



Banco de imagens de grãos de milho *Zea mays* L. amarela

Alberto Yoshihiro Nakano¹, Diego Ferreira¹, Felipe Walter Dafico Pfrimer¹

¹Universidade Tecnológica Federal do Paraná, Toledo-PR, Brasil (nakano@utfpr.edu.br)

Resumo: Banco de dados tornaram-se importantes para o desenvolvimento de técnicas de detecção, classificação e reconhecimento baseados em aprendizado de máquina (*machine learning*) envolvendo principalmente visão computacional (*computer vision*). Neste trabalho é apresentado detalhes adicionais de um pequeno banco de imagens de grãos de milho da variedade *Zea mays* L. amarela criado inicialmente para estudos e classificação de grãos por métodos estatísticos. Espera-se que a divulgação destes dados estimule o estudo de técnicas de classificação de grãos baseadas em inteligência artificial (IA).

Palavras-chave: Milho; banco de dados; *machine learning*; processamento digital de imagens; classificação.

INTRODUÇÃO

Os avanços tecnológicos aplicados ao agronegócio permitiram ao Brasil se destacar no cenário mundial de produção de grãos. A agricultura de precisão tem auxiliado no aumento da produtividade no campo, no retorno econômico ao produtor e, principalmente, na redução de impactos ambientais. De uma forma geral, a agricultura de precisão (AVILA et al., 2023) envolve a integração de sensores, georreferenciamento e visão computacional permitindo ações localizadas, seja no plantio de sementes, na aplicação de fertilizantes, ou em etapas do desenvolvimento da cultura até a colheita.

Pouco conhecido é o processo posterior à colheita que ocorre nos silos e armazéns envolvendo o processo de classificação no recebimento, na armazenagem e no transbordo de grãos. Para dar continuidade a este trabalho, o grão escolhido para este trabalho foi o milho da variedade *Zea mays* L. amarela por ser a variedade cultivada mais comum em território brasileiro. Para o milho *Zea mays* L. a Instrução Normativa 60/2011 do Ministério da Agricultura, Pecuária e Abastecimento (MAPA, 2011) estabelece um regulamento técnico definindo o padrão oficial de classificação do milho, considerando seus requisitos de identidade e qualidade, a amostragem, o modo de apresentação e a marcação ou rotulagem, nos aspectos referentes à classificação do produto. Em parte do processo técnico de classificação, há necessidade de mão-de-obra especializada que demanda principalmente inspeção visual (SILVA; FLORES, 2022) que podem ser comprometidos por fadiga humana. Assim, uma automatização do processo reduziria problemas, vícios e a subjetividade inerente a análise humana. Porém, para realizar tal automatização são necessários dados e o

acompanhamento da realização de todas as etapas pelos técnicos responsáveis. Uma vez que o processo é compreendido há a possibilidade de implementação de soluções que resolvam em parte ou totalmente o problema analisado.

Neste contexto, os dados podem se estender aos provenientes de análises estatísticas, os obtidos experimentalmente com o uso de instrumentação adequada (por exemplo de umidade), e os visuais obtidos por câmera. Este último se tornou especialmente atrativo pela redução no custo de produção de câmeras digitais presentes em *smartphones* e em câmeras de segurança. Assim, o emprego de imagens obtidas de câmeras tem a possibilidade de substituir a avaliação visual técnica suscetível de erro.

É importante destacar que o banco de imagens descrito neste trabalho não aborda todos os defeitos definidos na IN 60/2011, pois se trata de um conjunto simplificado de dados para atender o seu intuito inicial que é o uso em estudos acadêmicos. Um banco de dados mais completo e complexo pode ser criado a partir das diretrizes estabelecidas neste trabalho. Destaca-se também a importância de visitas às empresas graneleiras para: entender o atual nível de automatização do processo de classificação; a situação real dos grãos em silos e armazéns graneleiros; as condições das instalações de recebimento, armazenamento e transbordo de grãos; e ciência de possíveis atualizações de exigências adicionais do regulamento técnico do MAPA.

Deste modo, considerando-se as informações e o contexto apresentados um pequeno banco de dados referente a grão de milho da espécie *Zea mays* L. foi criado. O banco de dados de imagens foi empregado no estudo acadêmico (FERREIRA, 2023a) fo-

cando no problema de classificação de grãos por análise de histogramas para definição de limiares de decisão. Dessa forma, o objetivo deste trabalho é descrever a estrutura, o processo de aquisição, a rotulação e as limitações de um banco de imagens de grãos de milho amarelo, disponibilizado publicamente para estudos acadêmicos. Detalhes descritivos deste trabalho complementam informações constantes no repositório de dados do GitHub (FERREIRA, 2023b). Espera-se que as informações apresentadas aqui em conjunto com os dados de imagens permitam, com limitações, o seu empregado no treinamento de modelos de aprendizado de máquina (*machine learning*) como em sistemas YOLO (REDMON et al., 2016) para melhorar o processo de automatização da classificação.

MATERIAL E MÉTODOS

A criação de um banco de dados exige cuidados para que o usuário final possa compreender e empregar as informações apresentadas de forma correta. Outro fator importante é permitir a repetibilidade e a reprodutibilidade dos estudos e experimentos já desenvolvidos fornecendo um direcionamento para estudos futuros. No contexto da classificação de grão segundo a IN 60/2011 do MAPA, considera-se diversas condições do grão como, por exemplo, carunchados, avariados (ardidos, chochos ou imaturos, fermentados, germinados, gessados, mofados), quebrados e a presença de impurezas de matérias estranhas.

Entre as diversas condições em que o grão pode estar, quatro estados foram considerados para o estudo sendo eles:

- sadio: apresenta aspecto amarelado brilhante e sem defeitos;
- germinado: apresenta estado de germinação visível;
- ardido: apresenta escurecimento parcial ou total do grão;
- quebrado: apresenta apenas parte do grão.

Ressalta-se que pelo fato de apenas grãos sadios serem disponíveis no mercado, os defeitos foram simulados. Grãos germinados foram gerados por umedecimento do grão provocando a brotação. Grãos ardidos foram retocados em coloração escura característica deste tipo de estado. Grãos quebrados foram originados por impacto físico até apresentarem aspecto quebradiço ou pedaços de grão. A Figura 1 ilustra os quatro estados considerados na criação do banco de imagens.



Figura 1: Exemplos de grão sadio e grãos com os defeitos simulados (Adaptado de: (FERREIRA, 2023a))

Para a aquisição de imagens, três condições foram consideradas:

- porcentagem (%) de grãos com defeito (germinados, ardido e quebrado) dentro do conjunto de grãos da amostra considerada;
- quantidade de imagens, ou seja, quantidade de registros fotográficos capturados; e
- variação da quantidade de grãos por imagens,

sendo que as informações estão resumidas na Tabela 1. Assim, a porcentagem de grãos com defeito varia de 10 a 30% em intervalos de 5%. A quantidade de imagens capturadas para defeitos de 10 a 30% é de 20 imagens. Para situação sem grãos com defeito (0%) a quantidade varia entre 110 e 125 dependendo da quantidade de grãos consideradas. Por último a quantidade de grãos considerados é de 50 à 100 em intervalos de 10 grãos. A maior quantidade de imagens na situação sem defeito deve-se a necessidade de um maior número de dados para se extrair características estatísticas do grão sem defeito considerando-as como referência.

Definidas as condições experimentais prosseguiu-se para o sistema necessário para aquisição de imagens. A aquisição de imagens foi realizada empregando-se uma placa de desenvolvimento ESP32-CAM e uma câmera OV2640 posicionada na parte superior de uma estrutura metálica em formato cúbico (30 cm × 30 cm × 30 cm) construído por

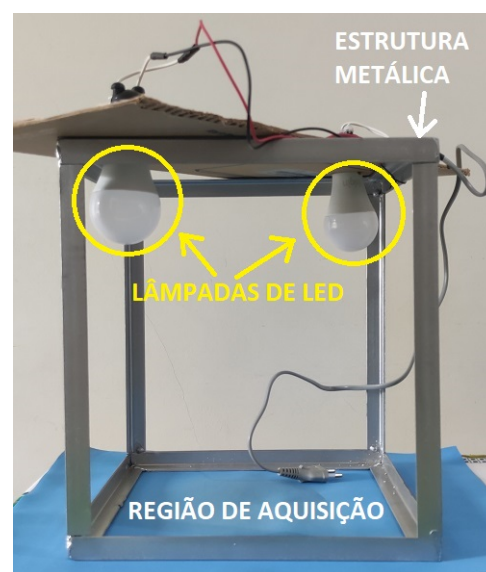
Tabela 1: Quantidade de imagens utilizadas para cada amostra, que contém diferentes percentuais de grãos com defeito e diferentes quantidades de grãos

Percentual de defeitos (%)	Quantidade de imagens					
	50 grãos	60 grãos	70 grãos	80 grãos	90 grãos	100 grãos
0	125	122	118	121	118	110
10	20	20	20	20	20	20
15	20	20	20	20	20	20
20	20	20	20	20	20	20
25	20	20	20	20	20	20
30	20	20	20	20	20	20

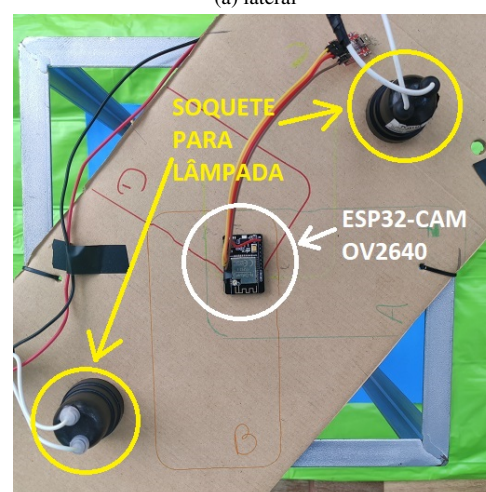
cantoneiras de aço de espessura 2 mm. A placa de desenvolvimento foi escolhida pelo custo de aquisição e facilidade de programação permitindo, caso necessário, a criação de outros sistemas a um custo reduzido. Para auxiliar na aquisição, lâmpadas de LED foram adaptadas para permitir uma iluminação adequada pois o local de captura pode variar, indo desde ambientes controlados como em laboratórios, como em galpões e armazéns fechados possibilitando aquisição *in loco*. As laterais da estrutura são vazadas permitindo a observação do processo de aquisição e da disposição dos grãos de milho. Isto é especialmente importante pois permite ajustes na disposição e quantidade de grãos, assim como uma intervenção sempre que necessário. O sistema é posicionado sobre um fundo azul que permite um contraste com a coloração dos grãos facilitando no processamento das imagens posterior. A coloração do fundo deve destoar das características do grão analisado, assim, se a análise considerasse grãos de milho branco, um fundo mais escuro poderia ser uma melhor alternativa. A Figura 2 apresenta a estrutura metálica desenvolvida com detalhes da posição das lâmpadas de LED, sistema de aquisição e da região de aquisição.

Tabela 2: Descrição dos rótulos.

Rótulo	Descrição
tag01	'a1' possui 30% de grãos com defeitos 'a2' possui 25% de grãos com defeitos 'a3' possui 20% de grãos com defeitos 'a4' possui 15% de grãos com defeitos 'a5' possui 10% de grãos com defeitos 'a6' até 'a11' possui 100% de grãos saudios
tag02	Quantidade de grãos na imagem: 50, 60, 70, 80, 90 e 100
tag03	Número da imagem, para cada percentual de defeito e quantidade de grãos têm-se 20 imagens. Amostras com grãos saudios têm quantidades variadas de imagens.



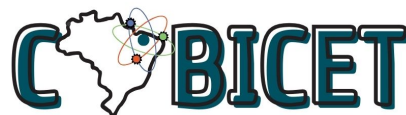
(a) lateral



(b) superior

Figura 2: Sistema de aquisição: (a) Vista lateral; e (b) vista superior (Adaptado de: (FERREIRA, 2023a))

O processo de aquisição pode ser resumido na Figura 3 que se inicia com a configuração do ESP32-CAM. Em seguida, verifica-se a aquisição foi realizada. Esta etapa é necessária pois é possível realizar



aquisições consecutivas. Caso não tenha sido realizada é enviado um comando por comunicação serial para realizar a aquisição e armazenar o dado. O processo é então repetido para o número de aquisições que forem necessárias. Entre as aquisições é possível mudar a disposições e a quantidade de grãos na região delimitada da Figura 2(a).

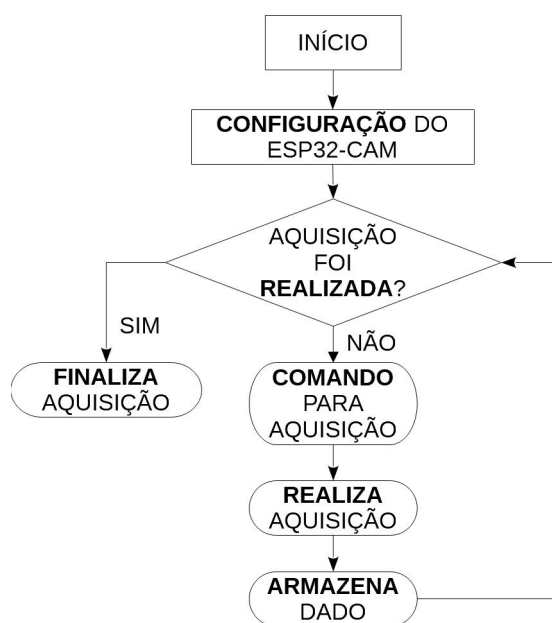


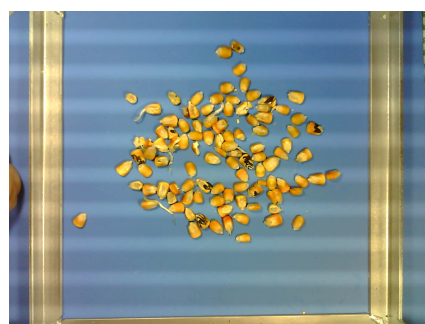
Figura 3: Fluxograma do processo de aquisição de imagens.

Após à aquisição há necessidade da rotulação das imagens para identificação das condições experimentais de aquisição para auxiliar o usuário na interpretação das imagens. Uma falha na identificação dos dados implicará em um viés na interpretação dos resultados. Assim, a Tabela 2 apresenta a rotulação definida para o banco de dados. O rótulo ‘tag01’ representa a porcentagem dos defeitos na amostra, ‘tag02’ a quantidade de grãos na amostra e ‘tag03’ a quantidade de imagens capturadas para compor o banco de dados. Assim, ‘a2_60_20’ representa uma amostra que contém 25% de defeitos em um total de 60 grãos sendo a 20^a captura realizada.

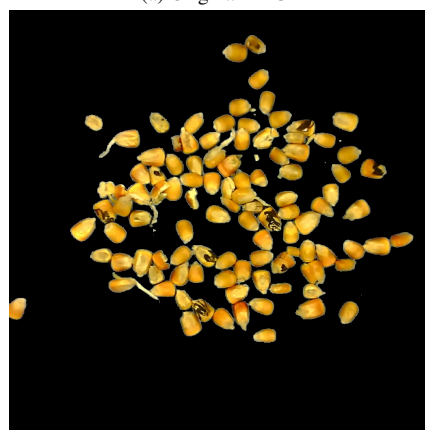
RESULTADOS E DISCUSSÃO

A Figura 4 apresenta a amostra ‘a1_100_1’ considerando-se: (a) a imagem original com 100 grãos capturada pelo ESP32-CAM/OV2640 com resolução de 1600×1200 pixels e 96×96 dpi; (b) a imagem processada em cores no sistema RGB (do inglês *red, green blue* referente as cores primárias vermelho, verde e azul) removendo-se o fundo e outros elementos não relacionados com os grãos em estudo com resolução de 1050×1050 pixels e 96×96 dpi; e

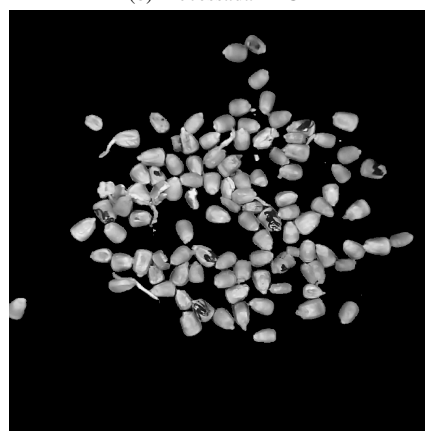
(c) a imagem processada na escala de preto-e-branco (grayscale) com resolução de 1050×1050 pixels e 96×96 dpi, sendo que apenas as imagens processadas são disponibilizadas publicamente no repositório. A disponibilização em formato JPG e em dois sistemas de busca oferecer ao usuário opções de imagens prontas para serem usadas. Observa-se que em algumas análises é interessante a informação de cores enquanto em outras basta a intensidade. O formato empregado dependerá da técnica de classificação estudada.



(a) Original - RGB

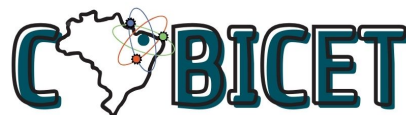


(b) Processada - RGB



(c) Grayscale

Figura 4: Amostra ‘a1_100_1’: (a) imagem original em RGB; (b) imagem processada no sistema de cores RGB; (c) imagem no sistema Grayscale.



A Figura 5 apresenta imagens do banco de dados considerando-se 50 grãos com diversas proporções de grãos com e sem defeito conforme rótulos definidos no banco de dados. É possível observar que as imagens abordam situações de grãos dispersos, concentrados e, em algumas vezes, sobrepostos permitindo uma avaliação mais real do que pode ocorrer no processo real de avaliação de grãos. A Figura 6 apresenta imagens do banco de dados considerando-se 90 grãos. As mesmas observações mencionadas na Figura 5 são válidas. Entretanto, quanto maior a quantidade de grãos maior a possibilidade de ocorrerem grãos sobrepostos.

A estrutura do repositório onde as imagens podem ser acessadas é dividida em 'gray', onde as imagens em intensidade representadas na escala cinza são listadas, e 'rgb', onde as imagens em cores são listadas.

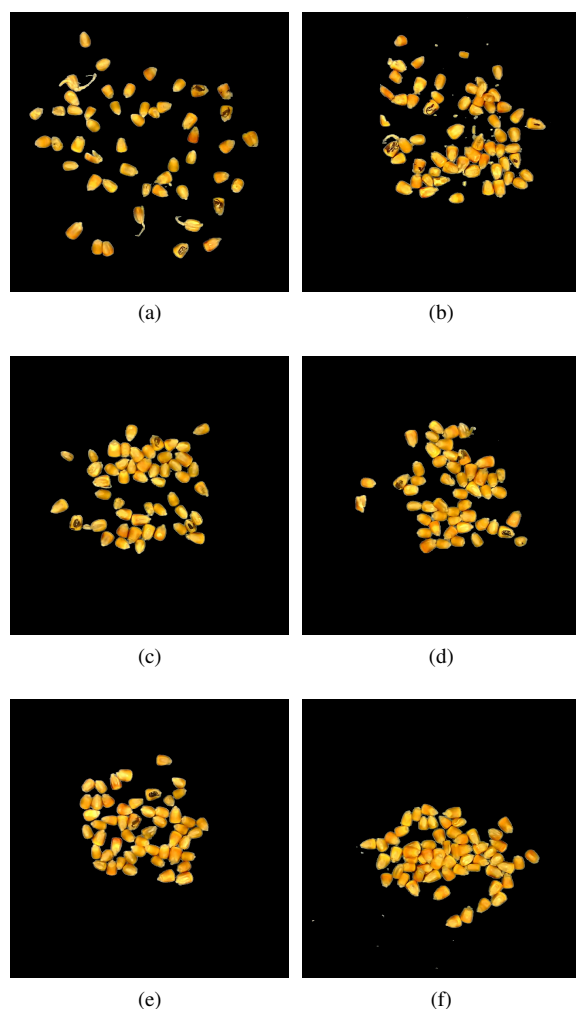


Figura 5: Amostras processadas em cores considerando 50 grãos com e sem defeito: (a) 'a1_50_01'; (b) 'a2_50_01'; (c) 'a3_50_01'; (d) 'a4_50_01'; (e) 'a5_50_01'; (f) 'a6_50_01'.

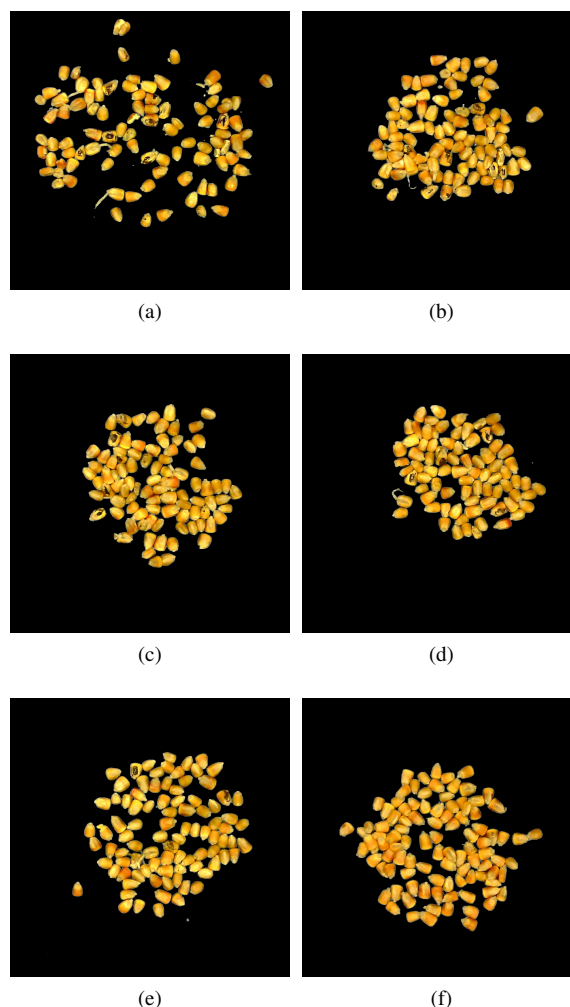


Figura 6: Amostras processadas em cores considerando 90 grãos com e sem defeitos: (a) 'a1_90_01'; (b) 'a2_90_01'; (c) 'a3_90_01'; (d) 'a4_90_01'; (e) 'a5_90_01'; (f) 'a6_90_01'.

CONCLUSÃO

Este trabalho apresentou uma descrição mais detalhada do banco de imagens criada para o processo de classificação de grãos de milho da espécie *Zea mays* L.. O processo de classificação de grãos em silos e armazéns graneleiros é pouco conhecido, mas grande importância para avaliar a qualidade na recepção, armazenagem e posterior envio dos grãos. Como limitação deste trabalho, destaca-se que parte dos defeitos foi simulada em ambiente controlado. Assim, o banco de imagens é adequado para estudos acadêmicos iniciais de segmentação e classificação, mas não substitui bases formadas por amostras reais classificadas por especialistas segundo critérios oficiais. Espera-se que este trabalho possa ser uma referência para o levantamento de outros bancos de imagens e que também possa ser empregado no desenvolvimento de métodos



computacionais de classificação de grãos em geral.

REFERÊNCIAS

AVILA, R. et al. Revisão sistemática de visão computacional para veículos autônomos agrícolas. In: *2023 15th IEEE International Conference on Industry Applications (INDUSCON)*. [S.l.: s.n.], 2023. p. 485–492.

FERREIRA, D. *Classificação de Grãos de Milho da Espécie Zea Mays da Classe Amarela Utilizando Processamento Digital de Imagens*. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Eletrônica) — Universidade Tecnológica Federal do Paraná (UTFPR), Toledo, PR, Brasil, jun. 2023.

FERREIRA, D. *Repositório de imagens para classificação de grãos de milho*. 2023.

<<https://github.com/ferreiradiego/maize-classification/tree/main/img/dataset>>. Acesso em: 13 de junho de 2026.

MAPA. *Instrução Normativa nº 60, de 22 de dezembro de 2011*. 2011. Acesso em: 13 de junho de 2026. Disponível em: <<https://sislegis.agricultura.gov.br/sislegis/action/detalhaAto.do?method=visualizarAtoPortalMapa&chave=1739574738>>.

REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, 2016. p. 779–788.

SILVA, G. Viais de B.; FLORES, F. C. Rot corn grain classification by color and texture analysis. *IEEE Latin America Transactions*, v. 20, n. 2, p. 208–214, 2022.