultura, com ênfase em redes neurais convolucionais e técnicas de segmentação de imagens, como o modelo YOLO.
- O **Capítulo 3** apresenta a fundamentação teórica, explicando os conceitos essenciais relacionados às redes neurais convolucionais, funções de ativação, camadas de pooling e dropout, além dos princípios da segmentação de imagens e das arquiteturas utilizadas, como ResNet50 e InceptionV3.

- O **Capítulo 4** apresenta a metodologia adotada no estudo, detalhando o processo de coleta de dados, o pré-processamento das imagens, o treinamento dos modelos e as métricas utilizadas para a avaliação do desempenho.
- No **Capítulo 5**, são discutidos os resultados obtidos, analisando o desempenho das diferentes arquiteturas testadas, bem como suas vantagens e limitações.
- O **Capítulo 6** sintetiza as conclusões do estudo, destacando as principais contribuições e implicações dos achados.
- Por fim, no **Capítulo 7**, são apresentadas sugestões para trabalhos futuros, explorando possíveis aprimoramentos e novas direções de pesquisa.

2 REVISÃO DE LITERATURA

A aplicação de IA na agricultura tem ganhado destaque na otimização de processos produtivos, reduzindo a dependência de métodos manuais para inspeção de qualidade. Técnicas como **visão computacional** e **aprendizado profundo** permitem a automação de análise de sementes, promovendo maior eficiência e padronização. Dentro desse contexto, CNNs se tornaram uma das abordagens mais eficazes para tarefas de classificação, enquanto algoritmos de detecção de objetos, como o YOLO, são empregados para segmentação e identificação de elementos relevantes em imagens agrícolas.

A qualidade dos grãos de soja é um fator determinante para a competitividade e eficiência da cadeia produtiva, impactando diretamente na comercialização e no aproveitamento industrial dessa cultura. As limitações dos métodos tradicionais de classificação, predominantemente manuais e sujeitos à subjetividade, impulsionam a adoção de técnicas automatizadas baseadas em inteligência artificial. Nesse contexto, destacam-se as redes neurais convolucionais (CNNs) e os algoritmos de segmentação, como o YOLO, que serão abordados a seguir nesta revisão de literatura.

Este capítulo apresenta o estado da arte das tecnologias aplicadas neste estudo, abordando o uso de CNNs na classificação de grãos, as vantagens de modelos pré-treinados e a importância da segmentação para otimizar o desempenho dos classificadores.

2.1 CNN NA CLASSIFICAÇÃO DE GRÃOS

As CNNs são um tipo de modelo de aprendizado profundo desenvolvido para análise de imagens. Sua arquitetura é composta por **camadas convolucionais**, que extraem características visuais como bordas, texturas e formas, seguidas por camadas de pooling, que reduzem a dimensionalidade dos dados sem perder informações relevantes (LECUN *et al.*, 1998).

No setor agrícola, CNNs vêm sendo utilizadas para a classificação de sementes e grãos, permitindo a detecção de defeitos com maior precisão e consistência do que métodos tradicionais. Estudos recentes, como os de (ZHANG; ZANG; LI, 2023), mostram que CNNs aplicadas à classificação de soja podem alcançar precisão superior a 98% quando combinadas com técnicas de balanceamento de dados e aprendizado profundo. Já (YANG *et al.*, 2024) destacam que a incorporação de imagens hiperespectrais melhora a diferenciação entre grãos saudáveis e defeituosos, tornando o processo ainda mais confiável.

As CNNs podem ser implementadas de duas formas principais:

- **Modelos personalizados:** Desenvolvidos do zero e treinados especificamente para o problema em questão, exigindo um volume considerável de dados e ajustes finos de hiperparâmetros.
- **Modelos pré-treinados:** Arquiteturas previamente treinadas em grandes bases de dados, como **ResNet50** e **InceptionV3**, que podem ser ajustadas para novas tarefas por meio de transferência de aprendizado (**transfer learning**), que permite aproveitar conhecimentos anteriores adquiridos em tarefas de aprendizado semelhantes (GOMES *et al.*, 2023), acelerando o treinamento e melhorando a generalização (KAMILARIS; BOLDÚ, 2018).

Trabalhos na área indicam que modelos pré-treinados tendem a superar arquiteturas personalizadas em termos de precisão e eficiência computacional, especialmente quando há um volume limitado de dados disponíveis (SZEGEDY *et al.*, 2016). No presente estudo, essa abordagem foi explorada para avaliar o impacto do transfer learning na classificação geral de grãos de soja.

2.2 YOLO E SEGMENTAÇÃO DE IMAGENS NA AGRICULTURA

A **segmentação de imagens** é uma etapa crucial na análise automatizada de grãos, pois garante que apenas os elementos de interesse sejam processados pelos classificadores. Um dos métodos mais eficientes para essa tarefa é o YOLO, um modelo baseado em aprendizado profundo que permite a detecção e segmentação de objetos em tempo real (REDMON *et al.*, 2016).

Estudos como os de (ALMEIDA *et al.*, 2024) (CARBONI, 2021) demonstram que o uso do YOLO pode melhorar significativamente a classificação de grãos ao eliminar elementos irrelevantes da imagem, reduzindo a interferência de sombras, sobreposições e variações de iluminação. Essa abordagem tem sido amplamente aplicada na segmentação de sementes, identificação de pragas e monitoramento de lavouras, tornando-se um componente essencial em sistemas de visão computacional para a agroindústria.

Neste estudo, o YOLO foi empregado para isolar os grãos de soja antes da etapa de classificação, garantindo que as CNNs trabalhassem apenas com as regiões relevantes da imagem. Essa estratégia contribuiu para a minimizar ruídos e otimizar a eficiência computacional.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 CLASSIFICAÇÃO DOS GRÃOS DE SOJA: CRITÉRIOS E PADRÕES

Este estudo utilizou como referência os critérios definidos pelo Referencial Fotográfico dos Defeitos da Soja – 3ª Edição (BRASIL, 2008), "Referenciais fotográficos de grãos, Aplicados à classificação de produtos de origem vegetal"(BRASIL, 2023) e pelo artigo "Defeitos da Classificação Comercial da Soja Devido a Danos Mecânicos e Fungos"(FRANÇA-NETO; HENNING; KRYZANOWSKI, 2011), que padronizam a identificação e diferenciação entre os tipos de grãos. A seguir, são descritas as categorias de grãos consideradas neste estudo:

- **Grãos Ardidos:** Caracterizam-se por uma coloração marrom a avermelhada, resultante de processos fermentativos ou deterioração causada por umidade excessiva e temperatura inadequada durante a armazenagem. Esses grãos podem ter sua qualidade nutricional comprometida e, em casos graves, tornam-se impróprios para consumo.
- **Grãos Bons:** São aqueles que apresentam coloração uniforme, textura íntegra e ausência de defeitos visíveis. São considerados próprios para o consumo e processamento, atendendo aos padrões de qualidade exigidos pelo mercado.
- **Grãos Esverdeados:** Apresentam uma coloração verde intensa ou parcial, causada por interrupção prematura do processo de maturação. Esse fenômeno pode ocorrer devido a condições climáticas adversas ou colheita antecipada. Grãos esverdeados podem ter impacto negativo na qualidade do óleo e do farelo de soja.
- **Grãos Mofados:** São aqueles que apresentam crescimento visível de fungos na superfície, resultado de condições inadequadas de umidade e ventilação. A presença de mofo pode indicar contaminação por micotoxinas, tornando esses grãos potencialmente prejudiciais à saúde humana e animal.
- **Grãos Picados:** São grãos que sofreram danos mecânicos severos, apresentando cortes ou quebras irregulares. Esses danos geralmente ocorrem durante a colheita, transporte ou beneficiamento e podem comprometer a qualidade do produto final.
- **Grãos Queimados:** Apresentam coloração escura ou carbonizada, decorrente de exposição excessiva ao calor em secadores ou incêndios acidentais. A aparência e as propriedades químicas desses grãos podem ser alteradas, tornando-os inadequados para algumas aplicações industriais.

3.2 SEGMENTAÇÃO E PRÉ-PROCESSAMENTO DE IMAGENS NA VISÃO COMPUTACIONAL

O pré-processamento é uma etapa essencial para garantir que os dados utilizados no treinamento do modelo de classificação sejam padronizados, limpos e otimizados para extração eficiente de características. No presente estudo, o pipeline de pré-processamento envolveu segmentação das imagens, recorte, redimensionamento, conversão para escala de cinza, normalização e data augmentation.

3.2.1 Segmentação com YOLO

A segmentação de imagens é um processo que consiste na separação de objetos de interesse dentro de uma imagem, permitindo que apenas as regiões relevantes sejam processadas pelo classificador.

O YOLO (You Only Look Once) é um dos algoritmos mais avançados para detecção e segmentação de objetos em tempo real (REDMON *et al.*, 2016). No presente estudo, foi utilizado o modelo YOLOv8n.pt, por ser uma versão mais eficiente e otimizada.

As vantagens da aplicação do YOLO incluem:

- **Detectar os grãos de soja dentro da imagem**, identificando automaticamente as áreas que continham as sementes.
- **Criar caixas delimitadoras (bounding boxes)** ao redor de cada grão, permitindo que fossem extraídos.
- **Reduzir a interferência do fundo**, removendo ruídos como sombras, reflexos e variações de cor.

Figura 1 – Comparação entre imagens de grãos de soja



(a) Imagem do grão de soja recortada



(b) Imagem de grãos com bounding box

Fonte: Elaborado pelo autor (2025)

A figura 1 ilustra a comparação entre imagem colorida sem a caixa delimitadora (bounding box) e outra imagem com a delimitação. A aplicação do YOLO garantiu que apenas os grãos segmentados fossem utilizados na etapa de classificação, eliminando informações desnecessárias e tornando o processo mais eficiente.

3.2.2 Recorte e Redimensionamento de Imagens

Após a segmentação com YOLO, os grãos identificados foram recortados individualmente variando entre cortes de 500 X 500 pixels a 900 X 850 pixels, para compor um conjunto de dados mais limpo e preciso. O recorte foi necessário pois os grãos ocupavam menos de 4.1% da área total da imagem, o que aumentaria o custo computacional e dificultaria a classificação posterior.

Para garantir uniformidade no treinamento da CNN, todas as imagens recortadas foram redimensionadas para 600 X 600 pixels. O redimensionamento é uma técnica fundamental na visão computacional, pois:

- **Garante que todas as imagens tenham o mesmo tamanho**, evitando distorções durante o processamento da rede neural.
- **Reduz a complexidade computacional**, diminuindo o tempo de treinamento sem comprometer a qualidade das informações visuais
- **Facilita o aprendizado da CNN**, pois redes neurais exigem entradas com tamanhos padronizados.

A técnica utilizada para redimensionamento foi a interpolação bilinear, que calcula novos valores de pixel, considerando a média ponderada dos quatro pixels

vizinhos mais próximos (GONZALEZ; WOODS, 2018), resultando em uma transição mais suave entre os pixels, reduzindo o efeito de serrilhado (aliasing).

Matematicamente, a interpolação bilinear pode ser representada como:

$$f(x, y) \approx (1-dx)(1-dy)f(x_1, y_1) + dx(1-dy)f(x_2, y_1) + (1-dx)dyf(x_1, y_2) + dxdyf(x_2, y_2), \quad (3.1)$$

onde:

- $f(x_1, y_1), f(x_2, y_1), f(x_1, y_2), f(x_2, y_2)$ são as intensidades dos quatro pixels vizinhos mais próximos;
- $dx = x - x_1$ e $dy = y - y_1$ representam as distâncias fracionárias entre os pixels originais e o novo ponto interpolado;
- A interpolação é realizada primeiro na direção x e depois na direção y , resultando em uma suavização mais eficiente da imagem.

3.2.3 Conversão para Escala de Cinza

A conversão para escala de cinza (grayscale) é uma técnica comum no pré-processamento de imagens para aprendizado de máquina. No presente estudo, essa conversão foi realizada para reduzir a complexidade computacional, pois imagens em cores possuem três canais de cor (RGB), enquanto imagens em tons de cinza possuem apenas um canal.

A conversão é realizada pela seguinte equação de luminância, que pondera os valores dos canais RGB de acordo com a percepção humana da luz:

$$Y = 0.299R + 0.587G + 0.114B, \quad (3.2)$$

onde:

- **R** é o valor do canal vermelho.
- **G** é o valor do canal verde.
- **B** é o valor do canal azul.
- **Y** é o valor final do pixel convertido para escala de cinza.

Esse método leva em consideração a sensibilidade do olho humano às diferentes cores. O canal verde (G) tem maior peso (0.587), pois o olho humano é mais sensível a essa cor, enquanto os canais vermelho (R) e azul (B) têm menor contribuição. Esse modelo de conversão de cores para escala de cinza é amplamente utilizado em

processamento digital de imagens devido à sua capacidade de preservar o contraste e os detalhes essenciais da imagem original (GONZALEZ; WOODS, 2018).

As vantagens incluem:

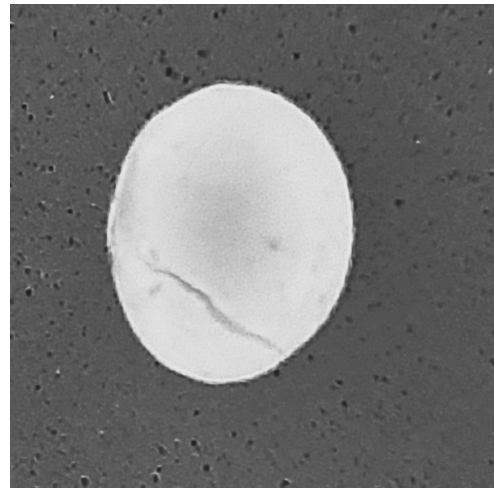
- Redução do número de parâmetros processados pela CNN.
- Eliminação de variações irrelevantes de cor
- Aceleração do treinamento

Para ilustrar esse processo, a Figura 2 apresenta um exemplo de imagem colorida convertida para escala de cinza.

Figura 2 – Comparação entre imagens colorida e em escala de cinza de grãos de soja



(a) Imagem colorida (a)



(b) Imagem em escala de cinza (b)

Fonte: Elaborado pelo autor (2025)

3.2.4 Normalização das Imagens

A normalização é um passo essencial para melhorar a estabilidade do treinamento de redes neurais, reduzindo variações nos valores de entrada e acelerando a convergência do modelo.

Normalmente, converte-se os valores de intensidade dos pixels de [0, 255] para uma faixa [0,1], especialmente quando se utilizam funções de ativação como a ReLU.

Neste estudo, foi adotada a seguinte equação:

$$X_{\text{normalizado}} = \frac{X}{255}, \quad (3.3)$$

onde:

- X representa o valor original do pixel, variando entre 0 e 255.

- $X_{\text{normalizado}}$ é o valor reescalado, agora pertencente ao intervalo $[0,1]$.

Principais benefícios da normalização:

- **Aceleração da convergência do modelo**, pois reduz variações nos valores dos pixels.
- **Evita saturação em funções de ativação**, como a ReLU, que pode ter problemas com valores muito altos.
- **Melhora a estabilidade do gradiente**, facilitando o aprendizado do modelo.

3.2.5 Aumento de Dados (Data Augmentation)

Data Augmentation é uma técnica utilizada para aumentar artificialmente o tamanho do dataset e melhorar a capacidade de generalização do modelo, prevenindo o **overfitting**.

Neste estudo, foram aplicadas as seguintes transformações:

- **Rotações aleatórias entre -15° e $+15^\circ$** : Ajudando a CNN a reconhecer grãos independentemente de sua orientação na imagem.
- **Inversões horizontais e verticais**: Simulando diferentes posições em que os grãos fossem fotografados.
- **Deslocamentos horizontais e verticais (até 10%)**: Evitando que o modelo dependa excessivamente da posição do grão na imagem.
- **Zoom aleatório entre 5% e 20%**: Permitindo que o modelo aprenda variações de tamanho nos grãos.
- **Ajuste de brilho entre -15% e +15%**: Alterando variações de iluminação, tornando o modelo mais robusto a diferentes condições.

Estas transformações ampliam a variedade de exemplos, tornando o modelo mais robusto e eficiente na classificação.

3.3 ARQUITETURA E FUNCIONAMENTO DAS REDES NEURAIS CONVOLUCIONAIS (CNN)

As redes neurais convolucionais (CNNs) são uma das abordagens mais eficazes para classificação das imagens, pois são capazes de aprender padrões visuais automaticamente, sem necessidade de extração manual de características. Esse aprendizado ocorre por meio de diferentes camadas estruturadas, incluindo **camadas convolucionais**, **camadas de pooling**, e **funções de ativação**, que trabalham em conjunto para identificar e distinguir características relevantes.

3.3.1 Camadas Convolucionais

As camadas convolucionais são a base das CNNs e responsáveis por extrair características das imagens. aplicando filtros (**kernels**) sobre regiões específicas para identificar padrões locais.

Exemplo de funcionamento:

Dada a matriz:

$$\text{Imagem} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

e o filtro 3×3 com pesos:

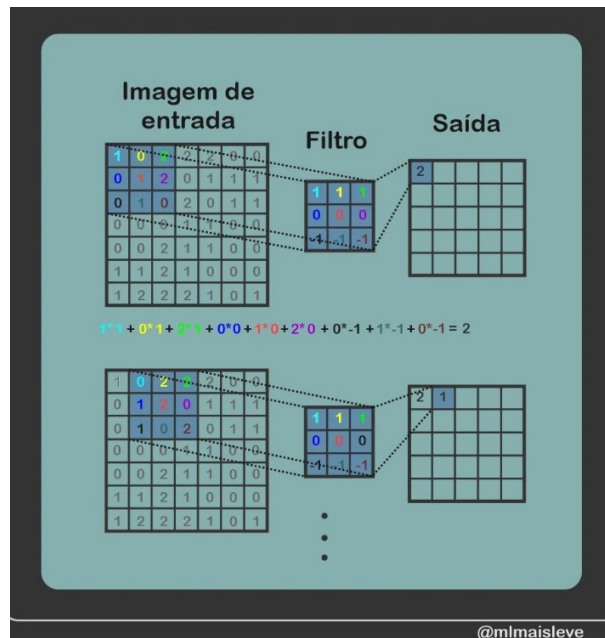
$$\text{Filtro} = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

o resultado da convolução para este recorte é:

$$\text{Resultado} = -1 - 2 - 3 - 4 + 40 - 6 - 7 - 8 - 9 = 0$$

Este cálculo é repetido para cada recorte da imagem, gerando um novo mapa de características.

Figura 3 – Exemplo simplificado do processo de convolução



Fonte: (RODRIGUES, 2021)

A figura 3, ilustra o processo realizado pelas camadas convolucionais de forma simplificada.

Tipos de Filtros:

- **Filtros de Detecção de Bordas:** Destacam mudanças abruptas na intensidade dos pixels.
- **Filtros de textura:** Identificam padrões repetitivos, como rugosidade em grãos.
- **Filtros de Detecção de Formas:** Que aprendem padrões específicos dos objetos na imagem.

A figura 4 abaixo, exemplifica alguns feature maps obtidos sobre uma imagem.

Figura 4 – Exemplo de feature maps extraídos do bloco 1 do modelo VGG16



Fonte: (BROWNLEE, 2019)

3.3.2 Camadas de Pooling

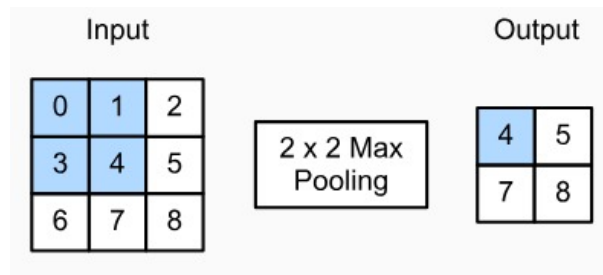
O **pooling** reduz a dimensionalidade dos feature maps, diminuindo o número de parâmetros e o risco de **overfitting**.

Tipos principais:

- **MaxPooling:** seleciona o maior valor.
- **AveragePooling:** calcula média dos valores na janela.

Esse processo permite que a CNN se concentre nas características mais relevantes.

Figura 5 – Exemplo de operação de max pooling



Fonte: (ZHANG *et al.*, 2021)

A figura 5, demonstra o funcionamento do MaxPooling. O pooling reduz o tamanho da matriz de saída enquanto preserva as informações mais relevantes. Isso ajuda a CNN a focar apenas nas características mais importantes da imagem.

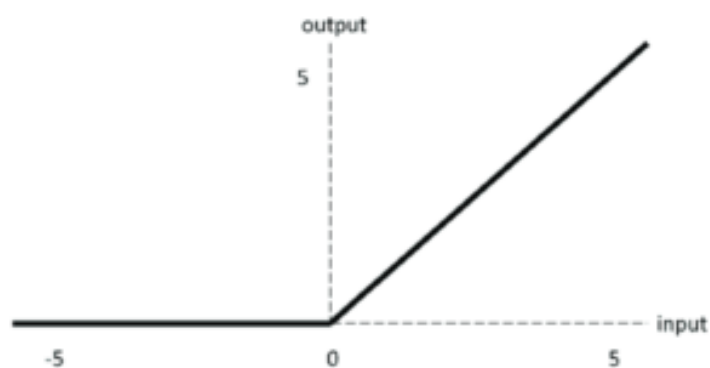
3.3.3 Funções de Ativação

As funções de ativação são aplicadas após cada camada convolucional para introduzir não-linearidade no modelo. Isso permite que a CNN aprenda padrões complexos, em vez de apenas realizar operações lineares sobre os pixels. Neste estudo, foram utilizadas duas funções, a Rectified Linear Unit (ReLU) e a Softmax.

3.3.3.1 ReLU

A função ReLU (Rectified Linear Unit) é uma das funções de ativação mais populares e recomendadas em redes neurais (GOODFELLOW; BENGIO; COURVILLE, 2016). Ela mapeia os valores negativos para zero, enquanto mantém os valores positivos inalterados. Isso ajuda a acelerar o treinamento, mitigando o problema de **gradient vanishing**, que ocorre quando os valores propagados para ajustar os pesos durante o treinamento se tornam extremamente pequenos, não causando alterações significativas. Esse problema ocorre principalmente nas camadas mais profundas da rede.

Figura 6 – Gráfico da função ReLU



Fonte: (SHUKLA, 2022)

A figura 6, ilustra seu gráfico, já sua representação matemática simplificada é dada por:

$$f(x) = \max(0, x) \quad (3.4)$$

Onde:

- Se x for maior que zero, a saída será o próprio valor de entrada (x).
- Se x for menor ou igual a zero, a saída será zero, eliminando valores negativos.

3.3.3.2 Softmax

A função Softmax é usada para transformar um vetor de valores em uma distribuição de probabilidade. Cada valor é exponenciado e normalizado para garantir que a soma das probabilidades seja igual a 1, tornando-a ideal para problemas de classificação que contenham muitas classes (multiclasse).

Matematicamente, a softmax é definida como:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (3.5)$$

Onde:

- z_i representa o valor bruto (logit) da classe i antes da normalização;
- e^{z_i} é a exponenciação do logit, ampliando diferenças entre valores;
- $\sum_{j=1}^n e^{z_j}$ é a soma de todas as exponenciações, garantindo que a soma total seja 1.

3.3.4 Camadas Densas

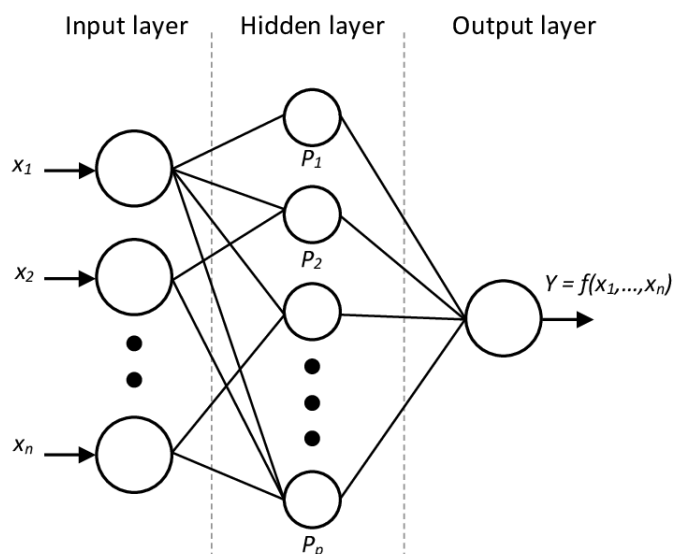
As camadas densas, também chamadas de camadas totalmente conectadas (fully connected layers), conectam todos os neurônios da camada anterior a todos os neurônios da camada seguinte. Elas são usadas para aprender representações complexas dos dados.

Matematicamente, a saída de uma camada densa é definida como:

$$Y = f(WX + b), \quad (3.6)$$

onde X é o vetor de entrada, W a matriz de pesos, b o viés e f a função de ativação, como ReLU, Sigmoid ou Softmax, dependendo da aplicação (GONZALEZ; WOODS, 2018).

Figura 7 – Diagrama de uma camada densa



Fonte: (FACURE, 2017)

A figura 7, ilustra a representação de uma camada densa (hidden layer) totalmente conectada como a camada de saída (output layer).

3.3.5 Camadas de Dropout

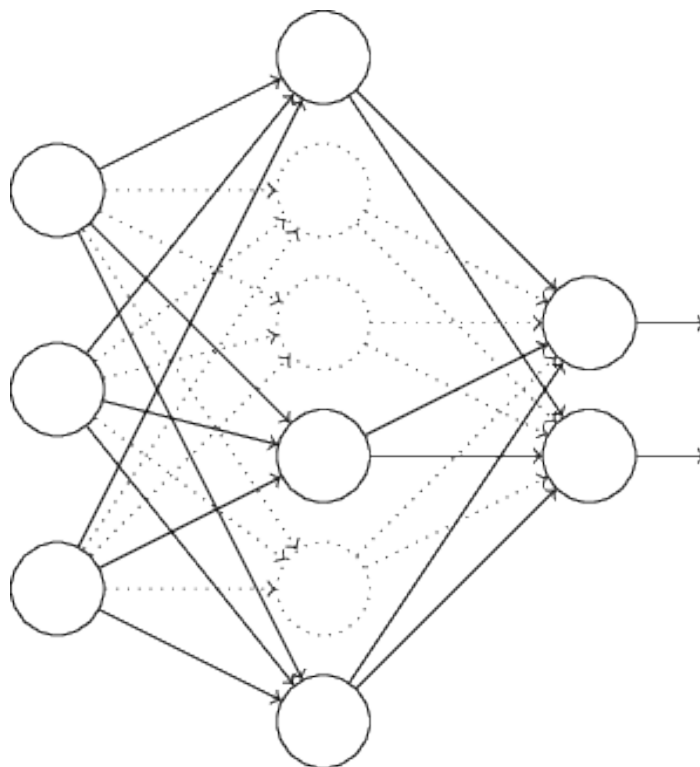
O Dropout é uma técnica de regularização que desativa aleatoriamente uma fração dos neurônios, como ilustrado na figura 8, durante o treinamento para evitar sobreajuste (overfitting), impedindo que a rede dependa excessivamente de neurônios específicos e melhorando a capacidade de generalização.

$$\text{Dropout}(x) = \begin{cases} 0, & \text{com probabilidade } p, \\ x, & \text{com probabilidade } 1 - p \end{cases}$$

Onde:

- x é a ativação da camada antes do dropout
- p é a taxa de dropout, a probabilidade do neurônio ser desativado, sendo seu valor substituído por zero.

Figura 8 – Ilustração do funcionamento do Dropout



Fonte: (DSA, 2024)

3.3.6 Fluxo Simplificado de Processamento da CNN

Com cada componente explicado, podemos visualizar o fluxo do processamento de uma imagem dentro da CNN:

- **Entrada da imagem:** Redimensionada e normalizada.
- **Camada Convolutiva:** Aplicação de filtros para extração de características.
- **Função de Ativação (ReLU):** Introdução de não-linearidade.
- **Pooling:** Redução da dimensionalidade.
- **Camadas Densas:** Conexões fully connected processam os padrões extraídos.
- **Dropout:** Redução do overfitting ao "desligar" aleatoriamente algumas unidades durante o treinamento, Podem ser utilizado tanto após camadas convolucionais quanto densas.
- **Softmax:** Converte a saída em probabilidades das classes.

3.4 ARQUITETURAS DE REDES NEURAIS PARA CLASSIFICAÇÃO DE IMAGENS

Para realizar tarefas complexas de classificação de imagens, podem ser utilizadas tanto arquiteturas **customizadas** quanto **pré-treinadas**.

3.4.1 Arquiteturas Personalizadas

Modelos customizados são desenvolvidos do zero, com diferentes configurações de camadas convolucionais, pooling e regularização. A personalização permite flexibilidade, mas exige ajuste fino dos hiperparâmetros e um volume considerável de dados para evitar overfitting.

As redes personalizadas seguem a estrutura típica de CNNs, combinando:

- múltiplas **camadas convolucionais** para extração de características,
- camadas de **pooling** para redução de dimensionalidade,
- técnicas de **regularização** como **Dropout**,
- e **camadas densas** para realizar a classificação final.

3.4.2 Arquiteturas Pré-Treinadas

Os modelos **pré-treinados** são arquiteturas desenvolvidas e treinadas em grandes bases de dados, como o ImageNet, e posteriormente ajustadas (fine-tuned) para novas tarefas específicas, utilizando a técnica de **transfer learning**.

Dentre as arquiteturas utilizadas neste estudo, destacam-se duas **ResNet50** e **InceptionV3**.

3.4.2.1 ResNet50

A ResNet50 é um modelo amplamente utilizado devido à sua capacidade de preservar o gradiente durante o treinamento, que é possível devido suas camadas residuais (**add layers**). Essas camadas criam conexões de atalho, (**skip connections**) entre camadas não consecutivas, permitindo que o gradiente seja transmitido de forma eficiente durante o processo de ajustes de pesos da rede (**backpropagation**), é essencial pois em redes muito profundas (que possuem muitas camadas), pode ocorrer o gradient vanishing (HE *et al.*, 2016).

As add layers, resolvem esse problema ao permitir que o gradiente flua diretamente para camadas anteriores, preservando informações essenciais durante o treinamento. Essa arquitetura permite que a ResNet50, possa ser muito mais profunda sem perder a capacidade de aprender, resultando em uma melhor extração de características e maior precisão, especialmente em tarefas com imagens visualmente complexas ou grandes, como o objeto deste estudo.

3.4.2.2 InceptionV3

O InceptionV3, por sua vez, adota uma abordagem única através dos blocos inception, que permitem à rede realizar convoluções com múltiplos tamanhos de

filtros simultaneamente, para capturar características visuais em diferentes escalas e resoluções dentro de um mesmo bloco. A vantagem dessa abordagem é que ela permite que o modelo seja mais eficiente em termos de tempo computacional, ao mesmo tempo que consegue representações mais robustas para diferentes tipos de texturas e padrões.

Cada bloco inception contém filtros 1x1, 3x3 e 5x5, que capturam, simultaneamente, padrões em diferentes escalas. Devido a essa flexibilidade, o InceptionV3 se torna eficaz na classificação de imagens, especialmente em casos onde imagens possuem grande variedade de características visuais ou imagens com muitos pixels. Trabalhos como o de (SZEGEDY *et al.*, 2016), detalham a fundo sobre essa estrutura inovadora.

3.5 TREINAMENTO DE MODELOS DE REDES NEURAIS

O treinamento das CNNs envolve o ajuste dos pesos dos neurônios com base no erro calculado entre as previsões do modelo e os rótulos reais. Esse processo utiliza técnicas de otimização e funções de perda, além de ser monitorado com métricas específicas.

3.5.1 Otimizador Adam

O **Adam** (Adaptive Moment Estimation) é uma técnica amplamente utilizada em redes neurais profundas, combinando as vantagens do **método do momento** e do **RMSprop** para ajuste eficiente dos pesos.

- O método do momento acelera a convergência ao considerar o histórico dos gradientes, suavizando atualizações e reduzindo oscilações.
- O RMSprop adapta a taxa de aprendizado para cada parâmetro, normalizando os gradientes com base na média exponencial dos valores passados.

O Adam incorpora ambas abordagens, utilizando estimativas adaptativas da média e variância dos gradientes, garantindo uma convergência eficiente e estável.

3.5.2 Taxa de Aprendizado (Learning Rate)

A taxa de aprendizado define o tamanho dos passos que o otimizador dá ao ajustar os pesos da rede. Um valor muito alto pode fazer o modelo não convergir, enquanto um valor muito baixo pode tornar o treinamento excessivamente lento ou resultar em um mínimo local ruim.

3.5.3 Função de Perda: Categorical Crossentropy

A função de perda Categorical Crossentropy é usada para medir o quão bem o modelo classifica múltiplas categorias. Ela calcula a diferença entre as probabilidades preditas pelo modelo e os rótulos reais, penalizando previsões incorretas de forma mais intensa, sendo ideal para problemas de classificação multiclasse

3.5.4 Épocas e Batch Size

As **épocas** representam o número de vezes que o conjunto de dados completo passa pelo modelo durante o treinamento. Mais épocas podem levar a um melhor ajuste da rede, mas um número excessivo pode resultar em overfitting. O **batch size** refere-se ao número de amostras processadas antes da atualização dos pesos do modelo. Valores menores podem melhorar a generalização, mas tornam o treinamento mais ruidoso, enquanto valores maiores exigem mais memória e podem levar a uma convergência mais suave.

3.5.5 Métricas de avaliação

Durante o treinamento, foram monitoradas duas métricas principais para avaliar a evolução do modelo:

- **Acurácia:** Mede a proporção de classificações corretas.
- **Perda:** Representa o erro do modelo em cada época.

Após o treinamento, os modelos foram avaliados pelas seguintes métricas:

- **Acurácia** (Accuracy): Calcula a proporção de previsões corretas em relação ao total de previsões feitas.

$$\text{Acurácia} = \frac{TP + TN}{TP + FP + TN + FN}$$

Onde:

- **TP** representam os verdadeiros positivos
- **TN** representam os Verdadeiros Negativos
- **FP** representam os Falsos Positivos
- **FN** representam os Falsos Negativos
- **Precisão** (Precision): Calcula a proporção de previsões positivas corretas em relação ao total de previsões feitas.

$$\text{Precisão} = \frac{\text{Verdadeiros Positivos (TP)}}{\text{Verdadeiros Positivos (TP)} + \text{Falsos Positivos (FP)}}$$

- **Revocação** (Recall): Avalia a capacidade do modelo de identificar todas as instâncias positivas, representando a proporção de positivos corretamente identificados em relação ao total de positivos reais.

$$\text{Recall} = \frac{\text{Verdadeiros Positivos (TP)}}{\text{Verdadeiros Positivos (TP)} + \text{Falsos Negativos (FN)}}$$

- **F1-Score**: Média harmônica entre precisão e recall, proporcionando uma métrica balanceada que penaliza desbalanceamentos entre essas duas métricas.

$$\text{F1-Score} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}}$$

Estas métricas permitem avaliar a capacidade do modelo de generalizar, especialmente em contextos com possíveis desequilíbrios entre classes.

4 METODOLOGIA

A metodologia deste estudo segue um pipeline estruturado em três fases principais: segmentação dos grãos com YOLO, pré-processamento das imagens e classificação com as CNNs. A abordagem adotada visa avaliar a eficácia de diferentes arquiteturas para a classificação de grãos de soja, comparando modelos customizados e modelos pré-treinados.

4.1 COLETA E ORGANIZAÇÃO DO DATASET

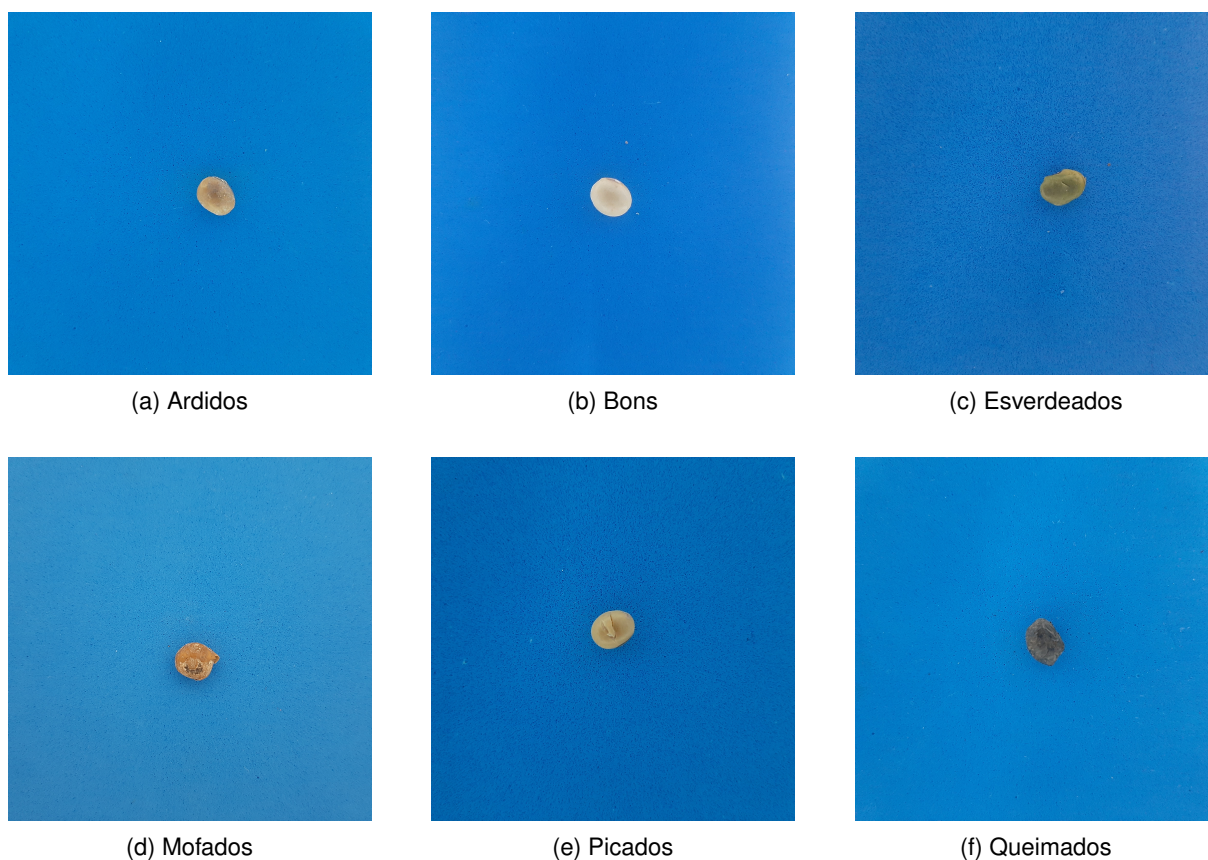
O conjunto de dados utilizado foi obtido por meio de um colaborador de pesquisa (RIBEIRO; PAULO, 2025), contendo 7.328 imagens de grãos de soja classificados manualmente em seis categorias distintas:

Tabela 1 – Distribuição das Classes de Grãos de Soja no Dataset

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

Fonte: Elaborado pelo autor (2024)

Figura 9 – Exemplos de imagens de grãos de soja por categorias



Fonte: Elaborado pelo autor (2024)

As imagens foram capturadas em condições controladas de iluminação e possuem resolução de 2.992 x 2.992 pixels.

4.2 PRÉ-PROCESSAMENTO DAS IMAGENS

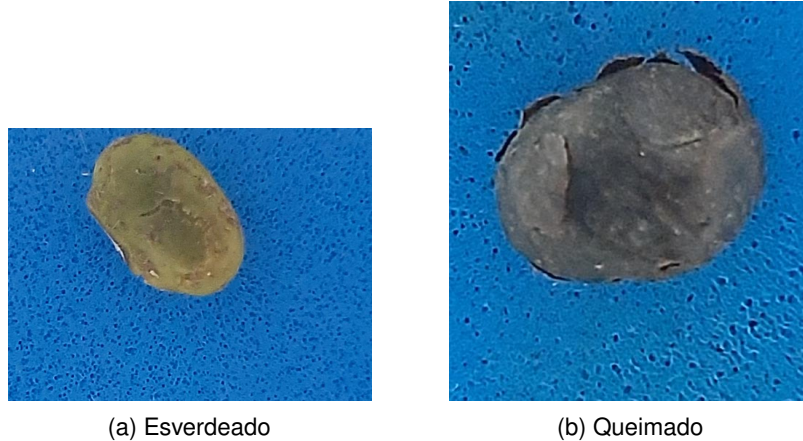
O pipeline de pré-processamento adotado neste estudo envolveu **segmentação das imagens com YOLO, recorte, redimensionamento, conversão para escala de cinza, normalização e data augmentation**.

4.2.1 Segmentação e Recorte

Após a segmentação com o modelo **YOLOv8n.pt**, os grãos identificados foram recortados individualmente, variando entre 500 x 500 pixels e 900 x 850 pixels, visando compor um conjunto de dados mais limpo e preciso.

A figura 10 abaixo, ilustra os grãos cortados.

Figura 10 – Exemplos de imagens de grãos de soja cortados



Fonte: Elaborado pelo autor (2024)

4.2.2 Conversão, Normalização e Aumento de Dados

Após o recorte e redimensionamento, as imagens foram convertidas para **escala de cinza** conforme a **equação de luminância** (ver Seção 3.2.3), reduzindo a complexidade computacional e preservando as informações essenciais.

Em seguida, todas as imagens foram **normalizadas** para a faixa $[0, 1]$ conforme descrito na Seção 3.2.4, garantindo maior estabilidade no treinamento das CNNs.

Para melhorar a generalização no modelo e mitigar o risco de **overfitting**, foi aplicada a técnica de **Data Augmentation** no conjunto de treinamento, utilizando transformações como:

- Rotações aleatórias
- Inversões
- Deslocamentos
- Zooms
- Ajustes de brilho

Essas transformações foram aplicadas exclusivamente no conjunto de treinamento, garantindo que os dados de validação e teste refletissem condições reais.

4.3 PROCESSAMENTO DAS REDES NEURAIS CONVOLUCIONAIS (CNN)

As CNNs foram utilizadas para realizar a classificação dos grãos de soja neste estudo, seguindo o fluxo descrito na Seção 3.3.

O pipeline incluiu:

- Aplicação de **camadas convolucionais** para extração de características.

- Redução de dimensionalidade com **pooling**.
- Introdução de não-linearidade com **ReLU**.
- Utilização de **camadas densas** para consolidar o aprendizado.
- Aplicação de **Dropout** para regularização.
- Uso de **Softmax** para produção das probabilidades de classificação

Esse fluxo garantiu a aprendizagem e a generalização do modelo, conforme descrito detalhadamente na **Fundamentação Teórica** (Capítulo 3)

4.4 ARQUITETURAS TESTADAS

Para avaliar o desempenho de diferentes estratégias de classificação, foram testadas três arquiteturas personalizadas e duas arquiteturas pré-treinadas.

4.4.1 Arquiteturas Personalizadas

Os modelos customizadas foram desenvolvidos do zero, com diferentes configurações de camadas convolucionais, pooling e regularização, como demonstrado na tabela 2.

Tabela 2 – Configurações dos Modelos de CNNs Personalizadas

Modelo/Camadas	Convolucionais (kernels)	Pooling (tamanho)	Dropout (%)	Densas (neurônios)
CNN Modelo 1	3 (32, 64, 128)	3 (3x3)	0	4 (64, 32, 32, 16)
CNN Modelo 2	3 (32, 64, 128)	3 (2x2)	6 (20% cada)	4 (128, 64, 32, 32)
CNN Modelo 3	3 (32, 64, 128)	3 (2x2)	7 (20% cada)	4 (128, 64, 64, 16)

Fonte: Elaborado pelo autor (2024)

4.4.2 Arquiteturas Pré-treinadas

As redes pré-treinadas ResNet50 e InceptionV3 foram utilizadas via **transfer learning**, adaptando seus pesos para a classificação de grãos.

Tabela 3 – Comparação das Arquiteturas ResNet50 e InceptionV3

Camadas/Arquitetura	ResNet50	InceptionV3
Convolucionais	53 (64-2048 kernels)	94 (32-448 kernels)
Batch Normalization	53	94
Pooling	2 (Max, Average)	4 (Max, 64-768 maps)
Densas	2 (1024, 6 neurônios)	2 (1024, 6 neurônios)
Add Layers	16	-
Concatenate	-	15 (256-2048 maps)

Fonte: Elaborado pelo autor (2024)

4.5 TREINAMENTO DOS MODELOS

O treinamento das CNNs foi realizado para garantir que os modelos fossem capazes de aprender os padrões presentes nas imagens de grãos de soja.

4.5.1 Configuração do treinamento das CNNs Personalizadas

Os modelos customizados foram treinados utilizando os seguintes parâmetros:

- **Otimizador:** Adam
- **Taxa de aprendizado:** 0,0001.
- **Função de perda:** Categorical Crossentropy.
- **Número médio de épocas:** 120.
- **Batch size:** 16 imagens por interação.

4.5.2 Configuração do treinamento das CNNs Pré-treinadas

Os modelos pré-treinados (ResNet50 e Inceptionv3) foram ajustados por meio de ajustes finos (fine-tuning), utilizando um conjunto de hiperparâmetros ligeiramente diferentes:

- **Otimizador:** Adam
- **Taxa de aprendizado:** 0,0001.
- **Função de perda:** Categorical Crossentropy.
- **Número médio de épocas:** 15.

- **Batch size:** 32 imagens por interação.

Diferente das CNNs customizadas, que exigiriam 120 épocas para um treinamento completo, as redes pré-treinadas já possuíam pesos ajustados com base em grandes bases de dados (ImageNet). Assim, o treinamento foi realizado com um número menor de épocas (entre 10 e 20), focando no ajuste dos pesos para o domínio específico dos grãos de soja.

4.5.3 Monitoramento durante o treinamento

Durante o treinamento, foram monitoradas duas métricas principais para avaliar a evolução do modelo:

- **Acurácia:** Mede a proporção de classificações corretas.
- **Perda:** Representa o erro do modelo em cada época.

Essas métricas foram calculadas tanto no conjunto de treinamento quanto no conjunto de validação, permitindo identificar possíveis problemas como overfitting.

Após o término do treinamento, os modelos foram avaliados em um conjunto de teste separado, utilizando as métricas de acurácia, precisão, recall e f1-Score.

5 RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados obtidos a partir do treinamento e avaliação das CNNs para classificação de grãos de soja, destacando a eficácia das arquiteturas personalizadas e pré-treinadas. Além disso, são analisadas métricas como acurácia, precisão, recall e F1-Score, bem como o impacto da segmentação com YOLO na qualidade das predições.

5.1 COMPARAÇÃO DO DESEMPENHO DAS ARQUITETURAS

Os testes de validação foram realizados com métricas padronizadas de precisão, recall e F1-score.

Tabela 4 – Comparação das Arquiteturas de CNNs Testadas

Modelo	Acurácia (%)	Precision	Recall	F1-Score	Treinamento (min/época)
Modelo 1	6.1%	0.01	0.06	0.01	3.3
Modelo 2	16.4%	0.02	0.13	0.03	3.6
Modelo 3	16.6%	0.02	0.14	0.03	3.9
ResNet50	63.1%	0.57	0.63	0.59	10.2
InceptionV3	95.9%	0.96	0.96	0.96	9.9

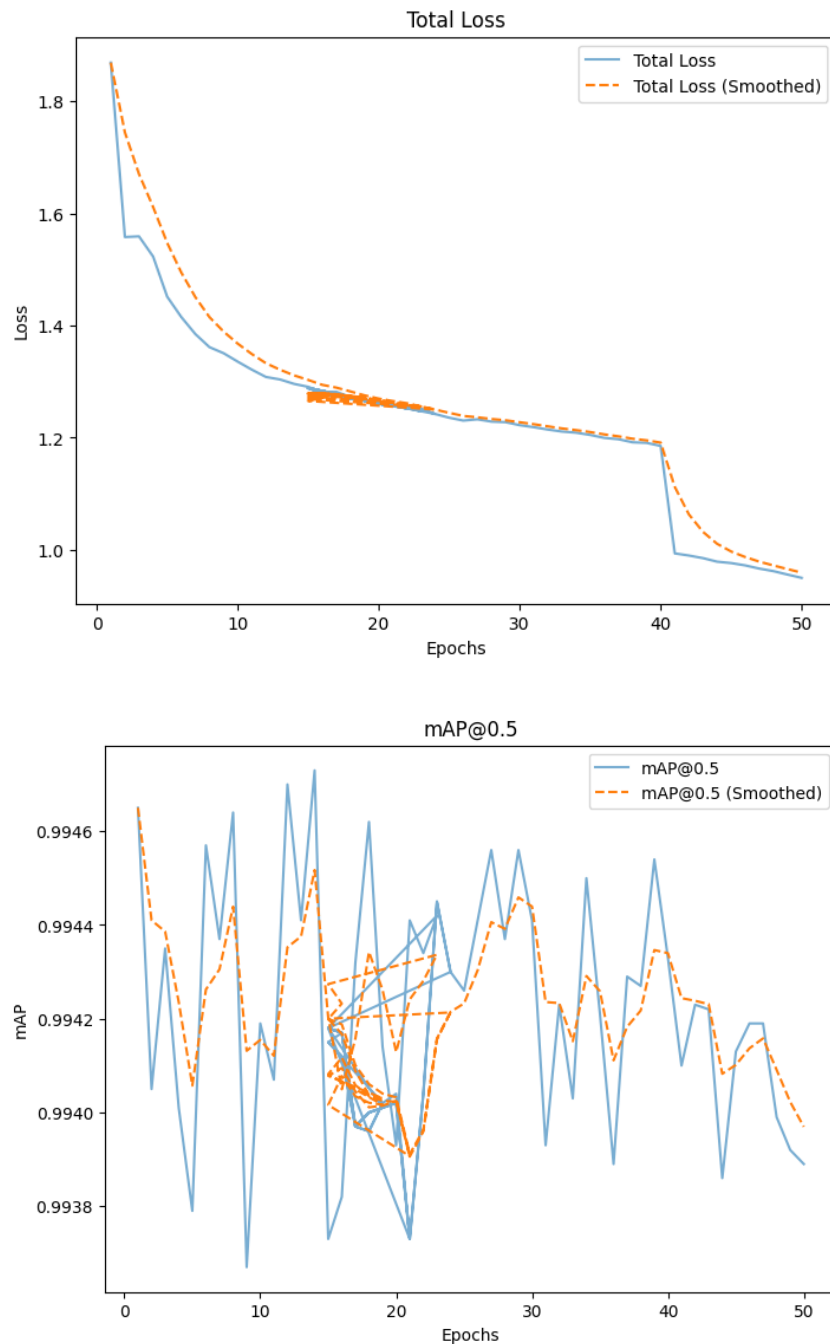
Fonte: Elaborado pelo autor (2024)

A tabela 4 ilustra de forma resumida os resultados obtidos pelos modelos personalizados e pré-treinados, os valores apresentados como métricas são referentes ao seu **weighted average**, que reflete o impacto de cada classe no conjunto de dados com base na quantidade de instâncias, evitando que classes com instâncias menores tenham impacto desproporcional na média total, fornecendo uma visão mais realista do desempenho geral do modelo. Observa-se de que os modelos pré-treinados (ResNet50 e InceptionV3) superaram amplamente as arquiteturas personalizadas.

5.2 IMPLICAÇÕES E UTILIZAÇÕES PRÁTICAS

Para garantir a qualidade dos dados enviados aos modelos classificatórios, o YOLO foi utilizado para identificar e recortar as imagens. O modelo alcançou métricas elevadas, como o mean Average Precision (**mAP**), que combina a precisão e o recall para medir o desempenho de detecção de modelos de objetos, alcançou de 0.995 e recall de 1.0, assegurando que apenas os grãos de soja fossem analisados nas etapas seguintes.

Figura 11 – Métricas do modelo YOLO: (a) Total Loss, (b) mAP0.5



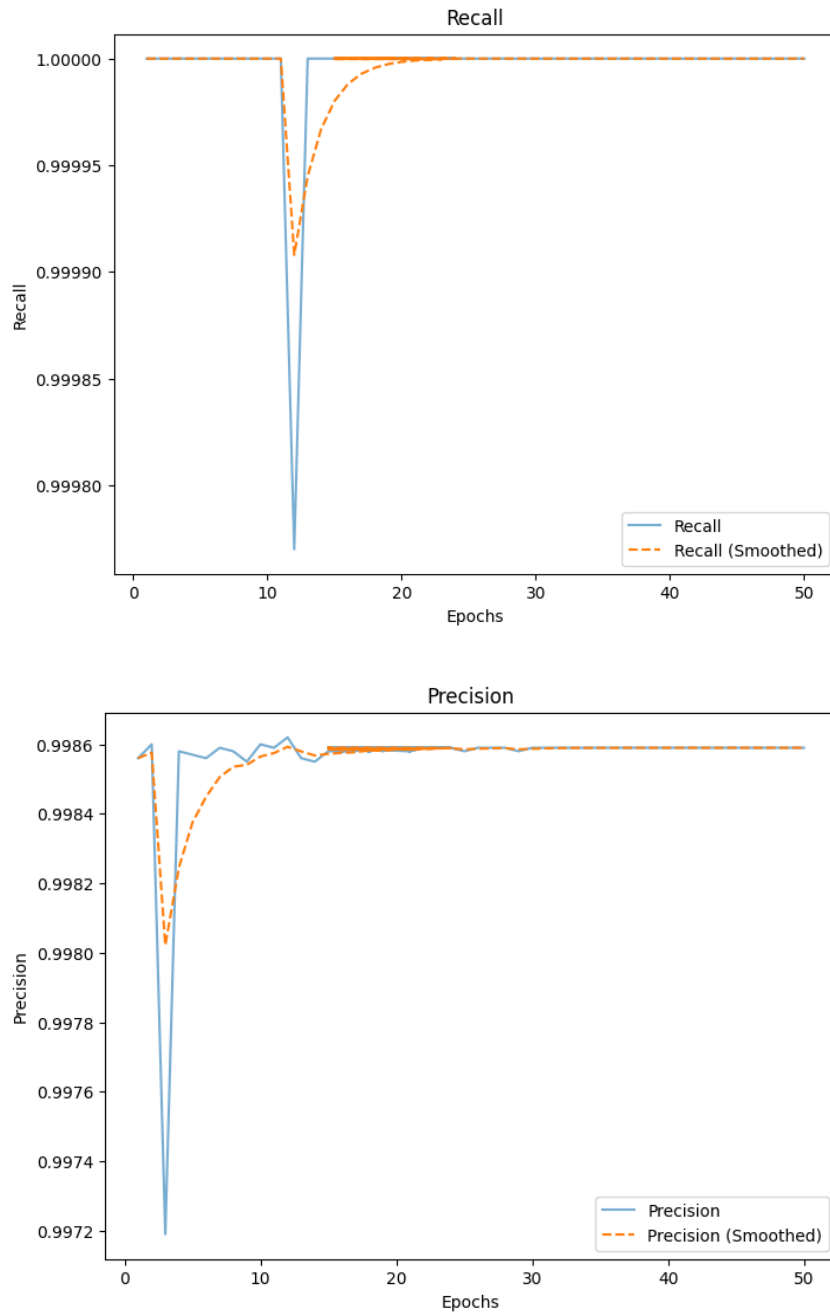
Fonte: Elaborado pelo autor (2024)

A figura 11, apresenta a perda total **(a)** e o mAP **(b)** do modelo ao longo das épocas, a perda decresce de forma geral, indicando que o modelo está aprendendo eficientemente. O gráfico de mAP, ilustra uma convergência gradativa ao longo das épocas com valores próximos de 1, indicando um alto desempenho na detecção de objetos.

Ambos os gráficos apresentam variações que podem estar relacionadas a ajustes no aprendizado, mas a perda suavizada (smoothed), que reduz a sensibilidade

a ruídos e **outliers** (valores atípicos que podem distorcer o aprendizado do modelo), demonstra uma tendência clara de redução da perda, maior estabilidade e alta precisão na detecção de objetos.

Figura 12 – Métricas do modelo YOLO: (a) Recall, (b) Precision

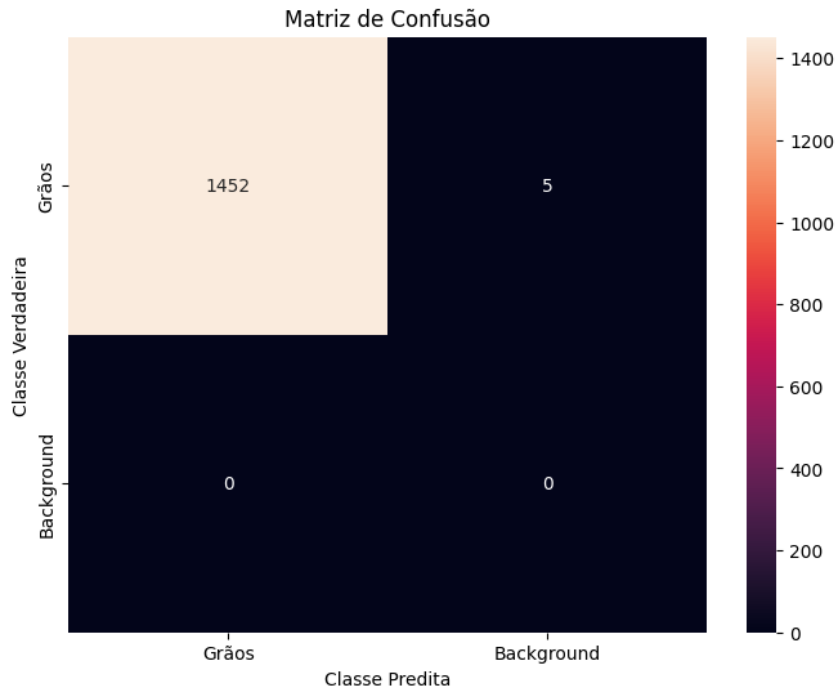


Fonte: Elaborado pelo autor (2024)

A figura 12, ilustra a recall (a) do modelo YOLO ao longo das épocas, os altos valores, indicados por serem próximos de 1, indicam que o modelo está detectando corretamente a maioria dos objetos presentes. Em (b), exibe a precisão, demonstrando uma tendência de alta estabilidade, com valores próximos de 1, o que indica que o modelo informa poucas vezes, falsos positivos.

Assim como nas outras métricas, as flutuações menores podem ser atribuídas a ajustes contínuos no processo de aprendizado, porém a precisão suavizada demonstra uma convergência consistente.

Figura 13 – Matriz de confusão do modelo YOLO



Fonte: Elaborado pelo autor (2024)

A figura 13 apresenta a matriz de confusão do modelo YOLO, informando a quantidade de previsões corretas e incorretas feitas pelo modelo, comparando as classes reais com as classes preditas. Observa-se um bom desempenho na detecção de grãos, com uma alta taxa de acerto e um baixo número de falsos positivos, indicando a eficiência em distinguir grãos de background, mesmo que ainda haja um espaço para ajustes para reduzir os erros mínimos observados.

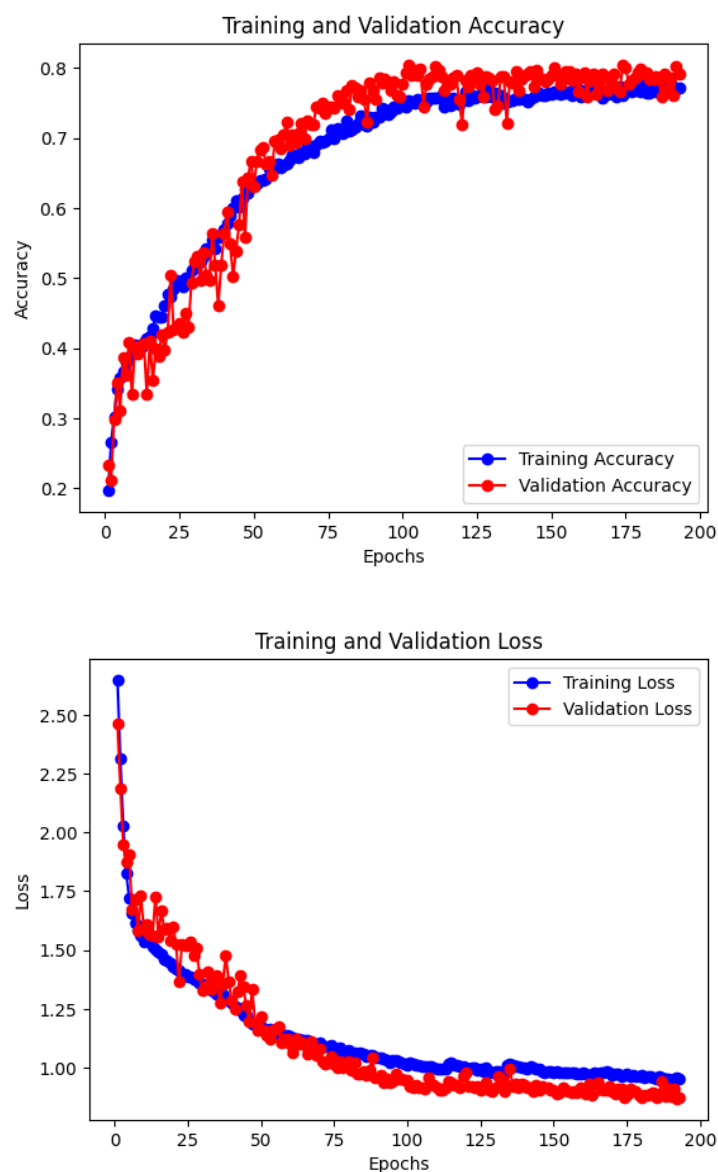
Com os dados segmentados previamente pelo YOLO, as redes classificatórias puderam focar na extração de características relevantes, maximizando a eficiência do pipeline como um todo. A aplicação das redes treinadas, na classificação dos grãos apresenta diversas implicações práticas, assim como várias possibilidades de utilização em ambientes agrícolas e industriais. As redes neurais, sobretudo as pré-treinadas, mostraram um grande potencial na classificação de grãos de soja. A escolha entre as redes customizadas e pré-treinadas deve ser baseada nos objetivos específicos de desempenho, custo computacional e aplicabilidade prática.

As CNNs customizadas, apresentaram resultados insatisfatórios em termos de acurácia, sendo muito inferiores às redes pré-treinadas. Porém, sua vantagem é o controle total sobre a arquitetura, que permite ajustes mais finos para atender a requisitos específicos de tarefas. Além disso, são mais eficientes em termos de uso

de memória e tempo de treinamento, o que pode ser útil em dispositivos de hardware mais limitados.

No entanto, a acurácia atual dos modelos é muito baixa, sendo a mais alta com 16.6%, isso implica que elas podem ser mais adequadas para tarefas preliminares ou em cenários onde a precisão não é o critério principal como sistemas para pré-triagem de grãos, que observam defeitos mais visíveis e não necessita de análises mais detalhadas. Elas também podem servir como base para implementações mais simples em dispositivos embarcados, onde o foco é mais na eficiência do que na precisão absoluta.

Figura 14 – Gráficos de treinamento do melhor modelo personalizado

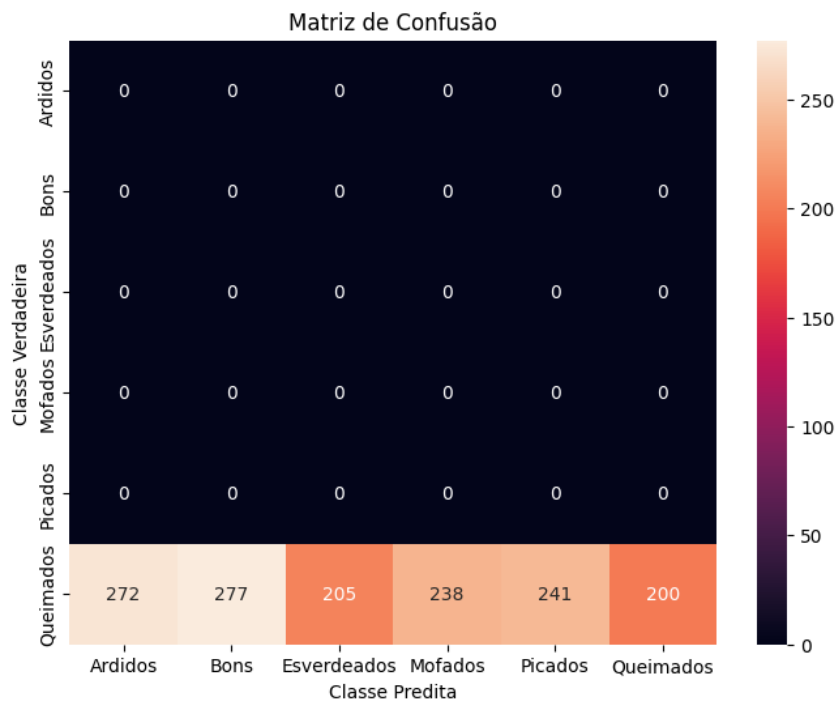


Fonte: Elaborado pelo autor (2024)

A figura 14 demonstra o desempenho da rede neural v6 durante o treinamento e validação do dataset. Apesar da perda tanto de treinamento quanto de validação

diminuírem ao longo das épocas, a baixa acurácia geral, indica que o modelo está tendo dificuldades em generalizar para novos dados. Essa baixa performance pode ser atribuída a diversos fatores, como a complexidade do problema, a qualidade dos dados, a arquitetura da rede ou a escolha dos hiperparâmetros. É importante investigar as possíveis causas dessa baixa acurácia e realizar ajustes no modelo para melhorar seu desempenho.

Figura 15 – Matriz de confusão do melhor modelo customizado



Fonte: Elaborado pelo autor (2024)

A figura 15 apresenta a matriz de confusão do modelo personalizado, informando a quantidade de previsões corretas e incorretas feitas pelo modelo. Observa-se um desempenho ruim na classificação de grãos, com uma nenhuma taxa de acerto para os tipos de grãos "ardidos", "bons", "esverdeados", "mofados" e "picados" classificando todos os grãos como sendo da classe "queimados", o que resultou em um acerto de 100% para essa classe em específico.

A tabela 5 apresenta o classification report do modelo personalizado, com métricas que detalham seu desempenho em cada classe. O relatório mostra que o modelo não conseguiu classificar corretamente as classes "ardidos", "bons", "esverdeados", "mofados" e "picados", com precisão, recall e F1-Score zerados. Para a classe "queimados", o modelo atingiu um recall de 1.00, identificando todos os exemplos reais, mas com uma precisão de apenas 0.14, indicando também que a maioria das previsões para essa classe são falsos positivos. O F1-Score de 0.24 para "queimados" reforça esse desequilíbrio.

Tabela 5 – Métricas de desempenho por classe do melhor modelo customizado

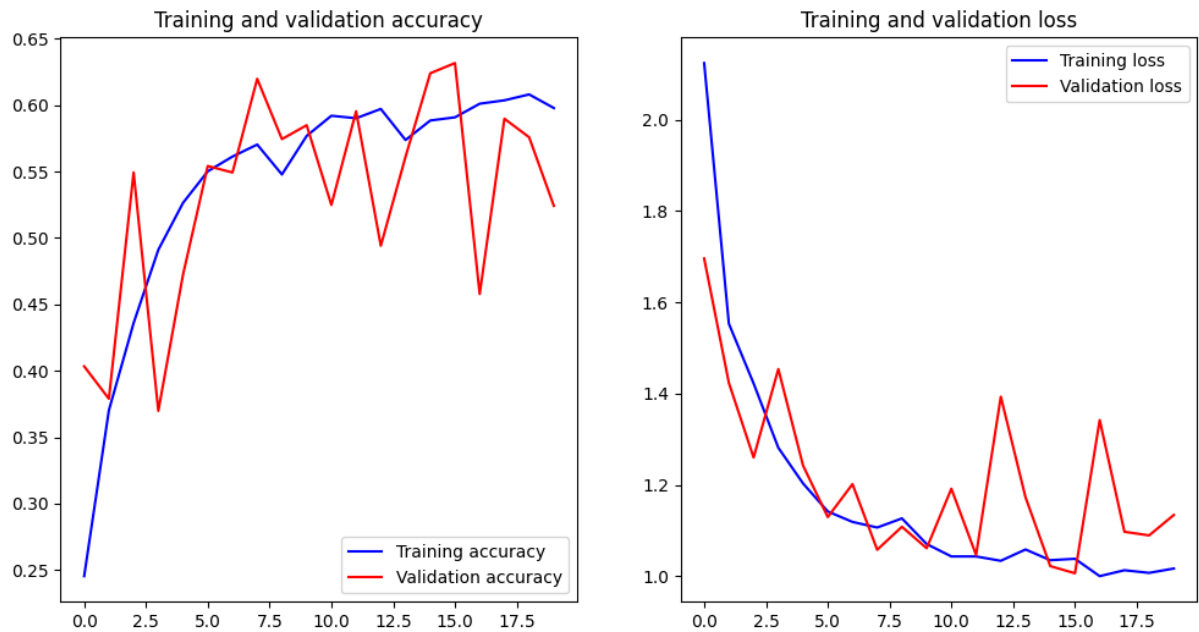
Classe	Precision	Recall	F1-Score	Support
Ardidos	0.00	0.00	0.00	272
Bons	0.00	0.00	0.00	277
Esverdeados	0.00	0.00	0.00	205
Mofados	0.00	0.00	0.00	238
Picados	0.00	0.00	0.00	241
Queimados	0.14	1.00	0.24	200
Accuracy	-	-	0.16	1433
Macro Avg	0.02	0.17	0.04	1433
Weighted Avg	0.02	0.14	0.03	1433

Fonte: Elaborado pelo autor (2024)

A acurácia geral do modelo é de 0.16, mostrando que apenas 16% das previsões estão corretas. As médias das métricas (Macro Avg e Weighted Avg) são próximas de zero, confirmando que o modelo não está performando bem na maioria das classes. Esses resultados reforçam a necessidade de ajustes no modelo para melhorar a classificação, especialmente para as classes negligenciadas.

O modelo pré-treinado ResNet50, destacou-se por sua capacidade de generalização graças às suas camadas residuais alcançando uma acurácia de 63.1%, se mostrando eficiente na extração de características visuais complexas. Seu uso é mais apropriado para situações em que se busca uma relação balanceada entre acurácia e custo computacional, especialmente com recursos computacionais suficientes como em workstations com GPUs poderosas em linhas de produção automatizadas ou servidores de processamento remoto onde grandes volumes de dados precisam ser processados rapidamente.

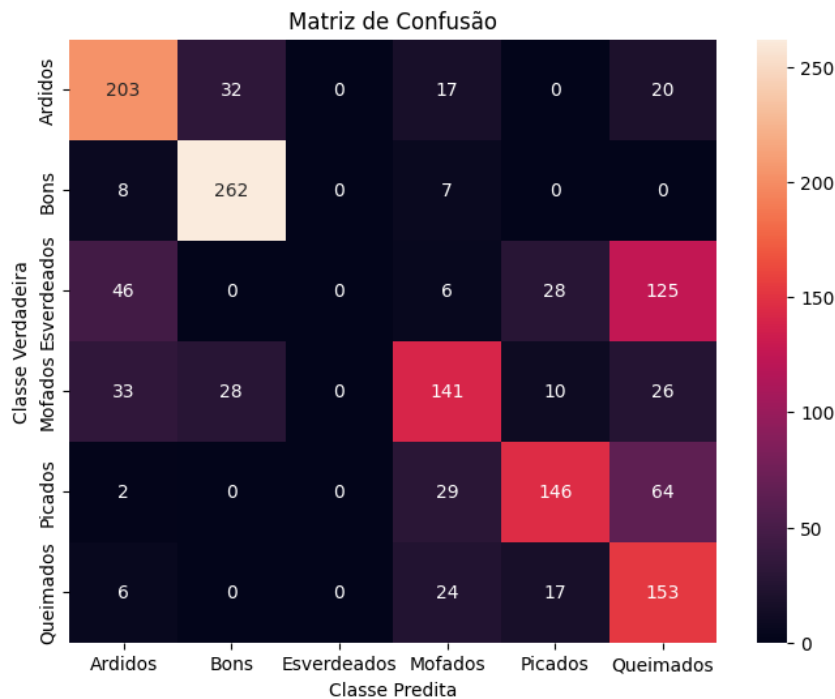
Figura 16 – Gráficos de acurácia e perda (loss) do modelo ResNet50 durante o treinamento e validação



Fonte: Elaborado pelo autor (2024)

A figura 16 apresenta o desempenho do modelo ResNet50. O gráfico de acurácia mostra uma convergência gradativa entre o treinamento e a validação, embora com variações mais acentuadas. A perda (loss) decresce de forma geral, mas exibe flutuações que podem estar relacionadas a pequenos ajustes no aprendizado do modelo.

Figura 17 – Matriz de confusão do modelo ResNet50



Fonte: Elaborado pelo autor (2024)

A figura 17 apresenta a matriz de confusão do modelo pré-treinado ResNet. Nota-se um desempenho satisfatório para a maioria das classes presentes, porém, a classe "esverdeados" não teve nenhuma previsão correta, sendo erroneamente classificada "ardidos", "mofados", "picados" e "queimados". Indicando que o modelo necessita de ajustes para melhorar a distinção de classes.

Tabela 6 – 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 (2024)

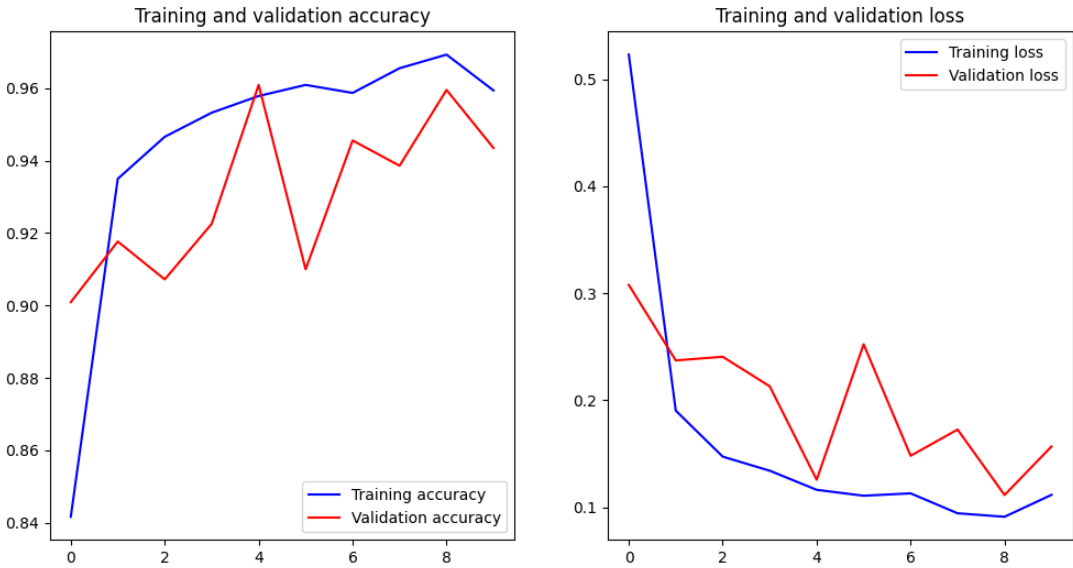
A tabela 6 apresenta o classification report do modelo ResNet50, detalhando seu desempenho por classe. O modelo teve bons resultados para "Bons"(F1-Score de 0.87) e "Ardidos"(F1-Score de 0.71), mas enfrentou dificuldades com "Esverdeados", que teve todas as métricas zeradas, indicando falha completa na classificação. Para "Queimados", o recall foi alto (0.77), mas a precisão baixa (0.39) revela muitos falsos positivos.

A acurácia geral foi de 0.63, com médias de F1-Score em 0.56 (Macro Avg) e 0.59 (Weighted Avg), mostrando que o desempenho foi impactado pelas classes problemáticas, fortificando a necessidade de ajustes para aprimorar a distinção entre as classes.

Já o modelo InceptionV3 demonstrou ser a melhor opção com acurácia de 95.9%. Sua capacidade de capturar variações sutis em diferentes escalas de imagem faz com que essa rede seja uma excelente escolha para classificação de alta precisão, ideal para produção em escala onde classificações mais precisas e rápidas dos grãos é um fator crucial.

Devido à sua sensibilidade na captura de padrões, faz dele a melhor escolha para produção industrial, como fábricas de processamento de grão onde o controle de qualidade deve ser rigoroso. Além disso, sua eficiência no uso de recursos computacionais permite que seja usado tanto em infraestruturas em nuvem quanto em dispositivos móveis mais avançados, como drones de monitoramento que podem classificar o grão em tempo real durante operações agrícolas.

Figura 18 – Gráficos de acurácia e perda (loss) do modelo InceptionV3 durante o treinamento e validação

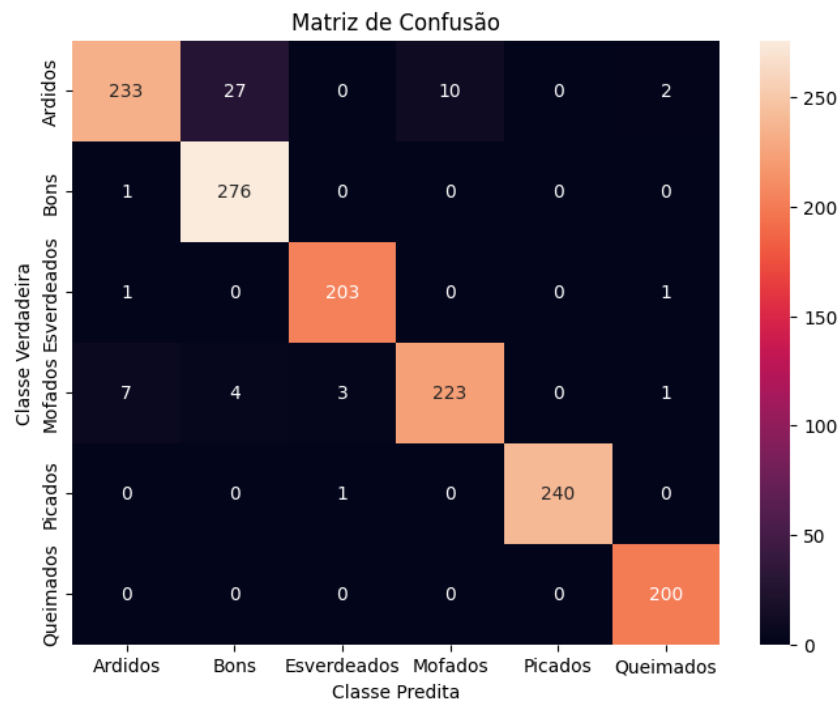


Fonte: Elaborado pelo autor (2024)

A figura 18 ilustra o desempenho do modelo InceptionV3 durante o treinamento e validação. Observa-se que a acurácia de validação se aproxima da acurácia de

treinamento ao longo das épocas, indicando boa generalização do modelo. A redução consistente na perda (loss) demonstra a eficiência do modelo em aprender as características dos dados.

Figura 19 – Matriz de confusão do modelo InceptionV3



Fonte: Elaborado pelo autor (2024)

A figura 19 apresenta a matriz de confusão do modelo pré-treinado Inception. Nota-se um ótimo desempenho para todas as classes presentes. Há poucos erros de classificação, sobretudo nas classes "ardidos" e "mofados", revelando que são ainda necessários pequenos ajustes para uma classificação mais precisa.

Tabela 7 – 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 (2024)

A tabela 7 apresenta o classification report do modelo, mostrando um desempenho excelente em todas as classes. Para "Bons", "Esverdeados", "Picados" e "Queimados", o modelo atingiu F1-Scores próximos ou iguais a 1.00, indicando alta precisão e recall. As classes "Ardidos" e "Mofados" também tiveram ótimos resultados, com F1-Scores de 0.91 e 0.95, respectivamente.

A acurácia geral foi de 0.96, com médias de F1-Score em 0.96 (Macro Avg e Weighted Avg), demonstrando um desempenho consistente e equilibrado em todas as classes, indicando que o modelo está classificando as categorias de forma eficaz, sem necessidade de ajustes significativos.

5.3 LIMITAÇÕES

Apesar do YOLO ter contribuído para a detecção de objetos nas imagens, sua aplicação não eliminou os desafios associados ao treinamento das redes classificatórias. Entre os problemas observados estão o overfitting e a baixa acurácia em arquiteturas customizadas, que carecem de camadas mais sofisticadas e técnicas avançadas de regularização. Esses fatores limitaram o desempenho de classificação, restringindo sua aplicabilidade a tarefas de menor complexidade e em cenários onde recursos computacionais são limitados.

Já na ResNet50, embora seja eficiente, exige maior capacidade computacional e tempo de treinamento, o que pode ser um impedimento para sua utilização em dispositivos embarcados ou ambientes com recursos limitados que dependem de processamento local. Apesar da acurácia significativamente maior se comparada às redes customizadas, essa rede não alcançou a acurácia desejada para tarefas de alta precisão como a classificação de grãos, indicando que mais ajustes e mais treinamento serão necessários para otimizar sua aplicação nesse cenário específico.

O modelo InceptionV3 apresentou ótimos resultados em termos de acurácia, mas sua complexidade o torna menos indicado para implementações em hardwares mais simples ou dispositivos de baixo consumo energético. O tempo de treinamento e a necessidade de grande capacidade de processamento também são fatores limitantes.

6 CONCLUSÃO

Este estudo explorou a aplicação de redes neurais convolucionais na classificação de grãos de soja, comparando três modelos personalizados de CNNs com as arquiteturas pré-treinadas, ResNet50 e InceptionV3. Os resultados mostram uma superioridade das arquiteturas pré-treinadas, particularmente a InceptionV3 em termos de acurácia. As redes CNN personalizadas, demonstraram muitas limitações, mas servem para incrementos futuros devido à sua alta flexibilidade de refinamento.

Além disso, o YOLO desempenhou um papel fundamental na segmentação das imagens, garantindo que apenas os grãos fossem analisados nas etapas classificatórias. Essa etapa de pré-processamento contribuiu para a organização dos dados e a eliminação de ruídos estabelecendo um pipeline eficiente para a classificação. Os resultados obtidos servem como base para o desenvolvimento de um sistema automatizado de classificação que visa otimizar os processos de classificação de grãos em tempo real na indústria da soja, garantindo maior previsão e eficiência.

Diante desses resultados, torna-se evidente que, embora o presente estudo tenha alcançado avanços importantes na classificação automatizada de grãos de soja, ainda há espaço para aprimoramentos e novas aplicações. Assim, na próxima seção, são apresentadas sugestões para trabalhos futuros que visam ampliar a aplicabilidade prática desse sistema e superar algumas limitações identificadas.

7 TRABALHOS FUTUROS

Embora os resultados deste estudo tenham demonstrado a viabilidade da classificação automatizada de grãos de soja utilizando redes neurais convolucionais (CNNs) e segmentação com YOLO, diversas possibilidades de aprimoramento podem ser exploradas.

A principal limitação identificada foi a utilização de imagens com grãos previamente segmentados individualmente. Assim, um dos principais avanços futuros consiste na expansão do dataset, incorporando imagens com múltiplos grãos simultaneamente, possibilitando que o YOLO realize a detecção e extração de cada grão de forma automática.

Essa expansão poderá ser realizada por meio da captura de imagens em diferentes configurações de iluminação, ângulos e fundos, aumentando a robustez do sistema. Além disso, técnicas avançadas de anotação e o uso de data augmentation poderão ser aplicadas para melhorar ainda mais a qualidade do treinamento, possibilitando que o modelo reconheça grãos em posições variadas ou sobrepostos.

Outro direcionamento importante consiste na implementação de um software de classificação em tempo real, integrando o modelo YOLO para segmentação e a CNN para classificação. Esse sistema poderia ser desenvolvido com frameworks como FastAPI ou Flask para o backend, e uma interface em React ou Vue.js para o monitoramento em tempo real, oferecendo praticidade e eficiência para ambientes industriais.

Para viabilizar a aplicação prática, o modelo deverá ser otimizado para execução em hardware especializado, como GPUs embarcadas (por exemplo, NVIDIA Jetson) ou FPGAs, garantindo baixa latência e alto desempenho. A integração com sensores IoT também poderá ser explorada, permitindo o monitoramento contínuo da qualidade dos grãos em processos automatizados.

Assim, os principais desafios para o futuro incluem:

- Construção de um dataset mais representativo com múltiplos grãos.
- Aprimoramento da segmentação e da generalização do modelo.
- Desenvolvimento de um sistema de inspeção automatizado e eficiente.

Esses avanços contribuirão significativamente para a modernização da agroindústria, promovendo processos de classificações mais rápidos, precisos e padronizados, com impacto direto na qualidade dos grãos comercializados e na redução de custos operacionais.

REFERÊNCIAS

ALMEIDA, G. P. S. de *et al.* Performance analysis of yolo and detectron2 models for detecting corn and soybean pests employing customized dataset. **Agronomy**, v. 14, n. 10, 2024. ISSN 2073-4395. DOI: <https://doi.org/10.3390/agronomy14102194>.

BEVERS, N.; SIKORA, E. J.; HARDY, N. B. Soybean disease identification using original field images and transfer learning with convolutional neural networks. **Computers and Electronics in Agriculture**, v. 203, p. 107449, 2022. ISSN 0168-1699. DOI: <https://doi.org/10.1016/j.compag.2022.107449>.

BOSCHIERO, B. N. **Subprodutos da soja: conheça os destinos e usos da soja brasileira**. 2023. Disponível em: <<https://agroadvance.com.br/blog-subprodutos-da-soja/>> Acesso em: 15 abr. 2025.

BRASIL. **Referencial Fotográfico dos Defeitos da Soja – 3ª Edição**. Brasília, 2008. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-vegetal/qualidade-vegetal-1/referencial-fotografico-pasta/referencial-fotografico-da-soja.pdf>>. Acesso em: 10 fev. 2025.

BRASIL. **Referencial Fotográfico para Grãos**. Brasília, 2023. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-vegetal/qualidade-vegetal-1/referencial-fotografico-pasta>>. Acesso em: 10 fev. 2025.

BROWNLEE, J. **How to visualize filters and feature maps in convolutional neural networks**. [S.l.]: Machine Learning Mastery, 2019. Disponível em: <<https://machinelearningmastery.com/how-to-visualize-filters-and-feature-maps-in-convolutional-neural-networks/>> Acesso em: 14 out. 2024.

CARBONI, D. F. **Sistema para identificação de plantas invasoras em lavouras com detectores de objetos aplicados a imagens e vídeos**. 2021. Disponível em: <<https://lume.ufrgs.br/handle/10183/222610>>. Acesso em: 05 jan. 2025.

DSA. **Como funciona o dropout**. [S.l.]: Deep Learning Book, 2024. Disponível em: [urlhttps://www.deeplearningbook.com.br/como-funciona-o-dropout/](https://www.deeplearningbook.com.br/como-funciona-o-dropout/). Acesso em: 20 set. 2024.

EMBRAPA. **Classificação e Padronização de Produtos, com ênfase na Agropecuária: Uma Análise Histórico-Conceitual**. [S.l.], 2001. Disponível em: <<https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/105998/1/seatexto10.pdf>>. Acesso em: 15 fev. 2025.

FACURE, M. **Introdução às Redes Neurais Artificiais**. 2017. Disponível em: <<https://matheusfacure.github.io/2017/03/05/ann-intro/>> Acesso em: 27 set. 2025.

FRANÇA-NETO, J. B.; HENNING, A. A.; KRYZANOWSKI, F. C. **Defeitos da Classificação Comercial da Soja Devido a Danos Mecânicos e Fungos**. Londrina, PR, 2011. Disponível em: <<https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/570849/1/tecno01.pdf>>. Acesso em: 10 fev. 2025.

GOMES, C. *et al.* Utilização de transfer learning no monitoramento de temperatura de transformadores de aerogeradores. In: BRAZIL WIND POWER. **Brazil Wind Power – Categoria 3: PAPERS-BWP-CATEGORIA3**. [S.l.], 2023. Disponível em: <<https://www.brazilwindpower.com.br/content/dam/markets/braz/brazil-windpower/pt/2024/pdfs/PAPERS-BWP-CATEGORIA3.pdf>>. Acesso em: 20 set. 2024.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. New York: Pearson, 2018.

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

HE, K. *et al.* Deep residual learning for image recognition. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2016. p. 770–778. DOI: <https://doi.org/10.1109/CVPR.2016.90>.

KAMILARIS, A.; BOLDÚ, F. X. P. Deep learning in agriculture: A survey. **Computers and Electronics in Agriculture**, v. 147, p. 70–90, 2018. ISSN 0168-1699. DOI: <https://doi.org/10.1016/j.compag.2018.02.016>.

LECUN, Y. *et al.* Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, IEEE, v. 86, n. 11, p. 2278–2324, 1998. DOI: <https://doi.org/10.1109/5.726791>.

MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. **Instrução normativa n.º 11, de 15 de maio de 2007**. 2007. Disponível em: <<https://sislegis.action/detalhaAto.do?method=visualizarAtoPortalMapa&chave=1194426968>> Acesso em: 20 abr. 2024.

PETERSEN, T. M. **Para que serve a soja? Veja 5 usos além do óleo e do farelo**. 2022. Disponível em: <<https://globo rural.globo.com/Noticias/Agricultura/Soja/noticia/2022/09/para-que-serve-soja-veja-5-usos-alem-do-oleo-e-do-farelo.html>>. Acesso em: 14 abr. 2025.

REDMON, J. *et al.* You only look once: Unified, real-time object detection. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2016. p. 779–788. DOI: <https://doi.org/10.1109/CVPR.2016.91>.

RIBEIRO, A. F. *et al.* **Testando a rede neural YOLOv5 para detecção de frutos em pomares de laranja**. Campinas: Instituto Agronômico, 2022. 1–11 p. Disponível em: <<http://www.alice.cnptia.embrapa.br/alice/handle/doc/1146123>>. Acesso em: 12 jan. 2025.

RIBEIRO, F. R. João Pedro de M. V.; PAULO, O. **soybeans-not-noted-cut**. [S.l.]: Mendeley Data, 2025. DOI: <https://doi.org/10.17632/xnx9bfbh6b>.

RODRIGUES, V. B. **Conhecendo a visão do computador: redes neurais convolucionais**. [S.l.]: Medium, 2021. Disponível em: <<https://vitorborbarodrigues.medium.com/conhecendo-a-visão-do-computador-redes-neurais-convolucionais-e1c2b14bf426>> Acesso em: 14 out. 2024.

SHUKLA, P. **ReLU Activation Function and Its Variants**. [S.l.]: PythonKitchen, 2022. Disponível em: <<https://www.pythonkitchen.com/relu-activation-function-and-its-variants/>> Acesso em: 20 set. 2024.

SZEGEDY, C. *et al.* Rethinking the inception architecture for computer vision. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2016. p. 2818–2826. DOI: <https://doi.org/10.1109/CVPR.2016.308>.

YANG, X. *et al.* Classification of soybean seeds based on rgb reconstruction of hyperspectral images. **PLOS ONE**, Public Library of Science, v. 19, p. 1–17, 09 2024. DOI: <https://doi.org/10.1371/journal.pone.0307329>.

ZHANG, A. *et al.* Dive into deep learning. **arXiv preprint arXiv:2106.11342**, 2021. Capítulo “Pooling”. Disponível em: <https://pt.d2l.ai/chapter_convolutional-neural-networks/pooling.html>. Acesso em: 14 out. 2024.

ZHANG, N.; ZANG, E.; LI, F. A soybean classification method based on data balance and deep learning. **Applied Sciences**, v. 13, n. 11, 2023. ISSN 2076-3417. DOI: <https://doi.org/10.3390/app13116425>.