

# Segmentação de Objetos em Tempo Real na Fruticultura: Contribuições para Agricultura 4.0

Valéria Ribeiro dos Santos, Elloá B. Guedes

<sup>1</sup>Grupo de Pesquisa em Sistemas Inteligentes  
Escola Superior de Tecnologia  
Universidade do Estado do Amazonas  
Av. Darcy Vargas, 1200 – Manaus – Amazonas  
{vrds.eng21, ebgcosta}@uea.edu.br

**Abstract.** *Aiming at contributing with the development of Agriculture 4.0 solutions for Fruit Farming, this work addressed the task of apple instance segmentation from images, considering different ripening stages. The proposed solution consisted of a YOLOv11 Small convolutional neural network trained and tested on a public dataset containing modal masks of partially occluded fruits, reflecting realistic agricultural contexts. The model proved to be effective and efficient, achieving an experimental mAP@0.5 of 0.823 at 52 FPS, contributing to real-time intelligent solutions that assist in production quantification, automatic harvesting, and strategic decision-making in Digital Agriculture.*

**Resumo.** *Com vistas a colaborar no desenvolvimento de soluções da Agricultura 4.0 para a Fruticultura, este trabalho abordou a tarefa de segmentação de instâncias de maçãs a partir de imagens, considerando diferentes estágios de maturação. A solução proposta consistiu em uma rede neural convolucional YOLOv11 Small treinada e testada em um conjunto de dados público contendo máscaras modais de frutos parcialmente oclusos, refletindo contextos agrícolas realísticos. O modelo mostrou-se eficaz e eficiente, tendo obtido mAP@0,5 experimental igual a 0,823 a 52 FPS, colaborando para soluções inteligentes em tempo real que auxiliam na quantificação da produção, na colheita automática e na tomada de decisões estratégicas na Agricultura Digital.*

## 1. Introdução

Os Objetivos de Desenvolvimento Sustentável (ODS) das Organizações das Nações Unidas (ONU), adotados em 2015, são um chamado à ação para pôr fim à pobreza, proteger o planeta e garantir que até 2030 todas as pessoas desfrutem de paz e prosperidade [ONU 2015]. Um dos objetivos dessa agenda (ODS 2) contempla aspectos relacionados à Agricultura, tais como dobrar a produtividade agrícola, implementar práticas resilientes que aumentem a produção, e promover a pesquisa e o desenvolvimento de tecnologias nesse âmbito. Tais metas visam mitigar os desafios decorrentes das mudanças climáticas e abordar a necessidade urgente de aumento da produção alimentar em pelo menos 70 % até 2050 [FAO 2009].

Soluções oriundas da Agricultura Digital podem colaborar para a perspectiva apresentada, pois promovem o aprimoramento da produção agrícola em termos de eficiência, qualidade e sustentabilidade por intermédio de novas tecnologias [Tang et al. 2002]. Quando essas novas tecnologias permitem a interconexão de maquinário, equipamentos, sensores, bem como o uso de ferramentas para coleta, armazenamento e análise de dados relevantes para a elaboração de recomendações práticas e acuradas para o contexto em questão, tem-se a Agricultura 4.0, fundamentada na Internet das Coisas, *Big Data*, Inteligência Artificial, Automação

e Robótica e Computação em Nuvem [Marçal de Queiroz et al. 2022]. Nesta perspectiva, as soluções para o contexto agrícola são orientadas a dados, há o monitoramento e controle em tempo real, otimiza-se o uso de recursos e a tomada de decisão é precisa e automatizada [Da Silveira et al. 2021]. De maneira geral, as inovações contempladas na Agricultura convergem para contemplar o ODS 2.

A Fruticultura, um setor de destaque no agronegócio, com papel significativo na economia global, segurança alimentar, geração de empregos, dentre outros [FAO 2023], oferece um campo fértil para a implementação de soluções inovadoras da Agricultura 4.0. A sustentabilidade deste setor enfrenta desafios multifacetados, que incluem a adaptação às mudanças climáticas, a otimização dos custos de mão de obra, o manejo integrado de pragas e doenças, e a mitigação das perdas pré e pós-colheita, estas últimas frequentemente resultantes de deficiências nutricionais e falhas nos processos de armazenamento e transporte [Pechan et al. 2023, Rutledge and Mérel 2022, Fachinello 2009].

Com vistas a colaborar no desenvolvimento de soluções da Agricultura 4.0 para a Fruticultura, o presente trabalho apresenta uma solução inteligente de Visão Computacional para monitoramento de pomares de maçãs com localização e classificação de frutos com diferentes graus de maturidade por meio de máscaras de segmentação. Os dados experimentais são oriundos de uma base de dados disponível na literatura contemplando cenários realísticos desse tipo de cultivo, os quais foram utilizados para treinar modelos inteligentes do estado da arte para a tarefa considerada. Os resultados comprovam a eficácia da metodologia, colaborando para o desenvolvimento de sistemas automatizados para monitoramento visual de frutos, colheita robótica e estimativas de produção aprimoradas, com potencial para incremento de produtividade.

Para apresentar o que se propõe, o presente trabalho está organizado como segue. A Seção 2 descreve os trabalhos relacionados. A metodologia de desenvolvimento da solução proposta é detalhada na Seção 3. Os resultados obtidos são apresentados na Seção 4. Finalmente, as considerações finais e as perspectivas de trabalhos futuros são discutidas na Seção 5.

## 2. Trabalhos Relacionados

Para compreender de forma abrangente as soluções existentes, identificar tendências e lacunas, realizou-se uma análise dos resultados de buscas no Google Acadêmico entre 2019 e 2024. Os critérios de inclusão contemplaram trabalhos baseados em Inteligência Artificial para detecção e segmentação na Fruticultura.

Métodos tradicionais de detecção de maçãs usualmente são voltados para um estágio específico de maturação. Para contornar esta limitação, o trabalho de Tian *et al.* [2019] propõe a classificação dos frutos em três estágios de crescimento, a citar: jovem, em expansão e maduros. Além disso, os autores demonstraram que o modelo YOLOv3 incorporado com a rede DenseNet superou outras versões da Família YOLO, como YOLOv2 e YOLOv3 tradicional, e ainda modelos como Faster R-CNN com VGG-16, tanto em termos de detecção ( $F_1$ -Score igual a 0,896 e IoU igual a 0,817) quanto no tempo de inferência (3 FPS). Embora a solução seja uma referência importante para abordagens baseadas em *Deep Learning* voltadas à fruticultura, ela não oferece uma delimitação mais precisa das formas das maçãs em comparação com a tarefa de segmentação de instâncias.

Utilizando fotos de laranjas em pomares, Ganesh *et al.* [2019] empregaram o *framework Mask R-CNN* para a tarefa de segmentação de instâncias. A abordagem adotada pelos autores considerou o uso de dados de entrada multimodais, com a incorporação de imagens no espaço de cores HSV e RGB simultaneamente, o que ocasionou uma redução significativa

na ocorrência de erros de detecção do tipo falsos positivos. Os resultados experimentais, com  $F_1$ -Score igual a 0,887 demonstram que essa estratégia melhorou a acurácia da tarefa sem comprometer substancialmente o tempo de inferência, que foi, em média, 11 ms por imagem (90 FPS). No entanto, o estudo apresentou limitações, como a ausência de categorização referente aos estágios de maturação das laranjas, o que poderia agregar valor significativo à aplicação do modelo em contextos agrícolas realísticos.

O estudo de Gené-Mola *et al.* [2023] investigou a segmentação amodal para detecção e predição do tamanho de maçãs, utilizando o *dataset* AmodalApple-Size\_RGB-D. A metodologia empregou uma Rede Neural Convolucional (CNN, do inglês *Convolutional Neural Networks*) para prever regiões visíveis e oclusas das frutas, permitindo a estimativa do diâmetro e da porcentagem de visibilidade por meio da combinação de máscaras modais e amodais, principal contribuição do trabalho (mAP@0,2 igual a 0,473 e  $F_1$ -Score igual a 0,857). As máscaras amodais previstas foram utilizadas para estimar o diâmetro dos frutos em pixels, enquanto as máscaras modais permitiram a identificação da região visível e a medição da distância entre as maçãs e a câmera com base em imagens de profundidade. A anotação manual das máscaras amodais em conjuntos de dados do mundo real, no entanto, mostra-se onerosa, imprecisa e potencialmente inconsistente devido às variações na percepção dos anotadores e na modelagem da parte oclusa dos frutos [Li et al. 2023]. Além disso, as imagens de profundidade foram obtidas por métodos de fotogrametria, o que restringe a aquisição de dados por conta da especificidade da técnica, a qual exige equipamentos especializados.

A partir da investigação realizada, foi possível constatar que a literatura carece de métodos de segmentação para Fruticultura que abordem máscaras modais e múltiplos estágios de maturação, bem como avaliações de modelos do estado da arte. Esta lacuna motivou o presente estudo.

### 3. Material e Métodos

Para contribuir com a Agricultura 4.0, o presente trabalho considerou o contexto de uma fazenda inteligente (*smart farm*) responsável pela produção de maçãs em larga escala, conforme ilustrado na Figura 1. O monitoramento da plantação é feito por meio da captura de imagens, tais como por câmeras e drones. A solução proposta no escopo deste trabalho visa, a partir de tais imagens, segmentar, de maneira automática e inteligente, instâncias dos frutos conforme os graus de maturação, isto é, prever máscaras que delimitam as partes visíveis de cada maçã, classificando-as em verdes ou maduras. As previsões elaboradas pela solução proposta podem então vir a ser utilizadas para quantificar a produção, a partir da contagem de instâncias, e também direcionar colheitadeiras robóticas para os frutos maduros.

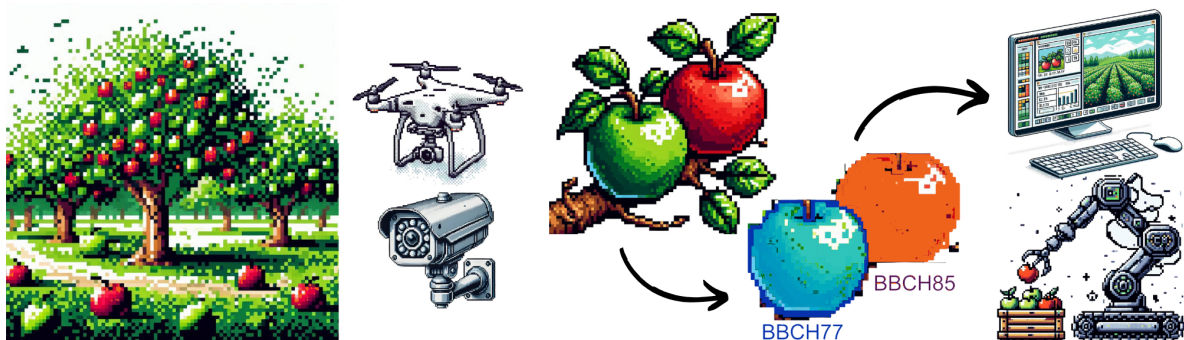


Figura 1: Visão geral da solução proposta.

Considerando o que se propõe, é necessário analisar a eficácia e eficiência de modelos inteligentes do estado da arte para a tarefa de Visão Computacional de segmentação de instâncias. Para este fim, as seções a seguir contemplam a descrição dos dados experimentais, a apresentação dos modelos, as métricas de desempenho e o ambiente de desenvolvimento.

### 3.1. Dados Experimentais

Os dados experimentais utilizados no escopo deste trabalho são oriundos do AmodalApple-Size\_RGB-D *dataset*, base de dados pública e gratuitamente disponível, contendo um total de 15.335 instâncias de maçãs Fuji, disponíveis em 3925 exemplos [Gené-Mola et al. 2024]. As imagens contemplam maçãs capturadas em dois estágios de crescimento, isto é, duas classes: BBCH77, com 5837 instâncias, correspondendo a frutas verdes e com cerca de 70 % do tamanho final; e, BBCH85, com 9498 instâncias, denotando frutas maduras. As máscaras de segmentação das frutas foram elaboradas por especialistas e encontram-se exemplificadas nas Figuras 2 e 3.



(a) Imagem original



(b) Imagem segmentada

Figura 2: Exemplos de máscaras de instâncias da classe BBCH77.



(a) Imagem original



(b) Imagem segmentada

Figura 3: Exemplos de máscaras de instâncias da classe BBCH85.



A partir de uma análise exploratória das anotações no conjunto de dados foi possível observar que o número de frutos por imagem possui variações, que o formato das máscaras é heterogêneo e que o tipo de tais máscaras é modal, isto é, contempla apenas a parte visível do fruto na imagem. Ademais, as condições de captura das imagens não são controladas: há variações de ângulo, iluminação e elementos de *background*, refletindo um cenário realístico da tarefa, o que é relevante para a avaliação da robustez dos modelos.

Os proponentes da base de dados utilizada dividiram as imagens em partições de treino (60 % dos exemplos), validação (20 %) e testes (20 %), com instâncias de cada classe distribuídas conforme histograma da Figura 4. A estratégia de validação cruzada proposta mostra-se adequada para fins de reprodutibilidade experimental e para comparação entre modelos. Cabe também ressaltar que o número de instâncias por imagem é variado, conforme Figura 5.

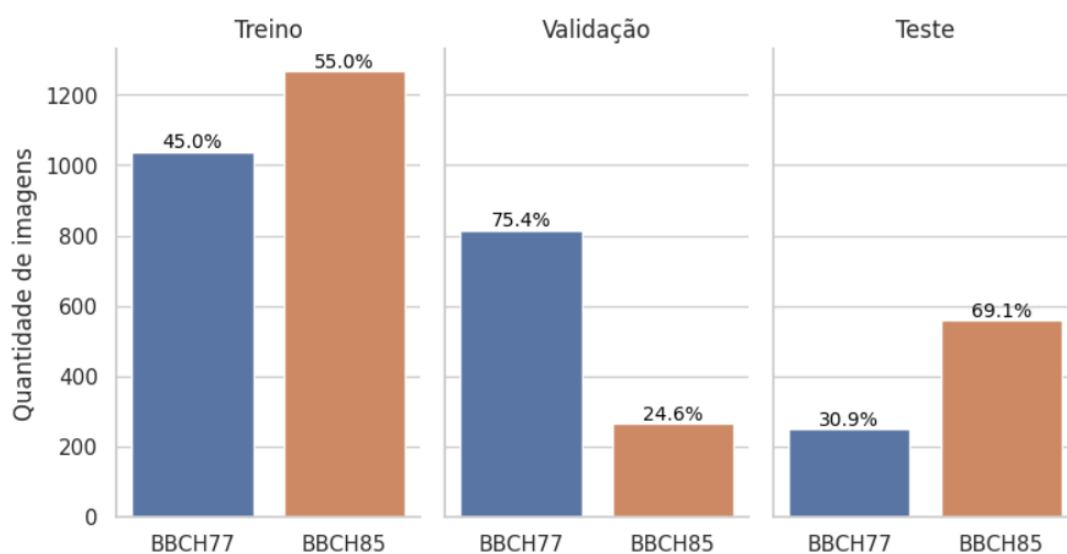


Figura 4: Histograma das partições do conjunto de dados.

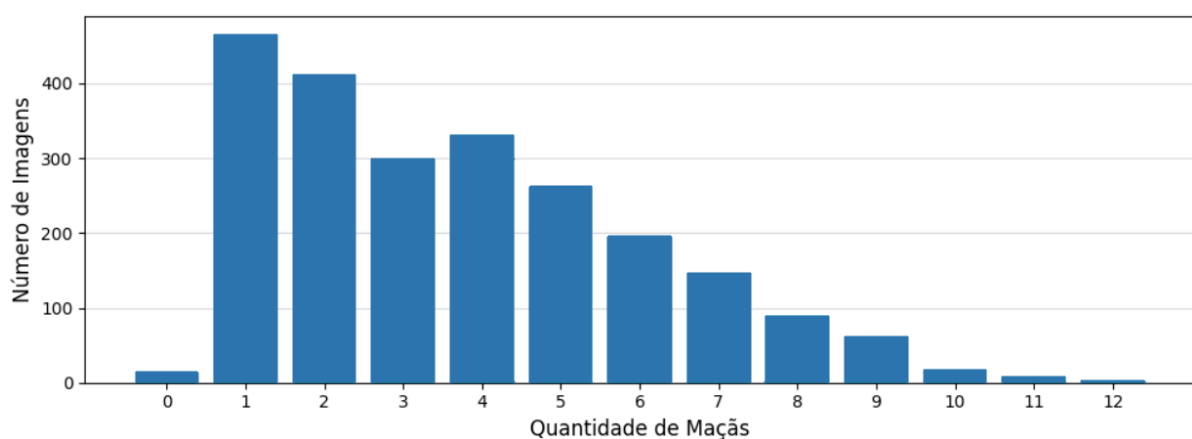


Figura 5: Histograma da quantidade de maçãs por imagem.

### 3.2. Modelos e Parametrização

O Aprendizado Profundo (AP), impulsionado por redes neurais artificiais multicamadas treinadas com grandes conjuntos de dados, consolidou-se como o paradigma dominante em Aprendizado de Máquina, com avanços notáveis em domínios como Visão Computacional

– a análise e interpretação automática de imagens. A maioria desses progressos decorre do emprego de CNNs, arquiteturas compostas por múltiplas camadas de convoluções, intercaladas com operações de *pooling* e, frequentemente, camadas densas nos estágios finais [Bishop and Bishop 2024].

A Família de CNNs denominada YOLO (acrônimo para *You Only Look Once*) destaca-se em tarefas de detecção de objetos e segmentação de instâncias, devido à sua eficiência computacional e capacidade de processamento em tempo real. Como detectores de estágio único, esses modelos dividem a imagem de entrada em uma grade, prevendo simultaneamente caixas delimitadoras ou máscaras de segmentação, pontuações de confiança e probabilidades de classe para cada célula [Redmon et al. 2016]. Essa arquitetura permite a detecção e segmentação simultâneas, otimizando o desempenho para aplicações que demandam alta velocidade de processamento [Vijayakumar and Vairavasundaram 2024].

A YOLOv11, versão mais recente da família YOLO, representa um avanço significativo em segmentação de instâncias, integrando inovações arquitetônicas e técnicas de otimização [Ultralytics 2024]. A principal inovação reside no *backbone* C3k2, projetado para extrair características multiescalares com alta eficiência computacional, reduzindo a complexidade através do particionamento da convolução, em comparação ao CSP (*Cross Stage Partial*) de versões anteriores. Adicionalmente, a YOLOv11 aprimora os avanços da YOLOv8, incorporando otimizações de parâmetros que resultam em desempenho superior na detecção de objetos [Khanam and Hussain 2024]. A combinação dessas inovações reforça a relevância da YOLOv11 como ferramenta do estado da arte em Visão Computacional, consolidando sua posição como arquitetura robusta e eficiente, justificando sua escolha para o presente estudo.

Para a família YOLO considerou-se a versão YOLOv11 e os modelos *Nano*, *Small*, *Medium*, *Large* e *Extra Large*. Todos esses modelos foram treinados com 500 épocas, paciência igual a 50, otimizador Adam e demais hiperparâmetros em valores padrão. Os tamanhos de lote variaram conforme o tamanho dos modelos, equilibrando a alocação de parâmetros e o uso da memória principal, em que 16 imagens por lote foram consideradas para a versão *Nano*, 8 para os modelos *Small* e *Medium*, e 4 para os modelos *Large* e *Extra Large*.

### 3.3. Avaliação de Desempenho

A avaliação de desempenho dos modelos propostos foi realizada a partir da análise das previsões efetuadas para o conjunto de testes conforme as seguintes métricas de desempenho:

1. **mAP@0,5.** Avalia a precisão média da detecção de objetos, aferindo a sobreposição (IoU – *Intersection over Union*) entre máscaras previstas e reais. Um IoU de 0,5 indica que a área de sobreposição deve ser de pelo menos 50 % para o cômputo de uma detecção correta. A métrica fornece um desempenho geral dos modelos na detecção e segmentação precisa de instâncias;
2. **F<sub>1</sub>-Score.** Denota a média harmônica entre a precisão e a revocação. A precisão avalia a exatidão das segmentações, ou seja, a proporção de pixels corretamente classificados como maçãs BBCH77 ou BBCH85 em relação ao total de pixels segmentados como tal. A revocação, por sua vez, mede a capacidade do modelo de identificar todos os pixels de maçãs na imagem. A métrica em questão foi escolhida para mostrar de forma isonômica o desempenho dos modelos apesar do desbalanceamento entre as classes na partição de testes;
3. **Frames por Segundo.** Refere-se à taxa de quadros de vídeo que o modelo de segmentação é capaz de processar em um segundo. Contribui para aferir a adequação do modelo para uso em sistemas de tempo real.

### 3.4. Ambiente de Desenvolvimento

Os modelos foram treinados e avaliados utilizando Python como linguagem de programação, juntamente com os *frameworks* TensorFlow e PyTorch. As execuções ocorreram em um servidor de alto desempenho equipado com processador Intel® Core™ i9-10900K com clock de 3,7 GHz, memória principal de 64 GB, 7 TB de memória secundária e 3 placas gráficas NVIDIA GTX 3060 com 12 GB VRAM cada. Essa configuração viabilizou a aceleração em *hardware* para o treinamento dos modelos.

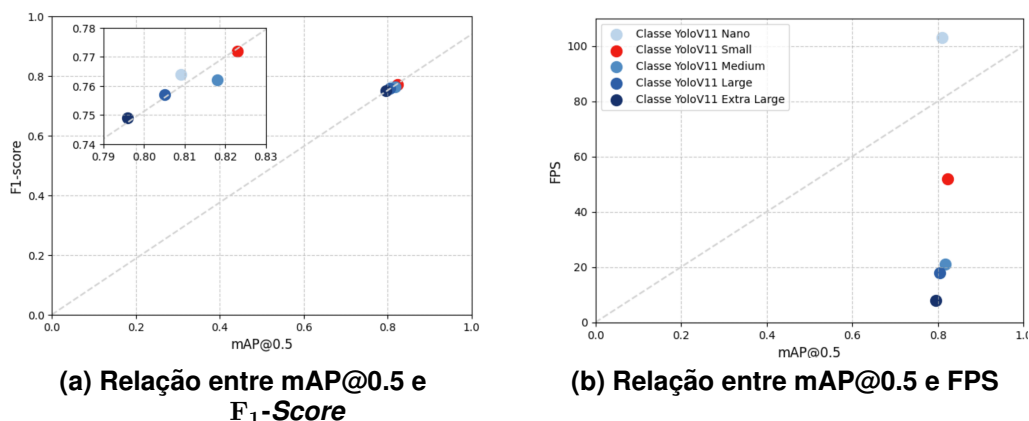
## 4. Resultados e Discussão

Os experimentos computacionais foram executados, e aferiu-se o desempenho dos modelos de segmentação de maçãs na partição de testes da base de dados selecionada. Os resultados dos modelos YOLOv11 encontram-se dispostos na Tabela 1.

**Tabela 1: Avaliação de desempenho dos modelos YOLOv11 perante a base de dados.**

Modelo	Treinamento			Teste		
	Parâmetros	Tempo	Épocas	mAP@0,5	F <sub>1</sub> -Score	FPS
YOLOv11 Nano	2.834.958	8h 50min	286	0,809	0,764	103
YOLOv11 Small	10.067.590	4h 15min	134	0,823	0,772	52
YOLOv11 Medium	22.336.854	10h 9min	154	0,818	0,762	21
YOLOv11 Large	27.586.134	11h 5min	139	0,805	0,757	18
YOLOv11 Extra Large	79.600.000	32h 33min	104	0,796	0,749	8

Tomando como referências as métricas de  $F_1$ -Score e de mAP@0,5, o modelo que maximizou ambas as métricas no cenário experimental considerado foi a YOLOv11 *Small*, informação esta corroborada pelo gráfico da Figura 6a. Tal modelo, embora mais acurado na segmentação de instâncias, não possui maior FPS, conforme Figura 6b, embora os valores obtidos sejam competitivos para cenários realísticos, considerando que os dispositivos para captura de imagens dos pomares podem ser estáticos ou ter sua velocidade de deslocamento controlada.



**Figura 6: Relação entre métricas dos modelos na experimentação**

Exemplos de segmentação com o modelo de referência proposto podem ser verificados nas Figuras 7-9, as quais contém instâncias de teste do *dataset* elencado. Observa-se que há elementos variados nas imagens, tais como trenas e suportes, que a situação de captura

das imagens não é controlada, que a oclusão dos frutos é elevada e que instâncias com graus de maturação distintos podem ocorrer em uma mesma imagem. Nesse contexto realístico, verifica-se que houve robustez da solução proposta.

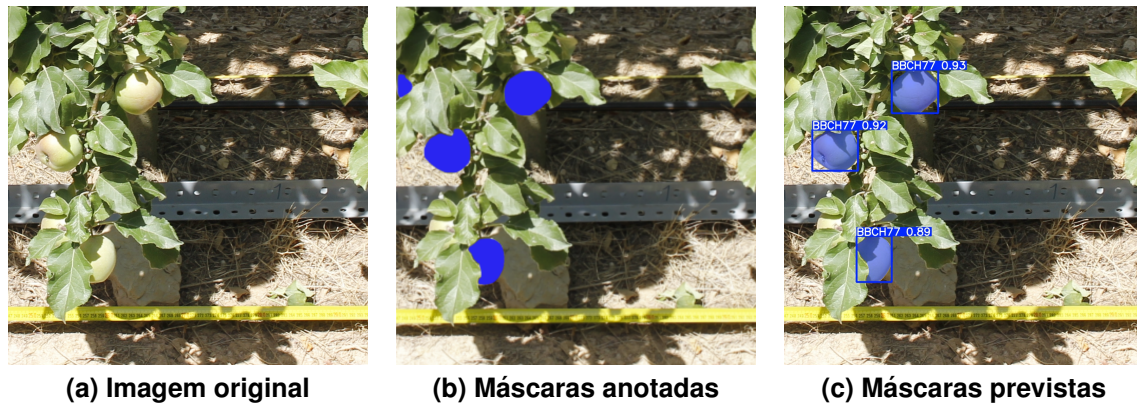


Figura 7: Exemplo de previsão do modelo YOLOv11 *Small* no exemplo \_MG\_6495\_16.png.

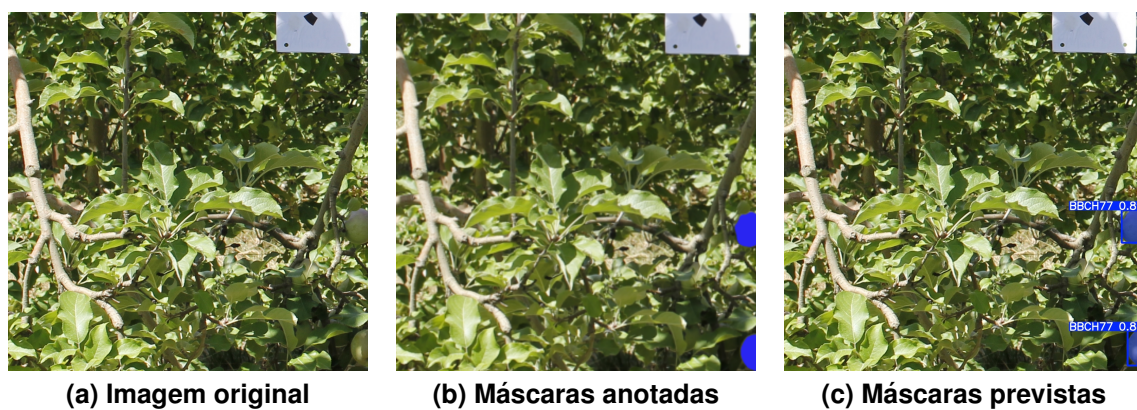


Figura 8: Exemplo de previsão do modelo YOLOv11 *Small* no exemplo \_MG\_6499\_14.png.

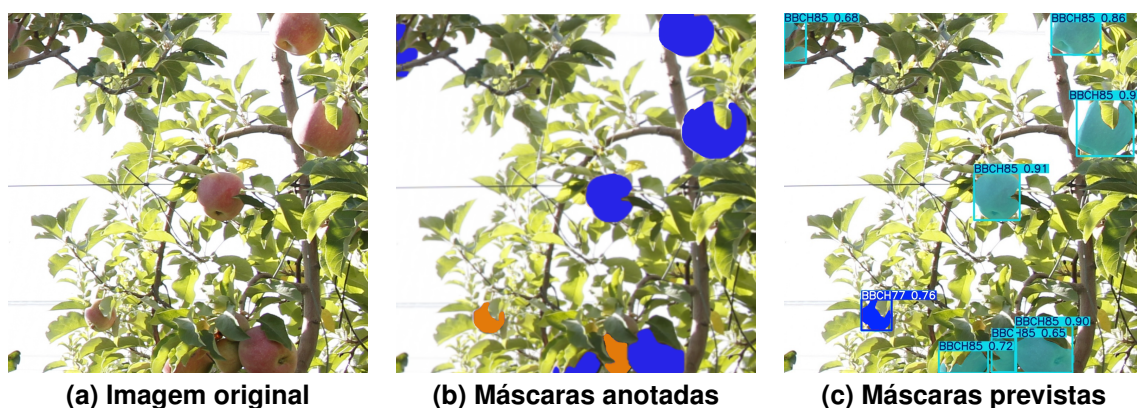


Figura 9: Exemplo de previsão do modelo YOLOv11 *Small* no exemplo \_MG\_2657\_18.png.

É interessante notar que todos os modelos possuem  $mAP@0,5$  próximo, variando em um intervalo com diferença numérica de 0,027. Embora se elenque a YOLOv11 *Small* como solução candidata, as demais não deixam de ter potencial prático. Ainda assim, destaca-se o menor tempo de convergência da solução proposta dentre todos os modelos avaliados, bem



como o segundo menor quantitativo de parâmetros. Esses aspectos tornam-se imperativos, considerando o crescente interesse em Computação de Borda e aplicações móveis para Agricultura Digital.

Uma semelhança da solução proposta com os trabalhos relacionados reside no uso de CNNs, porém considerou-se modelos YOLOv11 que foram mais recentemente propostos na literatura e que demonstraram eficácia e eficiência. Particularmente, há diferenças nos cenários experimentais, principalmente oriundos dos diferentes *datasets* utilizados. Logo, não é possível uma comparação mais objetiva entre eles. No entanto, este trabalho considera apenas imagens capturadas com câmeras comuns, o que viabiliza a utilização de dispositivos móveis, por exemplo, diminuindo o ônus financeiro e tirando proveito de infraestrutura de aquisição de imagens já existente. Além disso, ao utilizar máscaras modais dos frutos e contemplar diferentes estágios de maturação, favorece a aplicação e implantação da solução proposta em condições reais no contexto da Fruticultura perante a Agricultura 4.0.

## 5. Considerações Finais

O presente trabalho demonstrou a eficácia dos modelos YOLOv11 na segmentação de maçãs em diferentes estágios de maturação, destacando a parametrização da YOLOv11 *Small* para aplicações em tempo real na Agricultura 4.0. A segmentação permitiu a contagem automatizada de frutos e a análise de seu estágio de maturação, possibilitando estimativas de produção precisas e otimização do monitoramento do pomar, permitindo um acompanhamento mais detalhado do desenvolvimento das culturas e subsidiando a tomada de decisões eficazes.

Em trabalhos futuros, almeja-se expandir a avaliação de modelos de *Deep Learning* na tarefa em questão com vistas a uma maior eficiência, contemplando soluções recentemente propostas na literatura, como as YOLOv12 e modelos baseados em *Transformers*.

## Transparência da IA e Considerações Éticas

Os ícones apresentados na Figura 1 foram gerados utilizando o modelo DALL-E 3. A revisão gramatical e ortográfica foi realizada com a ferramenta Gemini 2.0 Flash. Além disso, todas as revisões automáticas foram complementadas por uma revisão humana para garantir a precisão e a qualidade do texto.

## Agradecimentos

As autoras agradecem o apoio financeiro concedido pela Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) por meio do Programa PAIC 2024/2025. EBG agradece o apoio financeiro da CAPES por meio do Projeto COFECUB-PJ3126170P. As autoras agradecem também o apoio material recebido do Laboratório de Sistemas Inteligentes (LSI) da Universidade do Estado do Amazonas (UEA).

## Referências

- Bishop, C. M. and Bishop, H. (2024). *Deep Learning: Foundations and Concepts*. Springer.
- Da Silveira, F., Lermen, F. H., and Amaral, F. G. (2021). An Overview of Agriculture 4.0 Development: Systematic Review of Descriptions, Technologies, Barriers, Advantages, and Disadvantages. *Computers and Electronics in Agriculture*, 189:106405.
- Fachinello, J. (2009). *Fruticultura: Fundamentos e Práticas*. Embrapa Clima Temperado.
- FAO (2009). Organização das Nações Unidas para a Alimentação e a Agricultura (FAO) – How to Feed the World in 2050. Disponível em: <https://tinyurl.com/4rvakbey>. Acesso: 04/mar/2025.

- FAO (2023). *The State of Food and Agriculture 2023*. Food and Agriculture Organization of the United Nations, Itália.
- Ganesh, P., Volle, K., Burks, T., and Mehta, S. (2019). Deep orange: Mask r-cnn based orange detection and segmentation. *Ifac-papersonline*, 52(30):70–75.
- Gené-Mola, J., Ferrer-Ferrer, M., Gregorio, E., Blok, P. M., Hemming, J., Morros, J.-R., Rosell-Polo, J. R., Vilaplana, V., and Ruiz-Hidalgo, J. (2023). Looking Behind Occlusions: A Study on Amodal Segmentation for Robust On-tree Apple Fruit Size Estimation. *Computers and Electronics in Agriculture*, 209:107854.
- Gené-Mola, J., Ferrer-Ferrer, M., Hemming, J., van Dalftsén, P., de Hoog, D., Sanz-Cortiella, R., Rosell-Polo, J. R., Morros, J.-R., Vilaplana, V., Ruiz-Hidalgo, J., and Gregorio, E. (2024). AmodalAppleSize\_RGB-D Dataset: RGB-D Images of Apple Trees Annotated with Modal and Amodal Segmentation Masks for Fruit Detection, Visibility and Size Estimation. *Data in Brief*, 52:110000.
- Khanam, R. and Hussain, M. (2024). YOLOv11: An overview of the key architectural enhancements. Disponível em <https://arxiv.org/abs/2410.17725>. Acesso em 15/mar/2025.
- Li, Z., Ye, W., Terven, J., Bennett, Z., Zheng, Y., Jiang, T., and Huang, T. (2023). Muva: A New Large-scale Benchmark for Multi-view Amodal Instance Segmentation in the Shopping Scenario. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 23504–23513.
- Marçal de Queiroz, D., M. Valente, D. S., de Assis de Carvalho Pinto, F., Borém, A., and Schueller, J. K., editors (2022). *Digital Agriculture*. Springer, Cham.
- ONU (2015). Objetivos de Desenvolvimento Sustentável. Disponível em <https://tinyurl.com/54tdheep>. Acesso: 04/mar/2025.
- Pechan, P. M., Bohle, H., and Obster, F. (2023). Reducing Vulnerability of Fruit Orchards to Climate Change. *Agricultural Systems*, 210:103713.
- Redmon, J., Divvala, S., Girshick, R. B., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788.
- Rutledge, Z. and Mérel, P. (2022). Farm Labor Supply and Fruit and Vegetable Production. *Am. J. Agric. Econ.*
- Tang, S., Zhu, Q., Zhou, X., Liu, S., and Wu, M. (2002). A Conception of Digital Agriculture. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3026–3028, Canada. IEEE.
- Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., and Liang, Z. (2019). Apple Detection during Different Growth Stages in Orchards using the Improved YOLO-V3 model. *Computers and Electronics in Agriculture*, 157:417–426.
- Ultralytics (2024). Ultralytics yolov11. Disponível em: <https://docs.ultralytics.com/models/yolov11/>. Acesso em 15/mar/2025.
- Vijayakumar, A. and Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. *Multimed. Tools Appl.*, 83:83535–83574.