

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**CornGrain: base de dados e classificação automática
de grãos de milho**

Adriana Casadei

Monografia - MBA em Inteligência Artificial e Big Data

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Adriana Casadei

CornGrain: base de dados e classificação automática de grãos de milho

Monografia apresentada ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientador: Prof. Dr. Tiago Agostinho de Almeida

Versão original

São Carlos

2023

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

C334c Casadei, Adriana
CornGrain: base de dados e classificação
automática de grãos de milho / Adriana Casadei;
orientador Tiago Agostinho de Almeida. -- São
Carlos, 2023.
58 p.

Trabalho de conclusão de curso (MBA em
Inteligência Artificial e Big Data) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2023.

1. Visão computacional. 2. Classificação de grãos
de milho. 3. Aprendizado de máquina. I. Almeida,
Tiago Agostinho de, orient. II. Título.

Adriana Casadei

**CornGrain: new dataset and automatic corn grain
classification**

Monograph presented to the Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, as part of the requirements for obtaining the title of Specialist in Artificial Intelligence and Big Data.

Concentration area: Artificial Intelligence

Advisor: Prof. Dr. Tiago Agostinho de Almeida

Original version

São Carlos

2023

Este trabalho é dedicado a todos que tenham interesse na área de aprendizado de máquina, em especial, aplicado a soluções para o agronegócio.

AGRADECIMENTOS

O desenvolvimento desse trabalho se deve ao encorajamento, paciência e auxílio de várias pessoas, dentre elas, agradeço:

Ao professor Tiago Agostinho de Almeida, por ter sido meu orientador e ter desempenhado tal função com dedicação e serenidade.

Aos meus familiares e amigos, por todo o apoio e incentivo, que muito contribuíram para a realização deste trabalho.

Por fim, agradeço a todas as pessoas que direta ou indiretamente colaboraram, de alguma forma, para que este trabalho fosse concluído.

RESUMO

CASADEI, A. **CornGrain: base de dados e classificação automática de grãos de milho**. 2023. 58p. Monografia (MBA em Inteligência Artificial e Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2023.

A população mundial continua aumentando. Já somos oito bilhões de pessoas e, para garantir a oferta de alimentos em quantidade suficiente para todos, é necessário que a produção também aumente, porém de forma sustentável e equilibrada. O Brasil pode desempenhar um papel importante nesse cenário, já que é um dos maiores produtores de grãos do mundo. Contudo, além de focar no aumento da capacidade produtiva, é importante garantir a qualidade dos grãos produzidos, tanto para o mercado interno quanto externo. A classificação de grãos é um processo minucioso que analisa amostras de grãos objetivando determinar a qualidade de um lote. O processo de análise, em geral, é realizado por uma pessoa treinada, porém um grande desafio é a subjetividade, ainda mais quando o grão está no limiar do que exige a normativa. A tecnologia pode auxiliar na otimização dessa tarefa, principalmente, nas áreas de visão computacional e aprendizado de máquina. Neste contexto, o objetivo deste trabalho é criar uma base de dados com amostras rotuladas de grãos de milho e, em seguida, treinar modelos no estado-da-arte para realizar a classificação dos grãos, oferecendo uma análise comparativa para identificar o modelo mais promissor. Este trabalho compara os resultados das redes YOLO v8, EfficientDet e Faster R-CNN com arquiteturas variando das mais simples e leves à mais completas respectivamente. Os resultados alcançados demonstram que, para o problema de classificação de qualidade de grãos, qualquer um dos modelos treinados alcança o desempenho necessário para realizar a tarefa com sucesso. Assim, conclui-se que a rede YOLO v8, sendo a mais leve entre as analisadas, é adequada para a solução do problema e automatização do processo de classificação de grãos.

Palavras-chave: Visão computacional, Grãos de milho, Aprendizado de máquina, YOLO, Faster R-CNN, EfficientDet.

ABSTRACT

CASADEI, A. **CornGrain: new dataset and automatic corn grain classification.** 2023. 58p. Monograph (MBA in Artificial Intelligence and Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2023.

The world's population continues to increase. We are already eight billion people and, in order to guarantee the supply of food in sufficient quantity for everyone, it is necessary that production also increase, but in a sustainable and balanced way. Brazil can play an important role in this scenario, as it is one of the largest grain producers in the world. However, in addition to focusing on increasing production capacity, it is important to ensure the quality of the grains produced, both for the domestic and foreign markets. Grain grading is a thorough process that analyzes grain samples in order to determine the quality of a batch. The analysis process, in general, is carried out by a trained person, but a major challenge is subjectivity, even more so when the grain is on the threshold of what the regulations require. Technology can help optimize this task, especially in the areas of computer vision and machine learning. In this context, the objective of this work is to create a database with labeled samples of corn grains and then train state-of-the-art models to perform grain classification, offering a comparative analysis to identify the most promising model. . This work compares the results of the YOLO v8, EfficientDet and Faster R-CNN networks with architectures ranging from the simplest and lightest to the most complete respectively. The results achieved demonstrate that, for the grain quality classification problem, any of the trained models achieves the necessary performance to successfully perform the task. Thus, it is concluded that the YOLO v8 network, being the lightest among those analyzed, is suitable for solving the problem and automating the grain classification process.

Keywords: Computer vision, Corn kernels, Machine learning, YOLO, Faster R-CNN, EfficientDet.

LISTA DE FIGURAS

Figura 1 – Exemplos de amostras de cada classe: ponto preto, padrão, película, inteiro e corpo estranho, respectivamente.	24
Figura 2 – Pipeline para classificação de grãos	28
Figura 3 – Arquitetura geral de uma Rede Neural Convolutacional	30
Figura 4 – Arquitetura da rede: Faster R-CNN	33
Figura 5 – Arquitetura das redes: MobileNet e MobileNetv2	34
Figura 6 – Arquitetura da rede EfficientDet	35
Figura 7 – Arquitetura da rede: YOLO v1	36
Figura 8 – Exemplo amostral usando o fundo cinza	42
Figura 9 – Exemplo de marcação dos grãos de milho	42
Figura 10 – Exemplo das transformações aplicadas nas imagens	43
Figura 11 – Exemplos de técnicas de aumento de dados	44

LISTA DE TABELAS

Tabela 1 – Distribuição das classes da base de dados CornGrain	43
Tabela 2 – Comparação da base de dados original e após a aumentação dos dados, com as respectivas proporções das classes (%)	45
Tabela 3 – Distribuição das amostras: conjunto de treino e validação	47
Tabela 4 – Hiperparâmetros dos modelos configurados empiricamente	48
Tabela 5 – Acurácia obtida pelos modelos na classificação dos exemplos de cada classe	49
Tabela 6 – Sensibilidade obtida pelos modelos na classificação dos exemplos de cada classe	49
Tabela 7 – F1-Score obtida pelos modelos na classificação dos exemplos de cada classe	49
Tabela 8 – Comparação da quantidade de parâmetros e do tempo de treinamento e de inferência	50

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Hipótese e Objetivos	24
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	Segurança alimentar, tecnologia e o agronegócio	27
2.2	Modelos de classificação de imagem	28
2.2.1	Aquisição de imagens	28
2.2.2	Pré-processamento	29
2.2.3	Segmentação	29
2.2.4	Classificação	29
2.2.5	Redes neurais convolucionais	29
2.2.5.1	LeNet-5	30
2.2.5.2	AlexNet	31
2.2.5.3	ZFNet	31
2.2.5.4	VGGNet	31
2.2.5.5	GoogLeNet	31
2.2.5.6	ResNet	31
2.2.5.7	Faster R-CNN	31
2.2.5.8	MobileNet	32
2.2.5.9	EfficientNet	34
2.2.5.10	YOLO	35
3	CONTEXTUALIZAÇÃO BIBLIOGRÁFICA	37
4	A BASE DE DADOS CORNGRAIN	41
4.1	Coleta das amostras	41
4.2	Captura das imagens	41
4.3	Anotação	42
4.4	Transformação	43
4.4.1	HueSaturationValue	45
4.4.2	RandomBrightnessContrast	45
4.4.3	HorizontalFlip	45
5	EXPERIMENTOS E RESULTADOS	47
5.1	Treinamento dos modelos	47
5.2	Resultados	48

6	CONCLUSÃO	53
	REFERÊNCIAS	55

1 INTRODUÇÃO

O Brasil desponta como um grande produtor e exportador de grãos, porém nas décadas de 1960 e 1970 embora o país estivesse em crescente industrialização e urbanização, o setor agrícola não acompanhava esse processo e sofria com a baixa produtividade. Uma parte considerável do abastecimento interno de alimentos era proveniente de importações. Outro ponto foi a intensificada migração rural-urbana, resultante do baixo rendimento do trabalho no campo. Buscando garantir segurança alimentar da população, principalmente a urbana que estava em expansão, e reduzir os preços dos alimentos, o governo instituiu políticas visando aumentar a produção e a produtividade agrícola, como investimentos em pesquisa e desenvolvimento, crédito rural subsidiado entre outras ações. Neste contexto, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) foi criada em 1973 pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA) para desenvolver um modelo de agricultura e pecuária adequado para o clima do país, sendo considerado um dos pontos essenciais para o aumento da produtividade e a melhora no desempenho da agropecuária e agricultura (PROCÓPIO, G. M. P. R., 2022).

Outro fator que atrasou o desenvolvimento da agricultura brasileira foi a extrema regulação do governo: o excesso de burocracia dificultava o processo de comercialização. A abertura comercial ocorreu no começo dos anos 1990 e a diminuição do controle regulatório sobre a produção e exportação de produtos agrícolas foi essencial para a expansão do agronegócio (IPEA, 2017).

Os objetivos de desenvolvimento sustentável estabelecidos sob a coordenação da Organização das Nações Unidas (ONU) visam garantir um planeta mais próspero, equitativo e saudável até 2030. Para esse objetivo ser alcançado, é necessário estabelecer uma relação de equilíbrio entre população e ambiente e o processo de produção de alimentos e energia. Espera-se que o Brasil desempenhe um papel de destaque nesse contexto, já que nas últimas cinco décadas, o país se estabeleceu como um dos mais importantes produtores e exportadores mundiais de alimentos, atendendo aproximadamente 1,5 bilhão de pessoas em todo o mundo (EMBRAPA, 2018).

Nas últimas décadas, a produção e a produtividade do plantio de milho cresceu significativamente. Além do aumento na produção, a produtividade por área plantada também aumentou, sendo em 1991 de 1,79 toneladas por hectare e, em 2021, 4,36. Esse aumento da produtividade está diretamente relacionado com o avanço em pesquisas, desenvolvimento de variedades de sementes, avanços na genética, entre outros fatores que tiveram relevância no aumento da eficiência no campo (CONAB, 2023).

Outro ponto que alavancou a produção de milho foi o cultivo nos solos do cerrado

brasileiro, pois antes o milho era plantado no verão das regiões Sul e Sudeste. Com apoio de estudos e pesquisas, foram desenvolvidas técnicas que viabilizaram o cultivo no cerrado, que, no caso do milho, ocorre nessa região predominantemente no inverno (EMBRAPA - ED.: 829, 2018).

O Brasil foi o terceiro maior produtor mundial de milho em 2020, com 100 milhões de toneladas, equivalente a 8,2% do total mundial, sendo superado apenas pelos Estados Unidos e China. Outro marco importante para o país é o crescimento das exportações de milho: em 2020, o Brasil exportou 38 milhões de toneladas do grão, equivalente a 19,8% das exportações totais do produto, segundo estudo da *Food and Agriculture Organization of the United Nations* (EMBRAPA, 2020).

Estudos arqueológicos e geológicos, realizados através de medições por desintegração radioativa, evidenciaram que o milho é cultivado há pelo menos cinco mil anos, sendo um dos cultivos mais antigos do mundo. Existem evidências de que sua origem tenha ocorrido no México, América Central ou sudeste dos Estados Unidos. Após o descobrimento da América, o milho foi levado para a Europa, onde a princípio era cultivado em jardins. Assim que seu valor nutricional foi conhecido, passou a ser plantado em grande escala e se espalhou por vários países (PESTANA *et al.*, 2014).

O milho é um cereal multifacetado, sua diversidade de utilizações, da alimentação animal até a indústria, faz com que seja um produto de grande importância econômica. O cultivo de milho é muito antigo e nos dias atuais se fixa como a cultura agrícola mais importante do mundo, ultrapassando 1 bilhão de toneladas produzidas anualmente. A estimativa é que há aproximadamente 3.500 aplicações diferentes para o milho, como alimentação de animais e pessoas, combustíveis, excipientes de medicamento, detergentes, polímeros, vitaminas, tintas, gomas de mascar, cervejas, entre muitas outras (EMBRAPA - ED.: 829, 2018).

Considerando o milho em grão, cerca de 70% da produção no mundo é utilizada como alimentação animal. No Brasil, essa porcentagem oscila de ano para ano, variando de 60% a 80%. Contudo, apesar de não ter uma participação expressiva, o uso do milho em grão na alimentação humana, através de seus derivados, é fator relevante no uso do cereal, principalmente, em regiões de baixa renda (DUARTE; MATTOSO; GARCIA, 2008). Devido a sua composição, o milho é fonte de energia, proteína, gordura e fibras. O grão seco, é composto aproximadamente de 72% de amido, 9,5% de proteínas, 9% de fibras e 4% de óleo (PAES, 2006).

Para ser considerado apropriado para o consumo, o milho precisa passar por um processo de classificação para definir pontos essenciais de qualidade e segurança. Conforme informa o MAPA, a classificação vegetal é o ato de se determinar a qualidade de um produto, mediante análises específicas e por comparação entre a amostra analisada e os padrões aprovados pelo Governo Federal. Além da classificação oficial, regulamentada pelo

MAPA, pode ocorrer a estabelecida entre o comprador e o vendedor conforme definição em contrato, conhecida como classificação comercial. Nesses casos, em geral, o comprador exige um padrão de classificação a ser cumprido pelo produtor dos grãos. Afinal, para a comercialização e processamento desse cereal, a qualidade dos grãos é um fator essencial e que pode afetar o valor final do produto. Em geral, a qualidade é assegurada realizando a classificação dos lotes dos grãos de milho (SENAR, 2017).

Por lei, é exigido que os grãos sejam classificados quando destinados diretamente à alimentação humana, devendo atender às normas vigentes de segurança alimentar, nas operações de compra e venda do poder público e para importação. O dispositivo legal que regulamenta a classificação de grãos é a Lei nº 9.972, de 25 de maio de 2000 e o decreto regulamentador nº 3.664, de 17 de novembro de 2000. No artigo 3º, é definido que todos os produtos vegetais, subprodutos e resíduos de valor econômico, que possuam padrão oficial definido pelo MAPA deverão ser classificados. Classificar um produto vegetal é a ação de determinar sua qualidade, além da qualidade de seus subprodutos e resíduos de valor econômico, com base nos padrões oficiais definidos pelo MAPA. A classificação, consiste na identificação do produto em determinado grupo, classe e tipo conforme a qualidade apresentada. O resultado da classificação é indicado no certificado ou laudo de classificação (BRASIL, 2000).

Atualmente, a classificação pode ser feita por dois profissionais, o classificador oficial, que além da formação superior ou técnica na área agrícola, deve ser devidamente registrado no MAPA e o classificador treinador/prático, que é o profissional graduado ou não que tem conhecimento e experiência na classificação (SENAR, 2017).

No referido trabalho será usado para classificação o grão de milho da classe amarela destinado, em geral, para produção de cereais (canjica). A canjica possui algumas variedades comercializadas no mercado brasileiro como a canjica branca e a amarela, ambas são feitas a partir dos grãos de milho, de variedades diferentes. Durante o processo de pesquisa foi encontrado apenas uma base de dados de milhos apodrecidos¹, não sendo relevante para o escopo da pesquisa.

Um ponto relevante, ao utilizar técnicas de aprendizado de máquina e visão computacional, é a qualidade da base de dados, pois o algoritmo irá aprender conforme as imagens apresentadas. Portanto, é crucial a obtenção de um conjunto de dados de qualidade, rotulados de maneira responsável e confiável, do contrário, o resultado pode não condizer com a realidade.

Conforme orientação da empresa que disponibilizou as amostras, será considerada as seguintes classes (Figura 1):

¹ Base de dados de grãos de milho apodrecidos: <https://ieee-dataport.org/open-access/rot-corn-grains>. Acessado em: 05/05/2023

1. Ponto preto: grãos ou pedaços de grãos atacados por insetos em qualquer uma das fases evolutivas;
2. Padrão: grãos quebrados e sem outros defeitos, usados para destinação final;
3. Película: grãos que apresentam uma película esbranquiçada independente de seu estado;
4. Inteiro: grãos inteiros (não quebrados); e
5. Corpo estranho: elementos como pequenos pedaços de terra, pequenos pedaços de folhas, que não fazem parte do escopo definido.



Figura 1 – Exemplos de amostras de cada classe: ponto preto, padrão, película, inteiro e corpo estranho, respectivamente.

Fonte: Autorial própria

1.1 Hipótese e Objetivos

Tendo em vista:

1. a importância do processo de classificação de grãos para sua comercialização e que, em geral, esse processo é feito de forma manual, tendendo assim a subjetividade do classificador, principalmente quando o grão de milho está no limiar do que exige a normativa; e
2. que com o crescimento populacional é necessário utilizar diversos recursos para aumentar a produtividade e otimizar os processos na produção dos alimentos.

Esse projeto avalia a hipótese de que é possível automatizar, potencializar e diminuir o tempo do processo de classificação de grãos utilizando métodos de aprendizado de máquina e visão computacional.

Os objetivos gerais deste trabalho são:

1. Criar e disponibilizar uma base de dados rotulada de grãos de milho;
2. Treinar modelos supervisionados de aprendizado de máquina e visão computacional;

3. Fornecer um marco com os resultados obtidos para comparações futuras.

Como a amostragem é uma atividade manual e demorada, com a automatização proposta neste trabalho será possível diminuir o tempo destinado a essa atividade, possibilitando a análise de uma quantidade maior de amostras e uma classificação mais apurada do lote de milho.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, é apresentada a conceituação teórica das técnicas usadas e desenvolvidas ao longo desse projeto. A fundamentação teórica está organizada da seguinte forma: a Seção 2.1 contextualiza a problemática da importância do aumento da produtividade de alimentos para abastecer toda a população; e, a Seção 2.2 apresenta métodos de tratamento de imagens, e conceituação dos principais modelos existentes.

2.1 Segurança alimentar, tecnologia e o agronegócio

Somos oito bilhões de pessoas vivendo no Planeta Terra, e o crescimento populacional continua aumentando. Consequentemente, a indústria de alimentos precisa aumentar sua produção e melhorar a qualidade dos produtos ofertados (WEFORUM, 2022).

Os cereais são insumos para diversos produtos finais, tanto para consumo humano como animal. De acordo com a *Food and Agriculture Organization*, em 2020, a produção mundial de cereais foi de 2.765 toneladas, sendo 2% a mais que em 2019 (FAO, 2023).

Seja por fatores climáticos ou de preservação ambiental, é difícil e insustentável aumentar a área de plantio, portanto, são necessárias melhorias em outros processos da cadeia produtiva, visando otimizar processos e aumentar a produtividade. Um exemplo dessa otimização é a utilização de Técnicas de Inteligência Artificial que podem auxiliar na automatização da classificação de grãos alimentícios. Um ponto que merece atenção no processo de classificação de grãos é entender que cada variedade apresenta necessidades e características específicas que devem ser seguidas para garantir a qualidade do processo (VELESACA *et al.*, 2020).

Segundo Faceli *et al.* (2004), apesar do Brasil ser um produtor de destaque no mercado mundial no setor de agronegócio, ainda existe muita margem para melhoria na exploração desta atividade econômica. Neste contexto, a incorporação da Inteligência Artificial, especialmente, de Aprendizado de Máquina (AM), pode ser melhor aproveitada na análise inteligente dos dados gerados.

Inclusive, Borth *et al.* (2014) apresentam algumas aplicações da visão computacional que auxiliam no desenvolvimento e no avanço do agronegócio, como a agricultura de precisão que considera as características e necessidades da propriedade ou plantio de forma precisa para aplicar insumos ou outras possibilidades de forma correta, na quantidade e no momento certos, trazendo um aumento na qualidade e produtividade. O uso de veículos aéreos não tripulados para mapeamento de áreas plantadas, traz celeridade nas análises. Outro ponto, é a aplicação e pesquisas sobre o controle de qualidade de alimentos, que, em geral, é feito de forma manual, mas a automatização proporciona uma inspeção certa e

rápida do processo.

A inserção de tecnologias digitais no agronegócio é uma realidade que se fortificou com a necessidade da redução de custos e a demanda por produtos de qualidade e procedência, uma vantagem de seu uso é a capacidade de processar volumes gigantescos de informações, automatizar processos e a troca de informações e conhecimentos (IPEA, 2022).

2.2 Modelos de classificação de imagem

A classificação de grãos é uma forma de automatização de processos que utiliza métodos de Inteligência Artificial e deve seguir algumas etapas. Velesaca *et al.* (2020) sugerem um pipeline geral (Figura 2) para a classificação de grãos que inclui as etapas de: aquisição das imagens, pré-processamento, segmentação e classificação, conforme apresentado a seguir.

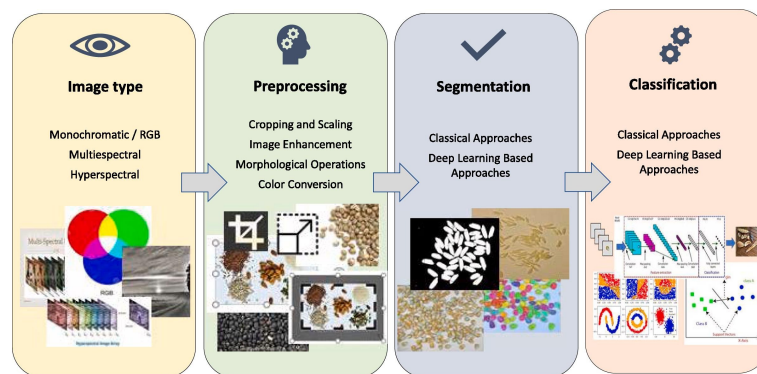


Figura 2 – Pipeline para classificação de grãos

Fonte: Velesaca *et al.* (2020)

2.2.1 Aquisição de imagens

O primeiro passo do processo de classificação começa com a aquisição das imagens para montar a base de dados. Para classificação de grãos, são consideradas imagens em escala de cinza, preto e branco e imagens coloridas. Assim, a maioria dos estudos presentes na literatura utiliza apenas um canal, devido ao baixo custo e facilidade de uso da câmera usual RGB (sem a necessidade de informação de profundidade ou infravermelho, por exemplo). Porém, na revisão feita por Velesaca *et al.* (2020) existem estudos que mostram o processamento de imagens em outras bandas espectrais, como o infravermelho. Outra abordagem mais recente é a utilização de imagens multiespectrais e hiperespectrais, que cobrem além do espectro visível, também as bandas espectrais ultravioleta e infravermelha. Uma vantagem da utilização dessas abordagens é obter informações mais específicas para o processo de classificação, que não estão disponíveis na utilização de apenas um canal.

2.2.2 Pré-processamento

As imagens devem ter um padrão de formato, então dependendo de como as imagens estão e a finalidade de uso, é necessário o pré-processamento para obter uma base de dados homogênea e abrangente. Entre as técnicas mais usuais de pré-processamento estão: corte e redimensionamento, melhorias na imagem, operações morfológicas e conversão de cores (GARCIA; CANDELO-BECERRA; HOYOS, 2019).

2.2.3 Segmentação

As técnicas de segmentação têm por objetivo dividir uma imagem em diferentes regiões que a constituem e a diferenciam para posterior análise e extração de atributos para cálculos descritivos (CONCI; AZEVEDO; LETA, 2007). Ainda, a etapa de segmentação é crítica, pois quaisquer erros ou distorções nas regiões de foco podem refletir nas demais etapas, podendo, inclusive, prejudicar a eficiência de todo o processo. A limiarização é uma técnica de segmentação amplamente utilizada para imagens de grãos. Ela lida com imagens em escala de cinza e realiza a binarização aplicando um valor limiar, que varia dependendo do tipo de grão e da cor de fundo da imagem (WANG; PALIWAL, 2003).

2.2.4 Classificação

A partir de um conjunto de dados consolidado, com as imagens devidamente tratadas, padronizadas e rotuladas, o próximo passo é treinar modelos de Aprendizado de Máquina e Aprendizado Profundo capazes de resolver o problema. Segundo Velesaca *et al.* (2020), as abordagens identificadas podem ser agrupadas em duas categorias: abordagem clássica de reconhecimento de padrões, como por exemplo: SVM, K-NN, *linear discriminant analysis* e abordagem de aprendizado profundo. Basicamente, na abordagem clássica, o objetivo é obter os atributos mais representativos e, em seguida, usar um algoritmo de aprendizado de máquina para o treinamento. Com a evolução da tecnologia e do hardware, como o aumento na capacidade de memória, que facilitou o processamento paralelo de grandes quantidades de dados, entre outros fatores, as técnicas de Aprendizado Profundo demonstraram uma vantagem nas atividades da área da visão computacional em relação às técnicas da abordagem clássica (ALTUNTAS; KOCAMAZ; YEROGLU, 2019).

2.2.5 Redes neurais convolucionais

As Redes Convolucionais Profundas são uma das redes mais representativas no campo do Aprendizado Profundo, principalmente por seu desempenho nas áreas de visão computacional e processamento de linguagem natural. A visão computacional baseada em Rede Convolucional Profunda permitiu que fossem realizadas tarefas consideradas impossíveis no passado, como reconhecimento facial, navegação autônoma, tratamentos médicos auxiliados por inteligência artificial, entre outros (LI *et al.*, 2021).

As Redes Convolucionais podem ter diferentes estruturas, porém também apresentam, em geral, alguns componentes principais, conforme exemplificado na Figura 3. Os *kernels* ou filtros são usados para determinar o mapa de características durante a convolução. A camada de convolução tem por objetivo aprender características discriminantes do objeto de entrada que alimenta a rede. A função de ativação, por sua vez, torna a rede mais preparada para tarefas mais complexas ao aplicar transformações não lineares nos dados de entrada. Em seguida, pode haver camadas de *pooling*, que atuam de forma a reduzir seletivamente a quantidade de parâmetros da rede, assim, agrupando os parâmetros importantes para a camada de saída. Ao final das camadas convolucionais, existe ao menos uma camada totalmente conectada (*fully connected layer*) que utiliza-se dos parâmetros treinados para produzir uma resposta ao final da rede (NARANJO-TORRES *et al.*, 2020).

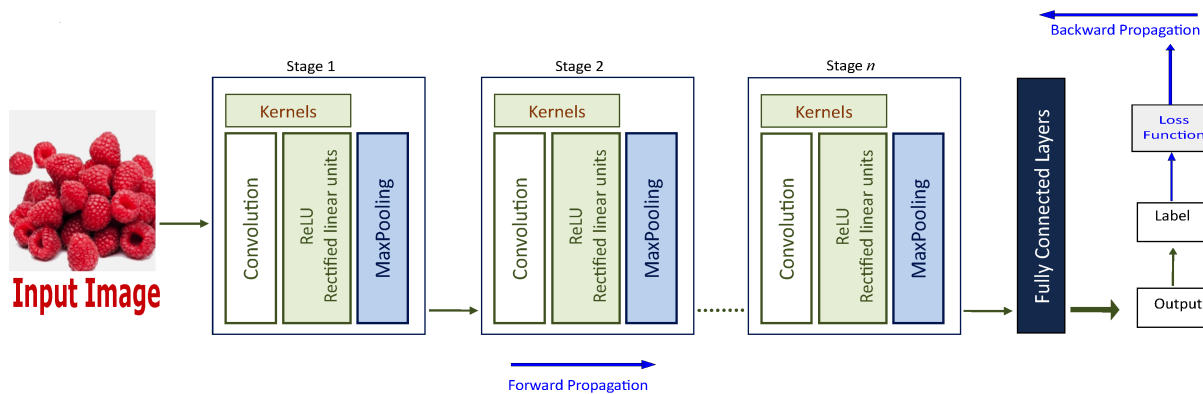


Figura 3 – Arquitetura geral de uma Rede Neural Convucional

Fonte: Naranjo-Torres *et al.* (2020)

Parte dessa evolução pode ser acompanhada pela competição *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC), realizada desde 2010, com objetivo de avaliar algoritmos de detecção de objetos e classificação de imagens em larga escala. A seguir, são apresentadas as principais arquiteturas de redes convolucionais profundas existentes na literatura.

2.2.5.1 LeNet-5

LeNet-5 é uma Rede Convolutional Profunda proposta para o reconhecimento de caracteres manuscritos e impressos à máquina de escrever. Esta rede emprega função de ativação *softmax* e é composta por 5 camadas de aprendizado, 3 conjuntos de camadas convolucionais combinadas com *average pooling* e 2 camadas totalmente conectadas. A sua importância na evolução das Redes Convolucionais Profundas se deve ao fato de ser uma das pioneiras em redes pré-treinadas disponíveis (LECUN *et al.*, 1998).

2.2.5.2 AlexNet

Em 2012, a rede AlexNet superou os demais concorrentes da competição ImageNet ao reduzir o erro top-5 para 15,3%, em comparação com o segundo melhor resultado que foi de 26,2%. A rede é estruturada com 8 camadas de aprendizagem, sendo 5 camadas convolucionais e 3 camadas totalmente conectadas, além de camadas de *max pooling* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

2.2.5.3 ZFNet

O vencedor da ILSVRC no ano de 2013 foi a Rede Neural Convolutiva Profunda ZFNet. Ela tem uma arquitetura similar à AlexNet, com o diferencial na mudança dos hiperparâmetros que, com isso alcançou a taxa de erro top-5 em 14,8% (ZEILER; FERGUS, 2014).

2.2.5.4 VGGNet

Em 2014, a vencedora foi a VGGNet, que garantiu o primeiro lugar na tarefa de localização e segundo lugar na área de classificação. Em vista das anteriores, foi considerada uma rede bem profunda, pois sua arquitetura foi planejada em várias versões, com a quantidade de camadas variando em 11, 13, 16 e 19 dependendo do tipo de implementação. Ela obteve a taxa de erro top-5 em 7,3% (SIMONYAN; ZISSERMAN, 2015).

2.2.5.5 GoogLeNet

O vencedor da competição ILSVRC em 2014 foi a GoogLeNet ou Inception, construída pelo Google. Esta rede alcançou uma taxa de erro top-5 de 6,67%, ficando inclusive bem próxima da taxa de erro top-5 alcançada por um especialista humano, que atingiu 5,51%. A GoogLeNet foi inspirada na LeNet, possui 22 camadas e reduzido número de parâmetros. Esta rede apresentou um novo conceito, o módulo inception, que tem por objetivo atuar como extrator de características (SZEGEDY *et al.*, 2015).

2.2.5.6 ResNet

Em 2015, o vencedor da competição foi a *Residual Neural Network* (ResNet) que apresentou uma nova arquitetura com o *skip connection* e normalização em lote. A rede possui 152 camadas e atingiu uma taxa de erro top-5 de 3,57%, superando o desempenho em nível humano no mesmo conjunto de dados (HE *et al.*, 2015).

2.2.5.7 Faster R-CNN

A Faster R-CNN é uma evolução da família de modelos R-CNN, principalmente em termos de eficiência computacional e melhoria no desempenho.

O sistema de detecção utilizado nas redes R-CNN (Figura 4) consiste em dois módulos distintos: No primeiro módulo, recebe-se uma imagem e são identificadas as regiões de interesse, ou seja, são delimitadas as áreas que possivelmente contêm objetos de interesse. Isso é realizado por meio de um algoritmo especializado nessa tarefa: a Rede de Proposta de Região (RPN).

Essa rede gera propostas de regiões de forma mais eficiente, permitindo ganhos significativos de processamento. O algoritmo da RPN percorre um mapa de características (imagem processada por um filtro que busca capturar as características dos objetos) usando janelas deslizantes, que possuem âncoras para encontrar regiões de interesse em cada instância das janelas. Diversas âncoras de diferentes tamanhos e proporções são aplicadas em cada janela, de modo que objetos de diferentes tamanhos possam ser identificados, tornando assim, mais preciso encontrar objetos independente de sua localização na imagem.

No próximo módulo da rede R-CNN, um detector utiliza as regiões delimitadas no primeiro módulo, possibilitado pelas camadas convolucionais compartilhadas entre os módulos, com o objetivo de otimizar o processo de treinamento e detecção de objetos. Para finalizar, o treinamento é feito de forma alternada entre o treinamento da rede RPN e da rede detectora conforme quantidade de iterações definida.

A Fast R-CNN usa um algoritmo de busca seletiva para agrupar pixels e identificar as regiões de interesse, que demanda, em média, 2 segundos por imagem. Com a Faster R-CNN, usando a RPN, esse tempo diminuiu para 10 milissegundos.

Essa nova proposta de modelo ficou em primeiro lugar nas trilhas de detecção do ImageNet, na competição ILSVRC e COCO de 2015, o que sugere que o método além de ser uma solução de uso prático, também é eficaz na melhora da precisão da detecção de objetos (REN *et al.*, 2015).

2.2.5.8 MobileNet

Em 2017, o Google apresentou a Rede Neural Convolutiva Profunda MobileNet v1 que possibilitou executar o modelo em sistemas embarcados, como celulares. Sua arquitetura é simples, usa convoluções separáveis em profundidade para formar a rede, e serviu de inspiração para diversos outros projetos (HOWARD *et al.*, 2017). A Figura 5 apresenta a arquitetura das redes MobileNet e MobileNet v2.

Na sequência, foi desenvolvida a MobileNet v2 que adicionou gargalos (*bottleneck*) que codificam na entrada e saída intermediárias do modelo, permitindo que a camada interna consiga encapsular conceitos de baixo nível, como pixels, e transformá-los em conceitos de alto nível, como categorias de imagem (SANDLER *et al.*, 2018).

Por último, foi apresentada a MobileNet v3, desenvolvida em dois modelos: MobileNet v3 Large e MobileNet v3 Small. Esta versão utiliza as técnicas MNASNet e NetAdapt,

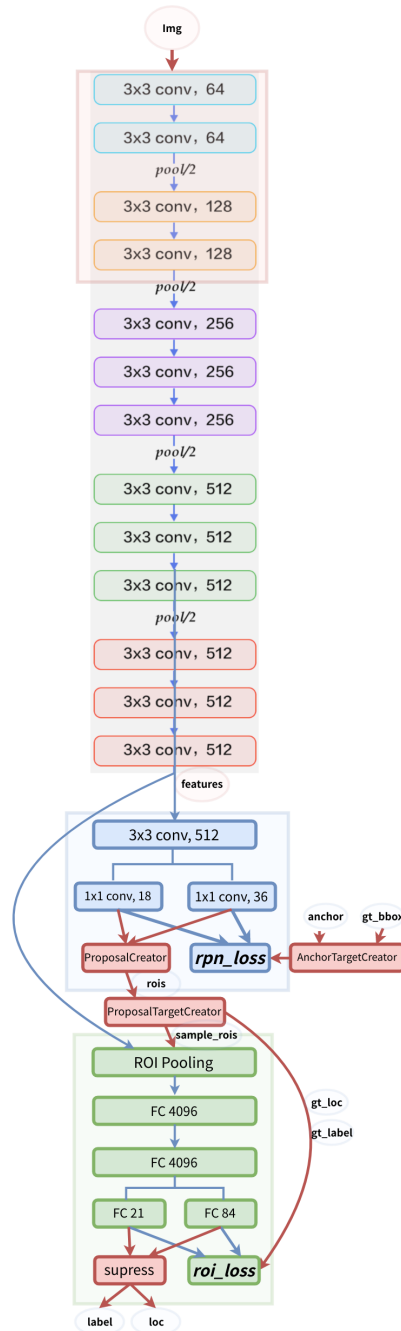


Figura 4 – Arquitetura da rede: Faster R-CNN

Fonte: <https://github.com/chenyuntc/simple-faster-rcnn-pytorch>. Acessado em: 10/04/2023

sendo que a primeira usa aprendizado por reforço para selecionar a configuração de rede ideal e, a segunda, ajusta a configuração indicada otimizando através da exclusão de canais de ativação subutilizados.

Outra novidade incorporada na MobileNet v3 são os blocos *squeeze* e *excite* que visam melhorar a qualidade das representações produzidas por uma rede ao permitir que ela recalibre os recursos e, com isso, consiga dar ênfase aos recursos informativos e suprimir os menos úteis.

A família de redes MobileNet apresentam bom desempenho para problemas de classificação, segmentação e detecção de objetos. A MobileNet v3 large apresentou uma acurácia 3,2% maior e uma latência 20% menor em comparação com MobileNet v2 no conjunto de dados ImageNet na tarefa de classificação. Já a MobileNet v3 small apresentou uma acurácia 6,6% maior com o MobileNet v2 se comparado com a mesma latência.

Para detecção de objetos a MobileNet v3 large se mostrou 25% mais rápida com aproximadamente a mesma precisão da MobileNet v2 no conjunto de dados COCO¹ (HOWARD *et al.*, 2019).

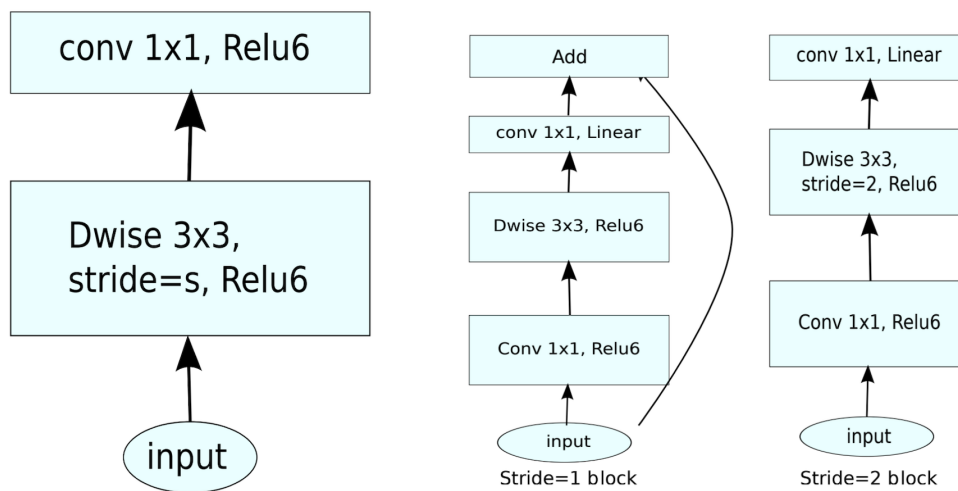


Figura 5 – Arquitetura das redes: MobileNet e MobileNetv2

Fonte: Sandler *et al.* (2018)

2.2.5.9 EfficientNet

Em 2020, foi introduzida a família de redes EfficientNet (B1 a B7) com o objetivo de melhorar a eficiência e acurácia dos modelos. Ela cumpre seu objetivo através do método de escalonamento que dimensiona uniformemente a largura, profundidade e resolução da rede com um conjunto de coeficientes de escala fixa. Divergindo da forma convencional que dimensiona de forma arbitrária esses fatores (TAN; LE, 2020).

A EfficientNet obteve bons resultados em comparação com outras redes neurais convolucionais. A EfficientNet-B7 atingiu o estado da arte no top-1 na ImageNet com 84,3%, sendo 8,4 vezes menor e 6,1 vezes mais rápida que a melhor rede neural convolucional até aquele momento (TAN; LE, 2020).

A família de redes EfficientNet também atingiu o estado da arte apresentando

¹ Base de dados COCO, disponível em: <https://cocodataset.org>. Acessado em: 28/04/2023

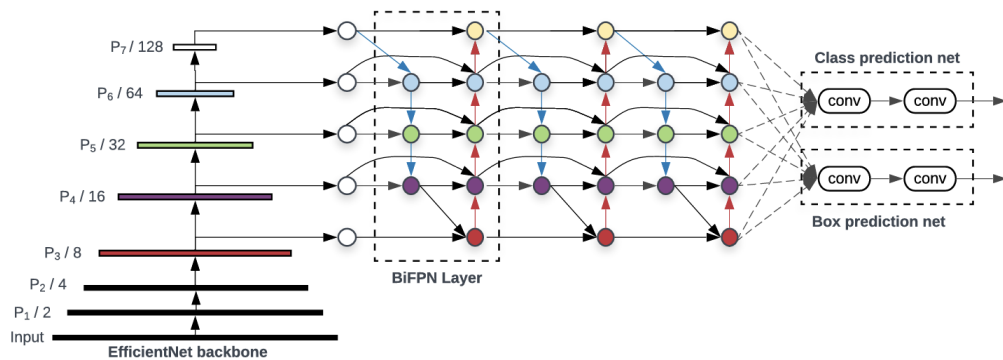


Figura 6 – Arquitetura da rede EfficientDet

Fonte: Tan, Pang and Le (2020)

acurácia de 91,7% no CIFAR-100² e 98,8% no Flowers³. Foi realizada uma prova de conceito, onde primeiro foi usado o Método de Escalonamento criado para EfficientNet em MobileNet e ResNet aplicado na ImageNet ao comparar com um único tipo de escalonamento. O Método de Escalonamento aumentou a acurácia de todos os modelos, sugerindo a efetividade do método proposto (TAN; LE, 2020).

A rede EfficientDet (D1 a D7) (Figura 6) trouxe alguns pontos de melhoria em relação a EfficientNet como a proposta do BiFPN *efficient bidirectional cross-scale connections and weighted feature fusion*, que é uma rede de recursos bidirecional que permite a fusão de recursos de forma rápida e o dimensionamento usando um único fator para modificar a profundidade, largura e resolução das redes backbone e dos recursos (TAN; PANG; LE, 2020).

2.2.5.10 YOLO

A família de modelos YOLO (*You Only Look Once*) foi apresentada em 2015 (Figura 7) e foi evoluindo para novas versões (v1 a v8) com foco na detecção de objetos em tempo real, visando bom desempenho tanto em dispositivos na ponta quanto em nuvem (REDMON *et al.*, 2016).

A YOLO v7, apresentada em 2022, melhorou a performance para dispositivos móveis, GPUs comuns e nuvem e buscou acelerar o processo de treinamento através de um módulo de otimização. A YOLO v7 apresenta uma arquitetura de rede mais rápida com um método de integração de recursos mais eficaz: foi proposto um método de dimensionamento composto que pode manter as propriedades que o modelo tinha

² Base de dados CIFAR-100, disponível em: <https://www.cs.toronto.edu/~kriz/cifar.html>. Acessado em: 28/04/2023

³ Base de dados Flowers, disponível em: <https://www.robots.ox.ac.uk/~vgg/data/flowers/102/>. Acessado em: 28/04/2023

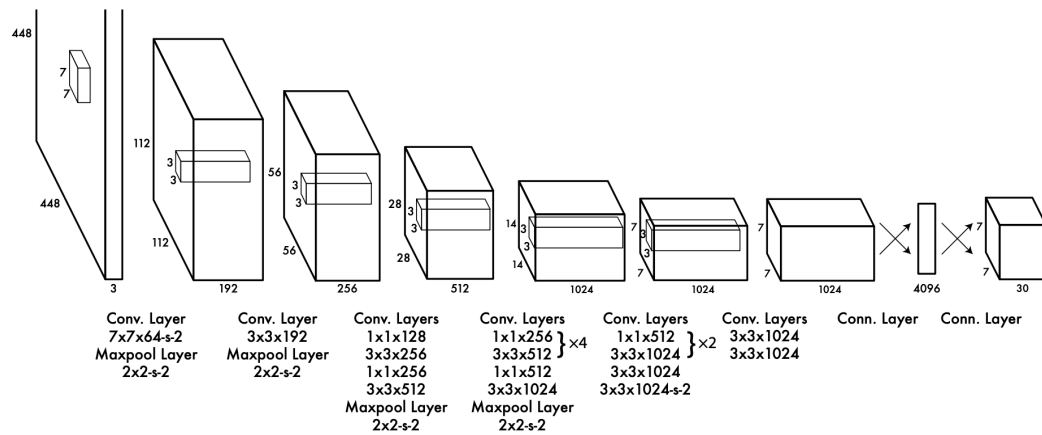


Figura 7 – Arquitetura da rede: YOLO v1

Fonte: Redmon *et al.* (2016)

originalmente ou manter uma estrutura opcional. Esse método, ao escalonar a profundidade de um bloco computacional, dimensiona também a largura considerando a ação anterior (WANG; BOCHKOVSKIY; LIAO, 2022).

Outro recurso adicionado foi o E-ELAN (Extended-ELAN) baseado em ELAN. Não foi modificado todo o gradiente de transmissão da arquitetura original, mas usando um grupo convolucional para aumentar a cardinalidade ao adicionar novos atributos e combiná-los a diferentes grupos de forma aleatória e, assim, mesclar a cardinalidade (WANG; BOCHKOVSKIY; LIAO, 2022). O E-ELAN pode guiar diferentes grupos de blocos computacionais para aprender mais sobre atributos diversos.

A YOLO v7 superou todos os modelos de detecção de objetos conhecidos considerando rapidez e acurácia. Foi treinada no conjunto COCO⁴ sem usar nenhum outro conjunto de dados ou pesos pré-treinados (WANG; BOCHKOVSKIY; LIAO, 2022).

Em janeiro de 2023, foi lançada a YOLO v8 que traz algumas mudanças, como um novo sistema de detecção sem âncora e novas camadas convolucionais para melhorar o desempenho. Outro ponto é que as alterações visam uma melhor experiência para o desenvolvedor, pois o modelo vem empacotado como uma biblioteca, facilitando a instalação. Adicionalmente, a rede YOLO v8 foi criada para funcionar com os modelos treinados em versões anteriores, o que simplifica a alternância entre elas e a comparação e avaliação do desempenho (ULTRALYTICS, 2023).

⁴ Base de dados COCO, disponível em: <https://cocodataset.org/>. Acessado em: 28/04/2023

3 CONTEXTUALIZAÇÃO BIBLIOGRÁFICA

O uso da visão computacional se expandiu para diversos setores devido a sua velocidade e precisão ao atender os requisitos cada vez maiores de produção e qualidade, auxiliando no desenvolvimento de processos totalmente automatizados (BROSNAN; SUN, 2002). Essa expansão se deu inclusive no setor do agronegócio, como, por exemplo, na tarefa de classificação de grãos, onde inúmeras abordagens foram consideradas para resolver diversos problemas ao longo dos últimos 20 anos (VELESACA *et al.*, 2021).

A inspeção de qualidade de produtos agrícolas e alimentícios, em geral, é realizada por pessoas treinadas, mas inspeções manuais são trabalhosas e demoradas. Além disso, em decorrência da subjetividade humana, não se pode garantir uma precisão consistente (BROSNAN; SUN, 2002).

Juntamente com o crescimento da produção (estimada em 312 milhões de toneladas de milho na safra 2022/2023), há também a exigência crescente dos consumidores por padrões de qualidade mais rígidos. O domínio dos padrões de qualidade e as boas práticas pós-colheita, além de atender as necessidades vigentes, agregam valor ao produto trazendo vantagens econômicas tanto para os produtores quanto para as indústrias (EMBRAPA MILHO E SORGO, 2020).

O atendimento dos padrões de qualidade pode acarretar múltiplos benefícios à cadeia de produção, ao produtor e à indústria, como na indicação de melhoria no processo, antecipação de problemas e principalmente, evitar perdas por rejeição e/ou condenação dos lotes, situação que acarreta prejuízos aos produtores e diminui a qualidade do produto final (SILVA, J. S., 2008).

Com a finalidade de definir um padrão oficial de classificação do milho, foi estabelecida a Instrução Normativa número 60, de dezembro de 2011, que considera questões de qualidade, amostragem e rotulagem da espécie *Zea mays L.*¹. Particularmente, após a pandemia de Covid-19, a qualidade dos grãos e as garantias sanitárias passaram a ser um fator importantíssimo, principalmente para desbravar diferentes mercados (SEIXAS, M. A., 2021).

A qualidade dos grãos é determinada a partir de alguns atributos por um especialista do setor, porém, além do erro humano, a análise visual é uma atividade custosa e demorada. Em contrapartida existem vários métodos usados para a classificação de grãos, mas a solução automatizada pode ter um custo inicial elevado. Para mitigar o alto valor de investimento na automatização entre outras limitações, Arora *et al.* (2020) propõe como solução implementar técnicas de processamento de imagem para que o modelo seja capaz

¹ Nome científico do milho pertencente à família Poaceae

de capturar com sucesso variados parâmetros, extrair características e avaliar a quantidade de grãos contida em cada imagem com base nas imagens de amostra, bem como avaliar vários algoritmos de processamento de imagens e de aprendizado de máquina (ARORA *et al.*, 2020).

Koklu and Ozkan (2020) lembram que a classificação das sementes é essencial para garantir sistemas agrícolas sustentáveis. Os autores empregaram a visão computacional para avaliar modelos que melhor classificam sementes e que consigam distinguir entre diferentes variedades. Foram usados algoritmos tradicionais, como *Multilayer perceptron* (MLP), *Support Vector Machine* (SVM), *k-Nearest Neighbors* (kNN) e *Decision Tree* (DT), que obtiveram os respectivos resultados considerando a precisão: 91,73%, 93,13%, 87,92% e 92,52%, sendo o SVM o modelo que alcançou melhor resultado.

Vlasov and Fadeev (2017) apresentaram, uma possível solução para classificação de sementes mais otimizada e precisa com o uso de uma Rede Neural Convolutiva. Como não foi encontrado conjunto de dados público apropriado, os autores criaram uma base com cerca de 500 imagens de sementes. Eles destacaram que apesar das técnicas tradicionais de AM serem mais rápidas e, por vezes, mais baratas, apresentaram um valor de acurácia menor se comparado com uso de técnicas de aprendizado profundo que, apesar de mais custosas para o processamento, obtiveram melhores resultados. Concluíram que, para a tarefa de classificação de grãos e sementes, técnicas de visão computacional são apropriadas para melhorar o processo.

Similarmente, Maheswaran *et al.* (2020) empregaram, uma Rede Neural Convolutiva com uma sequência de camadas convolucionais juntamente com camadas de *max pooling*, camada de ativação e cada camada conectada à camada anterior. Cada camada da RNC realiza uma convolução bidimensional (2D) dos dados e passam por filtros diferentes, como 3x3, 5x5 e 7x7. A rede foi treinada com um conjunto de 250 imagens de diferentes tipos de grãos e obteve 94,45% de precisão ao detectar partículas indesejáveis.

Li *et al.* (2019) ressaltam, a importância da classificação automatizada de milho na agricultura inteligente. Eles propuseram um método de classificação baseado em técnicas de visão computacional e aprendizado de máquina. Primeiro, para discriminar os diferentes tipos de grãos danificados, é usado um método de segmentação para separar os elementos muito próximos (tocando um no outro). Em seguida, são extraídas doze características de cor e cinco de forma de cada grão. Por fim, é usado um modelo treinado para classificar os grãos normais e danificados. Para avaliar o desempenho do modelo, um conjunto de dados com imagens de milho normal e seis tipos de milhos danificados (danos causados por: calor, germes, podridão de espiga, mofo do olho azul, insetos e superfície danificada pelo molde) é usado. O método proposto obteve 96,67% de precisão ao classificar os milhos normais e os quatro primeiros tipos de milho danificados e, 74,76% ao classificar os milhos normais e os seis tipos de milhos danificados.

Por fim, visando resolver o problema de grãos estarem muito próximos, Velesaca *et al.* (2020) propuseram um esquema de segmentação e classificação que visa funcionar em cenários irrestritos, não importando o número de milhos que encostam um ao outro na imagem. Primeiro, foram construídos dois conjuntos com imagens capturadas em ambiente controlado e com boa iluminação: um conjunto com imagens de grãos de milho individuais e o segundo com vários grãos juntos. No segundo conjunto, as imagens foram anotadas com o contorno dos grãos de milho e das impurezas que foram usadas para treinar e testar a fase de segmentação usando o modelo *Mask R-CNN* (HE *et al.*, 2018). Para classificação, foi proposta a CK-CNN, composta por 3 camadas convolucionais e 2 totalmente conectadas, a função de ativação usada é a RELU e *max pooling* para resumir os resultados da operação de convolução. Na última camada, totalmente conectada, é usada a função de ativação *softmax* para obter a distribuição de probabilidade correspondente a cada tipo de classe: milho bom, milho defeituoso e impureza. A rede CK-CNN foi treinada com o otimizador Nesterov ADAM (NADAM) e com uma taxa de aprendizado de 0,0002. O modelo proposto foi utilizado para classificar 2 classes (grãos de milho bons ou defeituosos) e para classificar 3 classes (grãos de milho bons, grãos de milho defeituosos e impurezas). Comparado com as redes VGG16 e ResNet50, o modelo CK-CNN obteve melhores resultados nos dois tipos de classificação.

Com base nos trabalhos apresentados e discutidos, a solução para o problema proposto neste trabalho não é simples e muito menos universal. Entretanto, é perceptível uma tendência de que modelos de Inteligência Artificial que utilizam redes convolucionais e redes profundas se saem melhor nas tarefas de identificação e classificação de grãos. Assim, o presente trabalho toma essa premissa como inspiração na escolha dos modelos.

Como o essencial para qualquer projeto da área de Ciência de Dados são os dados, na sequência, será descrito os detalhes para a criação, rotulação e transformação da base de dados que será utilizada para o treinamento dos modelos.

4 A BASE DE DADOS CORNGRAIN

Este capítulo descreve as etapas de coleta das amostras, captura das imagens, anotações e transformações dos dados realizadas no processo de criação da base de dados CornGrain.

4.1 Coleta das amostras

A primeira etapa realizada foi a coleta das amostras. Os grãos de milho usados pertencem a uma empresa que realiza processamento para produção de canjica. A amostragem foi realizada ao receber a carga com os grãos. A amostragem de grãos é a ação de obter uma fração representativa de um lote, com o objetivo de identificar sua qualidade. Para isso, a amostra obtida deve ter tamanho adequado e apresentar os mesmos componentes do lote de origem em proporções próximas (CONAB, 2015).

Os grãos utilizados neste projeto foram coletados de forma amostral: ao chegar o caminhão, foi verificado o lote e seu peso. Em seguida, um funcionário treinado retirou as amostras fazendo uso de um calador hidráulico¹. Na sequência, os grãos de milho foram reunidos, repartidos em 4 partes iguais e reduzidos no mínimo por 3 vezes até obter 3 kg do produto. Dessa forma, formaram-se 3 amostras com 1 kg em cada, que são representativas do lote. Por último, as amostras foram identificadas e acondicionadas.

4.2 Captura das imagens

O primeiro passo foi preparar um ambiente controlado que funcionasse para que as fotos dos grãos de milho fossem nítidas. Após alguns testes com diferentes cores de fundo, a cor cinza demonstrou ser a mais adequada, proporcionando maior contraste dos grãos (Figura 8). Para o registro, foi utilizada a câmera fotográfica DSLR da marca Nikon D3200, de 14 megapixels. Visando garantir a variabilidade de registros fotográficos, os grãos foram aleatoriamente espalhados, em média, 25 gramas por foto. Para capturar todos os ângulos dos grãos de milho e suas imperfeições, foram capturadas 3 fotos de cada conjunto de 25g, agitando-os aleatoriamente entre cada foto, com o objetivo de expor diferentes lados dos grãos.

¹ Equipamento operado hidraulicamente, usado para amostragem de produtos a granel.

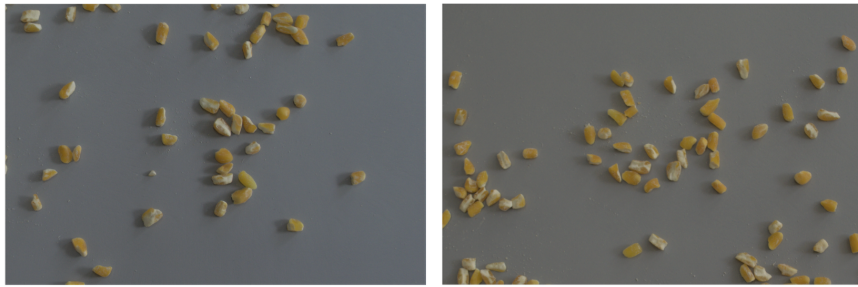


Figura 8 – Exemplo amostral usando o fundo cinza

Fonte: Autoria própria

4.3 Anotação

Nesta etapa, um agrônomo com experiência nos diversos defeitos encontrados nos grãos de milho para canjica compartilhou seu conhecimento. Ele apresentou características e exemplos dos seguintes tipos de grãos que podem ser encontrados na amostra: grão com ponto preto, grão padrão, grão com película, grão inteiro e corpo estranho. Dessa forma, a rotulação dos dados foi feita de forma acurada e condizente com a realidade dos grãos amostrais.

Para a marcação das imagens, foi utilizada a ferramenta CVAT (*Computer Vision Annotation Tool*), que é um software gratuito de anotação de imagem e vídeo voltado para visão computacional.

Para anotação dos rótulos nas imagens, primeiro foram adicionados os rótulos: ponto preto, padrão, película, inteiro e corpo estranho, identificados com cores diferentes e, em seguida, foram usadas as caixas delimitadoras (*bounding boxes*) para contornar o grão de milho e identificá-lo conforme seu rótulo correspondente (Figura 9).

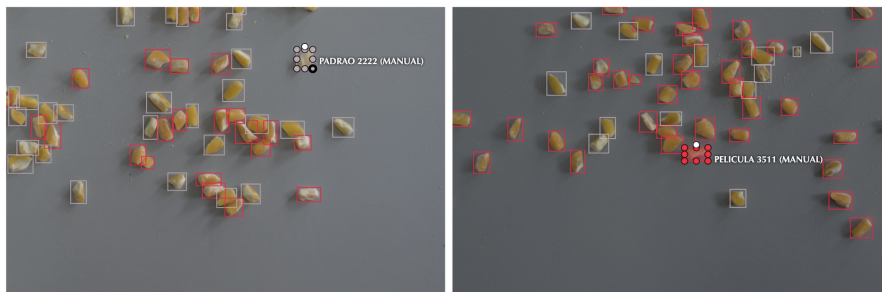


Figura 9 – Exemplo de marcação dos grãos de milho

Fonte: Autoria própria

Após a etapa de anotação, a base de dados denominada CornGrain continha 114

imagens, com a seguinte distribuição de classes (Tabela 1):

Classe	Quantidade	Proporção das classes (%)
Corpo estranho	635	15,8
Inteiro	1.696	42,3
Padrão	1.544	38,5
Película	97	2,4
Ponto Preto	41	1,0
Total	4.013	100

Tabela 1 – Distribuição das classes da base de dados CornGrain

Fonte: Autoria própria

4.4 Transformação

As imagens foram exportadas no formato YOLO com dimensão de 1280x1280 e, em seguida, aplicadas técnicas de aumento de dados e transformações utilizando a biblioteca Albumentations² (Figura 10).

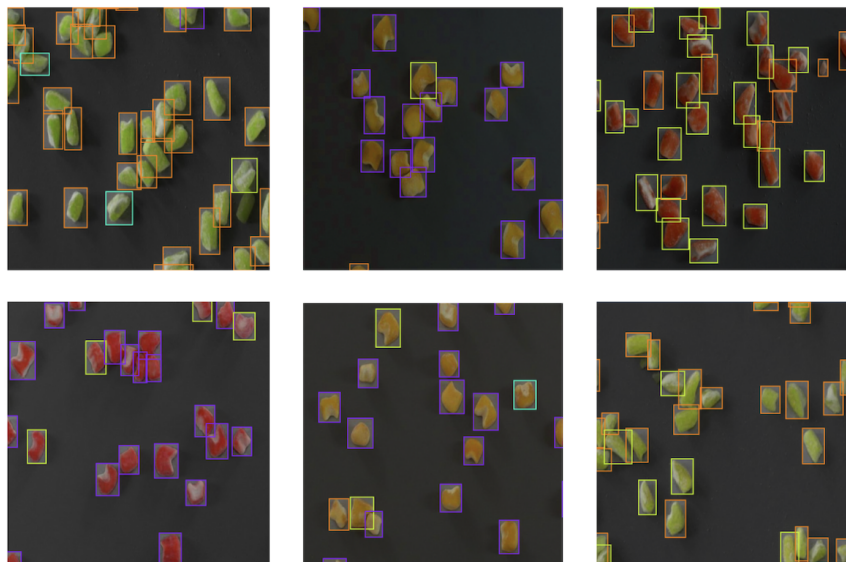


Figura 10 – Exemplo das transformações aplicadas nas imagens

Fonte: Autoria própria

A aumento de dados tem o objetivo de elevar a quantidade de exemplos rotulados. Esse processo engloba uma variedade de métodos que aperfeiçoam a amplitude e qualidade

² Biblioteca Albumentations, disponível em: <https://github.com/albumentations-team/albumentations>. Acessado em: 20/04/2023

do conjunto de dados de treinamento, permitindo a construção de modelos mais eficazes (SHORTEN C., 2019).

O uso das técnicas apresentadas na Figura 11 objetiva evitar o sobreajustamento (*overfitting*) dos modelos. O sobreajustamento ocorre quando o modelo se ajusta e tem ótimos resultados nos dados de treinamento, porém um desempenho ruim na partição de teste, ou seja, o modelo treinado não tem boa capacidade de generalização (YING, 2019).

Como a qualidade das imagens era elevada e os detalhes dos grãos de milho, muito sutis, foi utilizada a estratégia de recortar de forma aleatória as imagens no tamanho de 640 por 640 pixels (640x640px) múltiplas vezes para, ao mesmo tempo, aumentar o número de dados artificialmente e, também, dar ênfase nos detalhes sem perder a qualidade por compressão. No caso, foi optado por realizar as transformações e o recorte aleatório dez vezes em toda a base de dados.

Além do recorte aleatório (*RandomCrop*) de 640px, as seguintes técnicas de aumento de dados foram utilizadas: *HueSaturationValue*, *RandomBrightnessContrast* e *HorizontalFlip*, ilustradas na Figura 11 e brevemente descritas a seguir.



Figura 11 – Exemplos de técnicas de aumento de dados

4.4.1 HueSaturationValue

Altera, de forma aleatória, a matiz ou cor e a saturação da imagem. Essa transformação foi realizada dez vezes, porém, foi limitado a 50% da base de dados (BUSLAEV *et al.*, 2020).

4.4.2 RandomBrigtnessContrast

Realiza a alteração, de forma aleatória, do brilho e do contraste da imagem. Porém, foi limitada a aplicação dessa técnica a 20% na base de dados (BUSLAEV *et al.*, 2020).

4.4.3 HorizontalFlip

Gira a imagem horizontalmente em torno do eixo y . Dessa forma, traz uma variabilidade de posições diferentes das imagens. Optou-se por deixar a probabilidade de aplicação dessa técnica em 50% na base de dados (BUSLAEV *et al.*, 2020).

Transformações como as aplicadas na base de dados se mostram eficazes na aumento de dados, conforme demonstrado por Shorten C. (2019). Existem transformações mais complexas, entretanto, não foram empregadas neste trabalho devido às características inerentes ao problema a ser solucionado. Transformações mais complexas acarretam, muitas vezes, em alterações nos atributos que diferenciam grãos de milho da categoria padrão para as anomalias potencialmente existentes na base de dados, como por exemplo, a adição de ruído ou aberração de ISO. Entretanto, tais transformações podem intervir na detecção de pontos pretos do grão de milho.

A Tabela 2 compara a base de dados original e após as transformações usando as técnicas de aumento de dados descritas:

Classe	Base original	(%)	Base aumentada	(%)
Corpo estranho	635	15,8	2.062	15,2
Inteiro	1.696	42,3	5.563	40,8
Padrão	1.544	38,5	5.503	40,4
Película	97	2,4	355	2,6
Ponto Preto	41	1,0	141	1,0
Total	4.013	100	13.624	100

Tabela 2 – Comparação da base de dados original e após a aumento dos dados, com as respectivas proporções das classes (%)

Fonte: Autoria própria

Após o uso de técnicas de aumento de dados, o número de imagens multiplicou por dez e passou para 1.140, porém, após a remoção de imagens que não continham nenhum grão de milho, a base de dados final ficou com 1.036 imagens.

5 EXPERIMENTOS E RESULTADOS

Foram selecionados três métodos no estado-da-arte para serem treinados com a base de dados CornGrain: Faster R-CNN Resnet 152 (REN *et al.*, 2015) , YOLO v8 (ULTRALYTICS, 2023) e EfficientDet v2 (TAN; PANG; LE, 2020). A razão para a escolha desses modelos foi a seguinte: a rede Faster R-CNN Resnet 152 tem arquitetura mais tradicional e robusta, pesada e com muitos atributos em suas camadas; a rede YOLO v8, por outro lado, tem a premissa de ser extremamente leve, sem muita perda de precisão; e a rede EfficientDet v2 tem uma arquitetura intermediária, mas que também oferece ganhos de performance.

5.1 Treinamento dos modelos

Para treinar os modelos, a base de dados foi dividida com a proporção 70/30 entre treinamento e validação. Assim, para o treinamento constam 726 imagens e para validação, 310.

A Tabela 3 mostra a distribuição dos dados entre as partições de treinamento e validação:

Classe	Treinamento	Validação
Corpo estranho	1.484	578
Inteiro	3.867	1.696
Padrão	3.808	1.695
Película	237	118
Ponto Preto	106	35
Total	9.502	4.122

Tabela 3 – Distribuição das amostras: conjunto de treino e validação

Fonte: Autoria própria

Os hiperparâmetros empregados para treinar os modelos são especificados na Tabela 4:

Modelo	<i>Batch size</i>	Quantidade de Épocas	<i>Patience</i>
EfficientDet v2	4	200.000	-
Faster R-CNN Resnet 152	4	200.000	-
YOLO v8	64	1.000	100

Tabela 4 – Hiperparâmetros dos modelos configurados empiricamente

Fonte: Autoria própria

Complementando os hiperparâmetros, o *Learning rate* utilizado para o modelo EfficientDet v2 foi de $8e-2$, diminuindo 0,001 a cada 2.500 épocas e, de 0,04, diminuindo 0,01333 a cada 25.000 épocas para o Faster R-CNN Resnet 152. Em relação ao ajuste fino (*fine tune*), os modelos EfficientDet v2 e Faster R-CNN Resnet 152 foram pré-treinados na base de dados COCO¹ e feito o ajuste fino para a base de dados CornGrain.

Os treinamentos foram realizados no Google Colaboratory² usando uma GPU Tesla T4 de 16 GB VRAM baseada na arquitetura Turing, voltada para a aceleração da inferência nos modelos de aprendizado profundo.

5.2 Resultados

Foram computadas as seguintes métricas de avaliação: acurácia, sensibilidade (*recall*) e F1-Score. A acurácia indica o desempenho geral do modelo, demonstrando quanto ele é confiável em classificar corretamente as classes conhecidas, ou seja, a proporção de classificações corretas dentre todos os dados classificados. Já a sensibilidade indica a capacidade do modelo de classificar com sucesso a classe positiva. A F1-Score calcula a média harmônica com base no resultado das métricas: precisão e sensibilidade.

As Tabelas 5, 6 e 7 resumizam os resultados de acurácia, sensibilidade e F1-Score obtidos pelos modelos, respectivamente:

¹ Base de dados COCO, disponível em: <https://cocodataset.org>. Acessado em: 28/04/2023

² Google Colaboratory, disponível em: <https://colab.research.google.com/>. Acessado em: 20/03/2023

Classe	EfficientDet v2	Faster R-CNN Resnet 152	YOLO v8
Geral	0,926	0,960	0,957
Corpo estranho	0,933	0,994	0,993
Inteiro	0,952	0,983	0,959
Padrão	0,922	0,942	0,948
Película	0,926	0,942	0,948
Ponto Preto	0,898	0,939	0,937

Tabela 5 – Acurácia obtida pelos modelos na classificação dos exemplos de cada classe

Fonte: Autoria própria

Classe	EfficientDet v2	Faster R-CNN Resnet 152	YOLO v8
Geral	0,902	0,918	0,909
Corpo estranho	0,987	0,993	0,991
Inteiro	0,910	0,934	0,929
Padrão	0,872	0,916	0,868
Película	0,898	0,900	0,902
Ponto Preto	0,841	0,873	0,954

Tabela 6 – Sensibilidade obtida pelos modelos na classificação dos exemplos de cada classe

Fonte: Autoria própria

Classe	EfficientDet v2	Faster R-CNN Resnet 152	YOLO v8
Geral	0,914	0,939	0,932
Corpo estranho	0,959	0,993	0,992
Inteiro	0,931	0,958	0,944
Padrão	0,896	0,916	0,906
Película	0,912	0,921	0,924
Ponto Preto	0,869	0,905	0,894

Tabela 7 – F1-Score obtida pelos modelos na classificação dos exemplos de cada classe

Fonte: Autoria própria

A Tabela 8 mostra um comparativo da quantidade de parâmetros, do tempo de treinamento e de inferência de cada rede:

Rede	Qtde de parâmetros	Tempo de treinamento	Tempo de inferência (ms)
EfficientDet v2	105.377.816	12h 00 min	9.8
Faster R-CNN Resnet	60.192.808	12h 00 min	38.3
YOLO v8	3.011.823	3h 45min	7.64

Tabela 8 – Comparação da quantidade de parâmetros e do tempo de treinamento e de inferência

Fonte: Autoria própria

A rede YOLO v8 é uma rede leve e multifacetada, podendo ser utilizada em diversos dispositivos, GPUs comuns, em nuvem, além de ser rápida e oferecer bom desempenho. Esta rede apresentou um resultado geral de 0,957 de acurácia, apenas 0,003 a menos que a rede Faster R-CNN que obteve o melhor resultado. Comparando com a EfficientDet v2, ela obteve resultado superior de 0,031. Outro ponto a considerar é o tempo de inferência que foi de 7,64 milissegundos, com melhor resultado na comparação, superando as demais redes em: 30,66 milissegundos (Faster R-CNN) e 2,16 milissegundos (EfficientDet v2). Observando a acurácia na predição das classes, a YOLO v8 obteve melhor desempenho em relação às outras duas redes nas classes padrão e película, com acurácia de 0,948 para ambas. Analisando os valores de sensibilidade e F1-Score, ela alcançou o melhor resultado na classe película.

A rede Faster R-CNN é robusta, com uma arquitetura mais tradicional e usa como *backbone* a Resnet 152. Ela apresentou o melhor resultado geral, comparando com a YOLO v8 e a EfficientDet v2, com 0,960 de acurácia geral na classificação dos grãos de milho. Em comparativo as demais redes treinadas, teve desempenho 0,034 superior que a EfficientDet v2 e de 0,003, em relação a YOLO v8. Na predição por classes, foi a rede que obteve o melhor desempenho na assertividade da maioria das classes, com melhor performance nas classes: corpo estranho, inteiro e ponto preto. Considerando as métricas sensibilidade e F1-Score, o desempenho foi superior aos demais modelos, exceto na classe película. Em relação ao tempo de inferência, foi a rede mais demorada, consumindo 38,3 milissegundos, em média.

A rede EfficientDet v2 implementa uma nova técnica de escalonamento e emprega como *backbone* a EfficientNet. No comparativo com as demais redes testadas, obteve resultado inferior, apresentando 0,926 de acurácia na classificação geral. Em relação a predição por classes, não superou as demais redes em nenhuma categoria. Porém, no tempo de inferência, foi mais rápida que a Faster R-CNN, com diferença de 28,5 milissegundos.

Outro ponto relevante a ser comparado é o tempo de treinamento: a YOLO v8 apresentou um tempo de treinamento muito menor, se comparada com os outros dois modelos. O treinamento da YOLO v8 ocorreu em aproximadamente três horas e quarenta

e cinco minutos e dos modelos: EfficientDet v2 e Faster R-CNN Resnet 152 demoraram ao menos doze horas cada (Tabela 8). Considerando que os treinamentos foram realizados no Google Colaboratory que disponibiliza uma GPU Tesla T4 de 16 GB de VRAM baseada na arquitetura Turing.

Adicionalmente, deve ser considerado a quantidade de parâmetros de cada modelo: a EfficientDet v2 conta com 105.377.816 parâmetros, a rede Faster R-CNN Resnet 152 com 60.192.808 e a YOLO v8, com 3.011.823.

Assim, a quantidade de parâmetros da rede YOLO v8 é aproximadamente trinta e cinco vezes menor que da EfficientDet v2 e cerca de vinte vezes menor que da Faster R-CNN. Comparando o tempo de treinamento da EfficientDet v2 e da Faster R-CNN Resnet 152 foi quatro vezes maior que o da YOLO v8. Além disso, o tempo de inferência da YOLO v8 foi, em média, cinco vezes mais rápido que da Faster R-CNN Resnet 152 (Tabela 8).

Analisando o resultado num contexto mais amplo, é conclusivo que o modelo YOLO v8 é o que ofereceu o melhor balanceamento entre poder preditivo e complexidade do modelo. Ele obteve praticamente o mesmo desempenho preditivo da Faster R-CNN Resnet 152, demandando o menor tempo de treinamento e de inferência dentre todos os modelos comparados.

6 CONCLUSÃO

O Brasil, como um dos maiores produtores de grãos do mundo desponta como um forte candidato que terá benefícios com a crescente demanda global por alimentos ocasionado, entre outros fatores, pelo aumento da população mundial, assim é necessário pensar em estratégias para garantir o fornecimento de alimentos para todos. Porém, em geral, não é possível aumentar a área plantada, a solução então é investir em tecnologias que auxiliem no aumento da produtividade.

O milho é um dos cultivos mais antigos do mundo e, atualmente, um dos alimentos mais consumido. Sendo um cereal multifacetado, tem aproximadamente 3500 aplicações, da alimentação de pessoas e animais, até a indústria, na produção de combustível, de cerveja entre muitos outros usos. Porém, principalmente para os grãos destinados à alimentação, é necessário passar por um processo de classificação para garantir a segurança do consumo.

O processo de classificação dos grãos de milho é normalmente realizado de forma manual, por um especialista. Automatizar esse processo poderá trazer ganhos, como celeridade ao processo, maior assertividade e objetividade na classificação e aumento da quantidade de amostras classificadas. Esses benefícios trazem um impacto positivo em toda cadeia produtiva, pois de forma otimizada, o milho, conforme sua classificação, chega mais rápido ao seu destino final.

Este trabalho apresentou uma base de dados de imagens rotuladas de grãos de milho, chamada CornGrain Dataset¹ e comparou três modelos no estado-da-arte para a predição automática de grãos. A base de dados consiste em 1.036 imagens, com treze mil seiscentos e vinte e quatro (13.624) grãos de milho tendo a seguinte distribuição: 15,2% corpo estranho, 40,8% inteiro, 40,4% padrão, 2,6% película e 1,0% ponto preto.

Os resultados indicaram que a rede YOLO v8 é a que oferece melhor balanceamento entre poder preditivo e complexidade do modelo. O desempenho preditivo obtido por esse modelo foi praticamente igual ao da Faster R-CNN Resnet 152 (a rede com melhor resultado absoluto), mas demandando quatro vezes menos tempo de treinamento e cinco vezes menos tempo de inferência. Além de ter vinte vezes menos parâmetros.

Com esse resultado, demonstra-se que é factível que o processo de classificação de grãos de milho seja automatizado, eliminando assim o erro humano do processo. Adicionalmente, além do ganho em qualidade da classificação, também ganha-se em tempo de execução, possibilitando um aumento do número de amostras analisadas.

Trabalhos futuros podem explorar a expansão da base de dados com a inclusão

¹ CornGrain Dataset, disponível em: <https://huggingface.co/datasets/AdrianaCasadei/CornGrain>

de novas classes de milho. Nesse contexto, pode ser interessante o balanceamento das classes por meio da captura de novas imagens, já que atualmente a base de dados está desbalanceada, representando a distribuição natural das amostras. Por fim, é sugerido a implementação de um sistema automático de classificação de grãos e a otimização adicional do modelo treinado.

REFERÊNCIAS

- ALTUNTAS, Y.; KOCAMAZ, A. F.; YEROGLU, C. Identification of apricot varieties using leaf characteristics and knn classifier. **2019 International Artificial Intelligence and Data Processing Symposium (IDAP)**, p. 1–6, 2019.
- ARORA, B. *et al.* Rice grain classification using image processing machine learning techniques. **2020 International Conference on Inventive Computation Technologies (ICICT)**, p. 205–208, 2020.
- BORTH, M. R. *et al.* **A Visão Computacional no Agronegócio: Aplicações e Direcionamentos**. 2014.
- BRASIL. **LEI No 9.972, DE 25 DE MAIO DE 2000**. 2000. Acesso em: 16/03/2023. Available at: https://www.planalto.gov.br/ccivil_03/leis/19972.htm.
- BROSNAN, T.; SUN, D. W. Inspection and grading of agricultural and food products by computer vision systems—a review. **Computers and Electronics in Agriculture**, v. 36, n. 2, p. 193–213, 2002. ISSN 0168-1699. Available at: <https://www.sciencedirect.com/science/article/pii/S0168169902001011>.
- BUSLAEV, A. *et al.* Albumentations: Fast and flexible image augmentations. **MDPI**, 2020. Available at: <https://www.mdpi.com/2078-2489/11/2/125>.
- CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Instruções para Amostragem de Grãos**. Brasília, 2015. (ISSN: 2359-5612). Available at: https://www.conab.gov.br/outras-publicacoes/item/download/2524_9df63280b3c06ef33808100e23c8089f.
- CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira**. Brasília, 2023. Available at: <https://www.conab.gov.br/ultimas-noticias/4971-producao-de-graos-esta-estimada-em-312-5-milhoes-de-toneladas-na-safra-2022-23>.
- CONCI, A.; AZEVEDO, E.; LETA, F. **Computação Gráfica - Teoria e Prática**. [S.l.: s.n.], 2007. V2. Available at: <https://altabooks.com.br/produto/computacao-grafica-vol-2/>.
- DUARTE, J. O.; MATTOSO, M. J.; GARCIA, J. C. Importância socioeconômica. **Embrapa**, v. 1, n. 1, p. 71–97, 2008.
- EMBRAPA. **Uma história de sucesso sa civilização - Milho a riqueza dourada**. Brasília, 2018. Available at: <https://edcentaurus.com.br/agranja/edicao/829/materia/8972>.
- EMBRAPA. **Visão 2030 : o futuro da agricultura brasileira**. Brasília, 2018. Available at: <https://www.embrapa.br/documents/10180/9543845/Vis%C3%A3o+2030+-+o+futuro+da+agricultura+brasileira/2a9a0f27-0ead-991a-8cbf-af8e89d62829?version=1.1>.
- EMPRAPA. Empresa Brasileira de Pesquisa Agropecuária. **O Agro no Brasil e no mundo: uma síntese do período de 2000 a 2020**. Embrapa, 2020.
- EMPRAPA. Empresa Brasileira de Pesquisa Agropecuária. **Qualidade e Classificação de Grãos de Milho Colhidos e Armazenados nas Safras 2014/2015, 2015/2016 e 2016/2017**. Embrapa, 2020.

FACELI, K. *et al.* **Inteligência Artificial: uma abordagem de aprendizado de máquina**. 2. ed. [S.l.: s.n.]: LTC, 2004.

FAO: FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **World Food Situations**. Food and Agriculture Organization of the United Nations, 2023.

GARCIA, M.; CANDELO-BECERRA, J.; HOYOS, F. E. Quality and defect inspection of green coffee beans using a computer vision system. **MDPI - ppl. Sci.** **2019**, **9**, **4195**, p. 18, 10 2019.

HE, k. *et al.* Mask r-cnn. 2018. Available at: <https://arxiv.org/abs/1703.06870>.

HE, K. *et al.* Deep residual learning for image recognition. **Arxiv - Cornell University**, p. 12, 12 2015.

HOWARD, A. *et al.* Searching for mobilenetv3. **Arxiv - Cornell Universitys**, v. 25, 11 2019.

HOWARD, A. G. *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. **Arxiv - Cornell Universitys**, v. 25, 04 2017.

IPEA. **Agricultura e indústria no Brasil: inovação e competitividade**. Brasília, 2017. Available at: <https://repositorio.ipea.gov.br/handle/11058/7682>.

IPEA. **Expansão da produção agrícola, novas tecnologias de produção, aumento de produtividade e o desnível tecnológico no meio rural**. Brasília, 2022. Available at: https://repositorio.ipea.gov.br/bitstream/11058/11187/1/td_2765.pdf.

KOKLU, M.; OZKAN, I. A. Multiclass classification of dry beans using computer vision and machine learning techniques. **Computers and Electronics in Agriculture**, v. 174, p. 105507, 2020. ISSN 0168-1699. Available at: <https://www.sciencedirect.com/science/article/pii/S0168169919311573>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. Imagenet classification with deep convolutional neural networks. **Neural Information Processing Systems**, v. 25, 01 2012.

LECUN, Y. *et al.* Gradient-based learning applied to document recognition. **IEEE**, v. 33, n. 12, p. 6999–7019, 1998.

LI, X. *et al.* Corn classification system based on computer vision. **Symmetry**, v. 11, n. 4, 2019. ISSN 2073-8994. Available at: <https://www.mdpi.com/2073-8994/11/4/591>.

LI, Z. *et al.* A survey of convolutional neural networks: Analysis, applications, and prospects. **IEEE Transactions on Neural Networks and Learning Systems**, n. 46, 2021.

MAHESWARAN, S. *et al.* Development of machine learning based grain classification and sorting with machine vision approach for eco-friendly environment. **Journal of Green Engineering (JGE)**, v. 10, 2020. Available at: <https://iopscience.iop.org/article/10.1088/1742-6596/803/1/012177>.

- NARANJO-TORRES, J. *et al.* A review of convolutional neural network applied to fruit image processing. **Applied Sciences**, 2020. Available at: <https://www.mdpi.com/2076-3417/10/10/3443#>.
- PAES, M. C. D. Aspectos físicos, químicos e tecnológicos do grão de milho. **Embrapa**, v. 1, n. 6, p. 0-0, 2006.
- PESTANA, D. E. *et al.* Uso de molibdênio na cultura do milho. **Anais do 12º Encontro Científico Cultural Interinstitucional**, v. 1, n. 1, p. 5, 2014.
- PROCÓPIO, G. M. P. R. **A importância do agronegócio na economia brasileira e seu aumento de produtividade nas últimas décadas**. PUC-Rio, 2022. Available at: https://www.econ.pucio.br/uploads/adm/trabalhos/files/Gabriel_Martins_Penna_Rossi_Procopio_Mono_22.2.pdf.
- REDMON, J. *et al.* You only look once: Unified, real-time object detection. **Arxiv - Cornell University**, 2016.
- REN, S. *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. **Advances in Neural Information Processing Systems (NIPS)**, 2015. Available at: <https://arxiv.org/abs/1506.01497>.
- SANDLER, M. *et al.* Mobilenetv2: Inverted residuals and linear bottlenecks. **Arxiv - Cornell University**, v. 25, 01 2018.
- SEIXAS, M. A. **Segurança alimentar pós-Covid-19: megatendências dos sistemas alimentares globais**. Aprenda fácil - Embrapa, 2021. Available at: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/215506/1/Seguranca-alimentar-pos-Covid19.pdf>.
- SENAR. **Grãos: classificação de soja e milho**. Serviço Nacional de Aprendizagem Rural, 2017. Available at: <https://www.cnabrazil.org.br/assets/arquivos/178-GR%C3%83OS.pdf>.
- SHORTEN C., K. T. A survey on image data augmentation for deep learning. **Journal of Physics: Conference Series**, 2019. Available at: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0#citeas>.
- SILVA, J. S. **Secagem e armazenamento de produtos agrícolas**. Aprenda fácil - Embrapa, 2008. Available at: https://issuu.com/juarezufv/docs/cap_tulo_17_2013.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **Arxiv - Cornell University**, v. 25, 01 2015.
- SZEGEDY, C. *et al.* Going deeper with convolutions. **2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, 2015. Available at: <https://arxiv.org/abs/1409.4842>.
- TAN, M.; LE, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. **Arxiv - Cornell University**, p. 1905.11946, 9 2020.
- TAN, M.; PANG, R.; LE, Q. V. Efficientdet: Scalable and efficient object detection. **arXiv**, 2020. Available at: <https://arxiv.org/abs/1911.09070>.
- ULTRALYTICS. **YOLO v8**. 2023. Available at: <https://docs.ultralytics.com/>.

VELESACA, H. *et al.* Deep learning based corn kernel classification. **IEEE - CVPR**, 2020. Available at: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w5/Velesaca_Deep_Learning_Based_Corn_Kernel_Classification_CVPRW_2020_paper.pdf.

VELESACA, H. O. *et al.* Computer vision based food grain classification: A comprehensive survey. **Computers and Electronics in Agriculture**, v. 187, p. 106287, 2021. ISSN 0168-1699. Available at: <https://www.sciencedirect.com/science/article/pii/S0168169921003045>.

VLASOV, A. V.; FADEEV, A. S. A machine learning approach for grain crop's seed classification in purifying separation. **Journal of Physics: Conference Series, Volume 803, International Conference on Information Technologies in Business and Industry 2016 21–26 September 2016, Tomsk, Russian Federation**, 2017. Available at: <https://iopscience.iop.org/article/10.1088/1742-6596/803/1/012177>.

WANG, C. Y.; BOCHKOVSKIY, A.; LIAO, H. Y. M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. **Arxiv - Cornell Universitys**, 2022.

WANG, X.; PALIWAL, K. Feature extraction and dimensionality reduction algorithms and their applications in vowel recognition. **Pattern Recognition**, v. 36, p. 2429–2439, 10 2003.

WORLD ECONOMIC FORUM: WEFORUM. **World population just passed 8 billion. Here's what it means**. World Economic Forum, 2022.

YING, X. An overview of overfitting and its solutions. **Journal of Physics: Conference Series**, 2019. Available at: <https://iopscience.iop.org/article/10.1088/1742-6596/1168/2/022022/pdf>.

ZEILER, M. D.; FERGUS, R. Visualizing and understanding convolutional networks. *In*: FLEET, D. *et al.* (ed.). **Computer Vision – ECCV 2014**. Cham: Springer International Publishing, 2014. p. 818–833. ISBN 978-3-319-10590-1.