

## DESENVOLVIMENTO DE ALGORITMO PARA CLASSIFICAÇÃO DE FRUTOS POR MEIO DE REDES NEURAIAS CONVOLUCIONAIS

JOÃO CÉLIO LUNA DE CARVALHO <sup>1</sup>, ANDERSON GOMIDE COSTA <sup>2</sup>, MARCUS  
VINICIUS MORAIS DE OLIVEIRA <sup>3</sup>

<sup>1</sup> Eng. Agrícola e Ambiental, Mestrando, Depto. Engenharia, UFRRJ, Seropédica – RJ, joaoceleio1301@gmail.com

<sup>2</sup> Eng. Agrícola, Prof. Doutor, Depto. De Engenharia, UFRRJ, Seropédica – RJ

<sup>3</sup> Eng. Agrícola, Prof. Doutor, Depto. De Engenharia, UFRRJ, Seropédica – RJ

Apresentado no  
LIII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2024  
6 a 8 de agosto de 2024 – Natal – RN, Brasil

**RESUMO:** O crescimento da população mundial e produção alimentícia implica no aumento da demanda de produtos agrícolas. A classificação e seleção dos produtos de maneira adequada se faz necessária, e a inteligência artificial em conjunto com o processamento de imagens tem sido cada vez mais aplicados, como as Redes Neurais Convolucionais (CNN). O objetivo deste projeto foi desenvolver um algoritmo em linguagem Python, capaz de treinar uma Rede Neural Convolutiva customizada para identificar diferentes tipos de frutos. O banco de dados categórico foi composto por imagens de frutos de morango (137 imagens) e tomate (119 imagens) em diferentes ambientes obtidas na internet. O banco de dados foi dividido em 75% para treinamento e 25% para teste. O algoritmo foi implementado no microcontrolador Raspberry Pi, e a avaliação de desempenho se deu por meio da obtenção da acurácia, precisão, sensibilidade e F1-score da classificação dos dados de teste. O algoritmo se mostrou apto para a classificação dos frutos, com acurácia global de 75,00%. A precisão, sensibilidade e F1-Score para a classe dos tomates foi de 69,70%, 79,31% e 74,19%, respectivamente, enquanto para a classe dos morangos, foi de 80,65%, 71,43% e 75,76%, respectivamente.

**PALAVRAS-CHAVE:** APRENDIZAGEM PROFUNDA, PROCESSAMENTO DE IMAGENS, RASPBERRY PI.

## DEVELOPMENT OF ALGORITHM FOR FRUIT CLASSIFICATION THROUGH CONVOLUTIONAL NEURAL NETWORKS.

**ABSTRACT:** The growth of the world population and food production implies an increased demand for agricultural products. Proper classification and selection of products are necessary, and artificial intelligence, in conjunction with image processing, has been increasingly applied, such as Convolutional Neural Networks (CNNs). The goal of this project was to develop a Python algorithm capable of training a custom Convolutional Neural Network to identify different types of fruits. The categorical database consisted of strawberry (137 images) and tomato (119 images) fruit images from various environments obtained from the internet. The database was split into 75% for training and 25% for testing. The algorithm was implemented on the Raspberry Pi microcontroller, and performance evaluation was based on accuracy, precision, recall, and F1-score for the classification of test data. The algorithm demonstrated proficiency in fruit classification, achieving an overall accuracy of 75,00%. Precision, recall, and F1-score for the tomato class were 69,70%, 79,31%, and 74,19%, respectively, while for the strawberry class, they were 80,65%, 71,43%, and 75,76%, respectively.

**KEYWORDS:** MACHINE LEARNING, IMAGE PROCESSING, RASPBERRY PI.

**INTRODUÇÃO:** O desenvolvimento da mecanização e automação dos processos pós-colheita são impulsionados pelo crescimento populacional (Lima et al, 2019), onde consumidores exigem produtos de boa aparência e alto valor nutricional (Palumbo et al, 2023). Portanto, a utilização de redes neurais em conjunto com o processamento de imagens se torna cada vez mais comum na seleção e classificação de produtos agrícolas. No Brasil, a produção de Tomate e morango é significativa, sendo produzidos em 2023 15 milhões de toneladas (IBGE, 2024) e 183 mil toneladas (FAOSTAT, 2024), respectivamente. A visão computacional, portanto, em conjunto com a inteligência artificial, auxilia a entender os padrões e selecionar os frutos conforme as necessidades do mercado consumidor, trabalhando com os aspectos físicos e químicos de interesse (Oliveira, 2021), sendo mais rápidos e precisos que a classificação humanizada. O avanço da computação possibilitou o surgimento de microcontroladores de placa única e de baixo custo, que podem operar com sistemas de aprendizagem de máquina e inteligência artificial, como as Redes Neurais Convolucionais, capazes de detectar características de interesse em imagens e classificá-las sem a necessidade de intervenção humana (Gu et al., 2018). O desenvolvimento de CNNs próprias, leves e focadas, tem se tornado comum em pesquisas científicas. Portanto, o objetivo desse trabalho foi desenvolver um algoritmo baseado nas Redes Neurais Convolucionais (CNN) e implementada no microcontrolador Raspberry Pi 3B+, para criar um banco de dados com imagens de frutos de morango e tomate e identificar os diferentes tipos de frutos.

**MATERIAL E MÉTODOS:** Inicialmente, foram obtidas pela internet (<https://storage.googleapis.com/openimages/web/index.html>), imagens de frutos de tomate e morango. Para garantir um treinamento robusto para o algoritmo, com maior adaptabilidade as variações externas, aplicou-se como critério, a utilização de imagens com variadas quantidades de frutos, em diferentes estádios de maturação e em diferentes condições de ambientes e iluminações. A escolha destas culturas se deu com o intuito de observar a capacidade da CNN em identificar frutos da mesma cor, mas formas e tamanhos diferentes. Foram utilizadas 119 imagens de tomate e 137 imagens de morango, com 75% dos dados destinados ao treinamento e 25% para testes. Foi desenvolvido um algoritmo na linguagem Python 3.11 e implementado ao microcontrolador Raspberry Pi 3 B+, tendo como objetivo receber as imagens dos frutos de tomate e morango, compila-los em um único arquivo de dataset e realizar o treinamento da Rede Neural Customizada, intitulada OakMoonNet, para a classificação dos frutos. Para implementação do algoritmo foram utilizados os módulos Itertools, Keras, OpenCV2, OS, Matplotlib, Numpy, Pickle, Random, Sklearn, Tensorflow e Time. Ressalta-se que a CNN desenvolvida apresenta a capacidade de ser utilizada para o treinamento de banco de dados de imagens de frutos variados, classificados categoricamente. Inicialmente, o algoritmo pergunta ao usuário se deseja compilar o dataset ou realizar o treinamento da CNN com o dataset compilado. A compilação do dataset criou um único arquivo, adequado para utilização em uma Rede Neural Convolucional, com imagens redimensionadas para a escala 100x100, visando acelerar o processamento dos dados. O dataset foi embaralhado para evitar uma aprendizagem afetada pelo ordenamento dos dados. O tempo de processamento foi salvo em um arquivo de texto. O dataset processado foi carregado, sendo separado na proporção de 75% e 25% para treinamento e teste da rede neural, respectivamente. Posteriormente, a CNN customizada foi treinada, com 50 épocas e batch size 8. O modelo da rede neural convolucional customizada foi desenvolvido para distinguir imagens de frutos de tomate e morango, independentemente de sua quantidade, ambiente ou iluminação da imagem. A arquitetura da CNN (Figura 1) foi composta por seis camadas convolucionais, sendo duas camadas convolucionais com 32 filtros, duas camadas convolucionais com 64 filtros e duas camadas convolucionais com 128 filtros,

utilizando a função de ativação ReLU (Rectified Linear Unit) e filtro de proporção 3x3, sempre acompanhadas por camadas de batch normalization para aumentar a velocidade de processamento. A cada duas camadas convolucionais, houve uma camada de max pooling, com filtro de proporção 2x2, o que possibilitou o aumento da quantidade de filtros sem uma perda de desempenho. Em seguida, apresentou-se uma camada Flatten, uma camada densa de 128 unidades e função de ativação ReLU, uma camada dropout de 20%, para evitar overfitting (modelo ajustado excessivamente aos dados de treinamento, apresentando baixo desempenho para dados externos), e uma camada densa, com 2 unidades e função de ativação softmax. Ao final, o modelo foi compilado com a função loss “sparse categorical crossentropy”, e otimizador ADAM (Adaptative Moment Estimation).

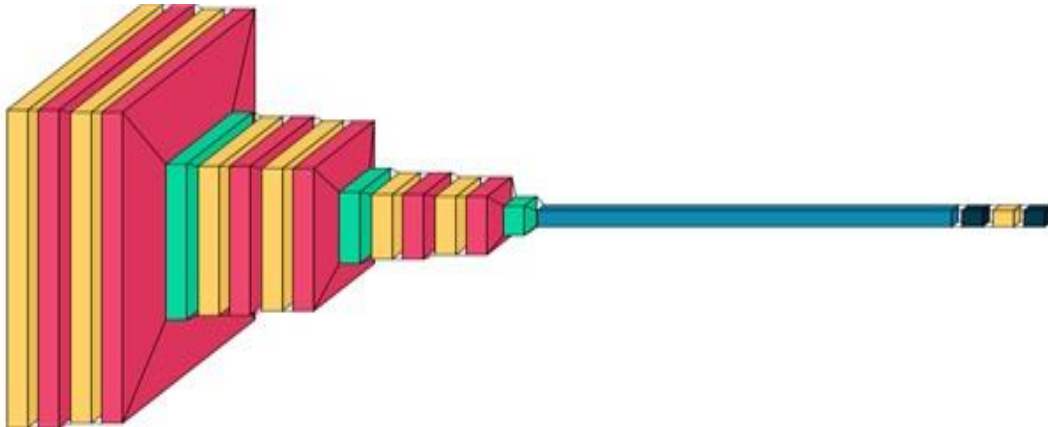


FIGURA 1: Arquitetura da rede neural convolucional customizada. Fonte: Autoria própria

O desempenho foi avaliado por meio do tempo de execução do algoritmo, além matriz de confusão, junto das métricas de acurácia (Equação 1), precisão (Equação 2), sensibilidade (Equação 3) e F1-score (Equação 4), conforme descrito em Chakraborty (2023). A matriz de confusão indicou como cada classe se comporta em relação a classificação, mostrando a quantidade de dados, em relação a uma determinada classe, corretamente classificados como pertencentes a classe (Verdadeiros positivos), dados corretamente classificados como não pertencentes a classe (Verdadeiros negativos), dados falsamente classificados como pertencentes a classe (Falsos positivos) e dados falsamente classificados como não pertencentes a classe (Falsos negativos). A acurácia identificou a quantidade de dados corretamente classificados em todo o modelo. A precisão identificou a quantidade de dados corretamente identificados por classe. A Sensibilidade identificou a tendência de cada classe em receber resultados, sejam eles corretos ou não. O F1-Score foi utilizado com atributo final de classificação do modelo, unindo precisão e sensibilidade.

$$\text{Acurácia} = (\text{VP} - \text{VN}) / (\text{VP} + \text{VN} + \text{FP} + \text{FN}) \quad (1)$$

$$\text{Precisão} = \text{VP} / (\text{VP} + \text{FP}) \quad (2)$$

$$\text{Sensibilidade} = \text{VP} / (\text{VP} + \text{FN}) \quad (3)$$

$$\text{F1-Score} = (2 * \text{Precisão} * \text{Sensibilidade}) / (\text{Precisão} + \text{Sensibilidade}) \quad (4)$$

em que,

- VP – Valores verdadeiro positivos em relação a classe analisada
- VN – Valores verdadeiro negativos em relação a classe analisada
- FP – Valores falso positivos em relação a classe analisada
- FN – Valores falso negativos em relação a classe analisada

Por fim, foi utilizada a função Loss, que quantifica o quão discrepante é a qualificação dos dados estimados em comparação com os valores reais do dataset (Shivashankar et al., 2022). Para isso, utilizam-se algumas funções, das quais calculam a diferença entre o valor estimado e real, aprimorando os pesos no próximo treinamento. Das funções utilizáveis, tem-se a Entropia Cruzada ou Softmax, Função de perda euclidiana (utilizada em problemas de regressão, também conhecida como erro quadrático médio) e Função de perda de Hinge (utilizada em problemas de classificação binária) (Alzubaidi et al., 2021). Nesta pesquisa, foi utilizada a Entropia Cruzada

**RESULTADOS E DISCUSSÃO:** Utilizando o microprocessador Raspberry Pi 3 B+, a compilação do dataset (imagens dos frutos) em um único arquivo composto pelos dados de entrada na rede foi realizada em 26 segundos. O tempo total de treinamento da CNN para classificação dos frutos foi de 1h e 40 minutos. Estes resultados indicam que, devido a capacidade de processamento baixa, o microcontrolador, se mostra aplicável para o processamento de bancos de dados compostos com um número reduzido de imagens. Para processar bancos de dados com grande número de imagens, recomenda-se a utilização de um computador doméstico ou servidores em nuvem, com maior capacidade de processamento. Para avaliar capacidade de identificar os frutos de morango e tomate, a CNN treinada em 75% do banco de imagens foi aplicada nas imagens de teste (25% do banco imagens, o que correspondeu a 35 imagens de morango e 28 imagens de tomate). A partir da matriz de confusão para os dados de teste (Tabela 1) foi observado que 25 imagens de frutos de morango foram classificadas corretamente enquanto 23 imagens de frutos de tomate foram classificadas corretamente. Assim, o algoritmo se mostrou mais tendencioso na detecção dos frutos de morango.

TABELA 1. Matriz de confusão gerada a partir da aplicação da CNN nos dados de teste (35 imagens de frutos de morango e 28 imagens de frutos de tomate).

		Real	
		Tomate	Morango
Previsto	Tomate	23	10
	Morango	6	25

As variações de ambiente, iluminação e condições dos frutos podem influenciar no desempenho do algoritmo. Imagens obtidas em ambientes controlados, como em laboratórios, possibilitam a padronização das disposições experimentais, do uso e configuração da câmera, das condições de ambiente e da iluminação, o que pode refletir no aumento do desempenho do algoritmo. Imagens obtidas em ambientes abertos, mas com destaque ao objeto de interesse, junto de boa iluminação, também podem apresentar bom desempenho. Tsai et al. (2023) propuseram um sistema de classificação dos frutos de tomate quanto a sua maturação, em campo, utilizando a associação da BoTNet, ShuffleNet e GhostNet com a Yolov5m, obtendo uma acurácia de 94%, 95% e 96% respectivamente. Neste experimento, é notável como o fruto de tomate se sobressai na imagem em comparação com o fundo, que em sua maioria é composto por somente folhas. A câmera foi sempre posicionada a 0,5 m do fruto, perpendicularmente, mantendo o padrão de tamanho dos frutos. Pahn et al. (2023) utilizou-se da mesma proposta, seguindo o mesmo padrão de captura das imagens dos frutos de tomate, mas dessa vez, utilizando a Yolov5m sem associações e associada com a ResNet-50, ResNet-101 e EfficientNet-B0, obtendo acurácias na classificação de tomates de 98%, 98%, 97% e 97%, respectivamente. Abekoon et al. (2024) buscou, por meio de dois modelos de CNN próprios, identificar a maturação de frutos de tomate cultivados em estufa. As imagens foram obtidas em ambiente controlado, sendo quatro imagens por fruto, com a câmera posicionada em 90° e 45° em relação ao fruto. Seu primeiro modelo

obteve 99% de acurácia para os dados de treinamento e teste, enquanto o segundo modelo obteve 99% de acurácia para os dados de treinamento e 98% de acurácia para os dados de teste. As métricas de desempenho (Tabela 2) foram geradas a partir da matriz de confusão. A precisão da classe do Tomate de 69,70%, menor que para a classe do Morango, de 80,65%, indicam que esta última apresenta maior capacidade de identificar corretamente a classe do fruto analisado. Fato reforçado pela sensibilidade, da qual a classe do Tomate apresenta 79,31%, mais alta que a classe do Morango, de 71,43%, esta última possuindo maior seletividade dos dados classificados. O F1-score, contudo, se mostra parecido entre as classes, sendo de 74,19% para a classe do Tomate e 75,76% para a classe do Morango. A acurácia geral do modelo foi influenciada pelo desempenho das classes, sendo de 75,00%. As métricas de desempenho poderiam apresentar melhores resultados caso fossem realizadas em ambiente controlado, com homogeneidade do ambiente externo, ou em ambiente aberto, mas com controle de posicionamento da câmera e quantidade máxima de frutos por imagem. Contudo, a ausência de controle nas variáveis citadas promove um algoritmo robusto, com maiores possibilidades de utilização.

TABELA 2. Métricas de avaliação do processamento dos dados de teste

	Acurácia	Precisão	Sensibilidade	F1-Score
Tomate	75,00%	69,70%	79,31%	74,19%
Morango		80,65%	71,43%	75,76%

Estudos visando a automação de sistemas a partir outros algoritmos de CNN customizados e pré-definidos também vem sendo testados por outros autores. Chakraborty et al. (2023) desenvolveram a CNN SortNet, focada na classificação de laranjas quanto a sua qualidade, classificando-as como aceitável ou rejeitável. Para isso, foi desenvolvida uma máquina, da qual era acoplada um Raspberry Pi 4 junto de um módulo de câmera, obtendo as imagens dos frutos. A classificação foi realizada em 0,32s por fruto, com a CNN apresentando 97,6% de acurácia, possuindo desempenho superior a outras duas CNNs já existentes, GoogLeNet e SortNet. De Luna et al. (2020) Desenvolveram um sistema de detecção de defeitos em frutos de tomate, utilizando o Raspberry Pi, acoplado com um módulo de Câmera. Foram treinadas as redes neurais convolucionais VGG16, InceptionV3 e ResNet50, apresentando acurácia de para os dados de treinamento de 95,75%, 56,38% e 90,58%, respectivamente, e acurácia para os dados de teste de 98,75%, 58,33% e 64,58%, respectivamente. Com isso, a VGG16 apresentou o melhor desempenho na identificação dos frutos de tomate quanto a presença de injúrias. Portanto, a implementação com sucesso da CNN no microprocessador Raspberry Pi 3 B+ mostra que é possível a construção de um dispositivo óptico e de baixo custo, capaz de classificar frutos quanto categorias pré-definidas, atendendo tanto as necessidades da agroindústria quanto do mercado consumidor. Industrialmente, pode-se escolher os melhores frutos de tomate e morango para o processamento, enquanto para o mercado consumidor, é possível entender a demanda de consumo, e assim obter um classificador automático que selecione os frutos que atendam a necessidade do mercado consumidor. A Figura 2 apresenta o gráfico de Loss e Acurácia para os dados de treinamento e teste, obtido durante o treinamento da CNN desenvolvida. Nela, é possível observar que o Loss para os dados de treinamento decresce rapidamente, permanecendo abaixo de 0,10 durante boa parte do treinamento. Para os dados de teste, embora apresentem variações, os valores de Loss se estabilizaram com o avanço do treinamento, mas ainda apresentando valores quase sempre acima de 1,00. Com isso, é notável a influência do dataset utilizado, do qual mesmo com um bom treinamento, ainda apresenta dificuldades na identificação dos dados de teste, promovendo um possível overfitting. Em relação a acurácia, o modelo apresentou valores muito próximos de 1,00 para os dados de treinamento, enquanto para os valores de teste, não ultrapassou 0,85, permanecendo a última

no modelo gerado, de 0,75. Os melhores valores de Loss e Acurácia foram obtidos na época 28, sendo de 0,55 e 0,84, respectivamente.

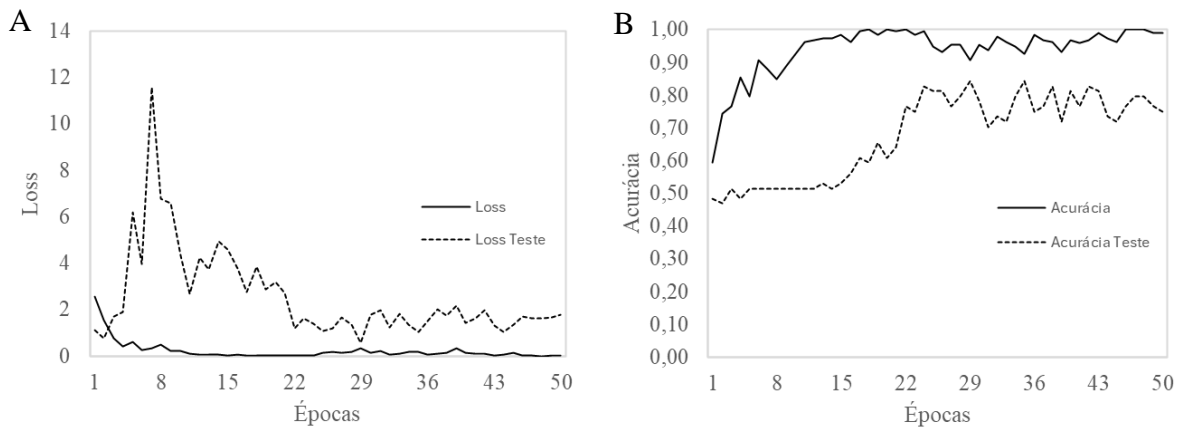


FIGURA 2: Gráfico de Loss (A) e Acurácia (B) para os dados de treinamento e teste. Fonte: Autoria própria

Chakraborty et al. (2023) obtiveram valores baixos e próximos de Loss durante o treinamento da rede neural convolucional customizada desenvolvida, o que indica que não houve overfitting, e que o modelo irá operar com bom desempenho na classificação dos frutos de laranja. Consequentemente, também foi obtido valores de acurácia próximos para treinamento e teste. O desempenho só foi possível pela padronização das imagens obtidas, em ambiente fechado, com iluminação igual e somente um fruto por vez. Pahn et al. (2023), mesmo com a captura das imagens de fruto de tomate em ambiente aberto, também obtiveram valores baixos e próximos de Loss e valores altos e próximos de Acurácia para os dados de treinamento e teste, devido ao controle no momento da captura das imagens, com posicionamento e quantidade de frutos por imagem, junto do destaque do fruto em relação ao fundo

**CONCLUSÕES:** O algoritmo se mostrou apto no processamento de imagens de morango e tomate para treinamento de uma rede neural convolucional customizada e classificação dos frutos, apresentando um acurácia global de 75,00% na detecção dos frutos. Contudo, devido ao tempo de processamento elevado, a utilização do microcontrolador Raspberry Pi 3 B+ é indicada para aplicações com banco de dados com quantidades reduzidas de imagens.

**AGRADECIMENTOS:** Agradecimentos a Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ pelo financiamento dos recursos para o custeio deste projeto e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo financiamento da bolsa de mestrado.

## REFERÊNCIAS:

- ABEKOON, Thilina et al. Image processing techniques to identify tomato quality under market conditions. *Smart Agricultural Technology*, v. 7, p. 100433, 2024.
- ALZUBAIDI, Laith et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, v. 8, p. 1-74, 2021
- CHAKRABORTY, Subir Kumar et al. Development of an optimally designed real-time automatic citrus fruit grading–sorting machine leveraging computer vision-based adaptive deep learning model. *Engineering Applications of Artificial Intelligence*, v. 120, p. 105826, 2023.

DE LUNA, Robert G. et al. Tomato fruit image dataset for deep transfer learning-based defect detection. In: 2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM). IEEE, 2019. p. 356-361.

FAOSTAT. 2024. Statistics Database. Crops and livestock products. Disponível em: <https://www.fao.org/faostat/en/#data/QCL>. Acesso em 2024

GU, J. et al. Recent advances in convolutional neural networks. *Pattern recognition*, v. 77, p. 354-377, 2018.

IBGE – Instituto Brasileiro de Geografia e Estatística. Estatística Mensal da Produção Agrícola. Levantamento Sistemático da Produção agrícola. Janeiro, 2024.

LIMA, A. F.; SILVA, E. G. de A.; IWATA, B. de F. Agriculturas e agricultura familiar no Brasil: uma revisão de literatura. *Retratos de Assentamentos*, v. 22, n. 1, p. 50-68, 2019.

OLIVEIRA, C. T. de. Monitoramento da maturação dos frutos e de doenças do cafeeiro utilizando modelos de deep learning. 2021.

PHAN, Quoc-Hung et al. Classification of Tomato Fruit Using Yolov5 and Convolutional Neural Network Models. *Plants*, v. 12, n. 4, p. 790, 2023.

PALUMBO, M. et al. Computer vision system based on conventional imaging for non-destructively evaluating quality attributes in fresh and packaged fruit and vegetables. *Postharvest Biology and Technology*, v. 200, p. 112332, 2023.

SHIVASHANKAR, M.; PANDEY, Manish; ZAKWAN, Mohammad. Estimation of settling velocity using generalized reduced gradient (GRG) and hybrid generalized reduced gradient–genetic algorithm (hybrid GRG-GA). *Acta Geophysica*, v. 70, n. 5, p. 2487-2497, 2022.

TSAI, Fa-Ta et al. Tomato Fruit Detection Using Modified Yolov5m Model with Convolutional Neural Networks. *Plants*, v. 12, n. 17, p. 3067, 2023.