



ANAIS

SEGMENTAÇÃO AUTOMATIZADA DE ALFACE EM CULTIVO HIDROPÔNICO UTILIZANDO YOLO

JAKELINE ANDRADE
jakelineandradeds@gmail.com
UTFPR

PEDRO DE PAULA
plpf2004@gmail.com
UTFPR

JORGE AIKES JUNIOR
jorgeaikes@gmail.com
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

GLAUCO MIRANDA
glaucovmiranda@gmail.com
UTFPR

RESUMO: A produção de alface em sistema hidropônico está sujeita a variações nutricionais e ambientais que podem impactar o desenvolvimento das plantas ao longo do ciclo produtivo. Nesse cenário, o monitoramento do crescimento torna-se fundamental para identificar possíveis alterações de forma precoce. Técnicas automatizadas baseadas em visão computacional surgem como alternativa eficiente para acompanhar esse processo de maneira objetiva e quantitativa. Este trabalho apresenta uma abordagem fundamentada exclusivamente no modelo YOLOv12 para detecção e segmentação de alface em imagens. O modelo foi treinado em regime supervisionado para identificar e segmentar individualmente as plantas. Nos experimentos de detecção (bounding boxes), o YOLOv12 alcançou 97,78% de precisão, 98,46% de recall, 99,24% de acurácia e 93,18% de mAP. Na segmentação, os resultados foram 97,53% de precisão, 95,93% de recall, 98,45% de acurácia e 81,45% de mAP. Embora tenham sido observadas falhas pontuais na detecção em cenários de maior adensamento foliar, tais ocorrências não comprometem o desempenho global, uma vez que a segmentação manteve elevada correspondência espacial mesmo sob maior complexidade visual.

PALAVRAS CHAVE: YOLOv12; Segmentação de imagens; Agricultura de precisão.

ABSTRACT: Hydroponic lettuce production is subject to nutritional and environmental variations that may affect plant development throughout the production cycle. In this context, growth monitoring becomes essential for the early identification of potential changes. Automated computer vision techniques have emerged as an efficient alternative for monitoring this process in an objective and quantitative manner. This study presents an approach based exclusively on the YOLOv12 model for lettuce detection and segmentation in images. The model was trained under a supervised learning framework to individually detect and segment lettuce plants. In the detection experiments (bounding boxes), YOLOv12 achieved 97.78% precision, 98.46% recall, 99.24% accuracy, and 93.18% mAP. For the segmentation task, the model achieved 97.53% precision, 95.93% recall, 98.45% accuracy, and 81.45% mAP. Although occasional detection failures were observed in scenarios with higher leaf density, these instances did not compromise the overall performance, as the segmentation results maintained high spatial correspondence even under increased visual complexity.

KEY WORDS: YOLOv12; Image Segmentation; Precision Agriculture.

1. INTRODUÇÃO

A produção de alface ocupa posição relevante na horticultura brasileira, especialmente em sistemas protegidos que permitem maior controle das condições ambientais (Meirelles; Baldotto; Baldotto, 2017). Dentre as tecnologias adotadas nesses sistemas, destaca-se a hidroponia, caracterizada pelo cultivo sem solo e pelo fornecimento controlado de solução nutritiva às raízes, o que favorece maior eficiência produtiva e o uso racional de recursos (Amaro et al., 2007; Góis et al., 2024; Herrera-Arroyo et al., 2025). Contudo, mesmo em sistemas tecnicamente estruturados, variações na solução nutritiva, na luminosidade ou na temperatura podem comprometer o desenvolvimento vegetal, exigindo estratégias de acompanhamento contínuo do crescimento (Didonet, 2009; Andrade, 2026).

Sob essa perspectiva, o monitoramento manual apresenta limitações, sobretudo em cenários de produção intensiva, nos quais a subjetividade das avaliações visuais e a dificuldade de análise em larga escala reduzem a precisão do acompanhamento (Andrade et al., 2022).

Diante dessas limitações, abordagens baseadas em visão computacional e aprendizado profundo têm se consolidado como alternativas promissoras para análise do crescimento automatizado. Entre os modelos disponíveis, a família YOLO (*You Only Look Once*) destaca-se pela capacidade de realizar detecção e segmentação com elevada precisão e eficiência computacional (Alif; Hussain, 2025), tornando pertinente o desenvolvimento de estratégias automatizadas capazes de acompanhar o crescimento da alface em sistemas hidropônicos de forma objetiva e escalável. Nesse sentido, modelos como o YOLO permitem não apenas a detecção e segmentação das plantas, mas também a extração de informações quantitativas, como a área foliar projetada, utilizada como indicador do desenvolvimento vegetativo.

Diante desse contexto, este trabalho teve como objetivo desenvolver e avaliar um sistema baseado no modelo YOLO para a detecção, segmentação e estimativa da área foliar projetada de plantas de alface cultivadas em sistema hidropônico, visando fornecer uma abordagem automatizada para o monitoramento do crescimento vegetal.

2. PROCEDIMENTOS METODOLÓGICOS

A metodologia proposta neste estudo tem como objetivo empregar o modelo YOLOv12 na segmentação de alfaces cultivadas em sistema hidropônico. Para alcançar esse objetivo, o procedimento foi estruturado em cinco etapas principais: organização do conjunto de dados, anotação das imagens, treinamento do modelo, avaliação de desempenho e análise da inferência.

Na etapa inicial, realizou-se a organização do conjunto de dados denominado Hidroponia em Santa Helena, desenvolvido por meio de uma parceria entre uma empresa privada do setor agrícola e a Universidade Tecnológica Federal do Paraná (UTFPR), câmpus Santa Helena. As imagens foram adquiridas ao longo do ano de 2022, entre os meses de fevereiro e setembro, abrangendo múltiplos ciclos produtivos da alface cultivada em ambiente protegido. Ao todo, foram capturadas 7.650 imagens, as quais foram organizadas em dois grupos principais: (i) imagens destinadas ao treinamento e validação do modelo e (ii) imagens correspondentes a ciclos produtivos completos, reservadas exclusivamente para a etapa de inferência e para a análise temporal do crescimento das plantas. Essa estratégia de particionamento possibilita o treinamento do modelo em um subconjunto específico e sua posterior aplicação em sequências temporais independentes, permitindo avaliar o comportamento do modelo ao longo do desenvolvimento das plantas.

Na segunda etapa, realizou-se a anotação das imagens, com a delimitação individual das instâncias de alface presentes nas bancadas de cultivo. As anotações foram conduzidas manualmente na plataforma CVAT (*Computer Vision Annotation Tool*), por meio da criação de máscaras poligonais que representassem o contorno visível de cada planta. Ressalta-se que, ao considerar o elevado volume de imagens — muitas das quais continham até 80 mudas de alface por cena — optou-se por não anotar integralmente todo o conjunto disponível. A anotação exaustiva demandaria um esforço manual significativamente elevado, tornando o processo operacionalmente inviável dentro do cronograma estabelecido.

No que se refere ao treinamento, foi selecionado o modelo YOLOv12n, uma versão de menor complexidade computacional dentro da família YOLOv12, que incorpora mecanismos de atenção à sua arquitetura. Diferentemente das versões anteriores da família YOLO, predominantemente baseadas em blocos convolucionais convencionais, o YOLOv12 integra módulos capazes de modelar dependências espaciais e contextuais de longo alcance. Essa característica mostra-se particularmente relevante em cenários agrícolas, nos quais ocorrem sobreposição foliar, elevada proximidade entre instâncias e variações graduais de textura ao longo do ciclo de crescimento das plantas.

Em relação ao treinamento, este foi conduzido de forma supervisionada, utilizando exclusivamente o subconjunto previamente anotado. O ambiente computacional adotado foi o Google Colab, com utilização de GPU NVIDIA A100, garantindo capacidade adequada de processamento para as iterações do modelo. Quanto aos hiperparâmetros, estabeleceu-se um limite máximo de 1.000 épocas, com aplicação do critério de *early stopping* baseado em paciência de 100 épocas. Além disso, todas as imagens foram redimensionadas para 640×640 pixels, conforme a configuração padrão da arquitetura.

Após a etapa de treinamento, procedeu-se à análise das métricas de desempenho, com o objetivo de avaliar o comportamento do modelo em diferentes cenários. Buscou-se verificar sua capacidade de discriminar adequadamente as instâncias de alface em relação ao fundo, lidar com a proximidade entre plantas adjacentes e manter consistência na detecção e delimitação das regiões de interesse. Para essa finalidade, foram consideradas as métricas Precisão (*Precision*), Revocação (*Recall*) e *mean Average Precision* (mAP), calculadas a partir da comparação entre as predições do modelo e as anotações de referência (*ground truth*). A Precisão (*Precision*) quantifica a proporção de predições positivas que estão corretas, refletindo o grau de confiabilidade das detecções realizadas. Já o *Recall* mede a capacidade do modelo de identificar todas as instâncias presentes no conjunto de referência, expressando sua sensibilidade à detecção dos objetos de interesse. Por fim, a métrica *mean Average Precision* (mAP) é então obtida a partir da área sob a curva *Precision–Recall*, agregando o desempenho do modelo em diferentes limiares de confiança e, quando aplicável, em diferentes valores de IoU. Essa métrica sintetiza, em um único valor, o equilíbrio entre precisão e revocação, sendo amplamente adotada na avaliação de modelos da família YOLO.

Com relação a etapa final, realizou-se o processo de inferência utilizando o modelo previamente treinado, aplicando-o ao conjunto de imagens destinado à avaliação. As predições geradas — incluindo caixas delimitadoras (*bounding boxes*) e máscaras segmentadas — foram analisadas com o objetivo de examinar o comportamento do modelo sob diferentes condições de iluminação e níveis de complexidade visual da cena. Durante essa etapa, foram considerados cenários caracterizados por sobreposição foliar, variações de contraste, regiões com maior intensidade luminosa e interferências estruturais do ambiente da estufa. Além disso, investigou-se o desempenho do modelo ao longo das diferentes fases de desenvolvimento das hortaliças dentro do período monitorado, observando sua capacidade de

manter consistência na detecção e na segmentação à medida que ocorria aumento de área foliar, maior proximidade entre plantas e intensificação de oclusões parciais.

3. APRESENTAÇÃO DOS RESULTADOS E DISCUSSÕES

O treinamento do YOLOv12 apresentou convergência antes do limite máximo estabelecido, sendo interrompido automaticamente na 625ª época em função do critério de *early stopping*. Esse comportamento indica estabilização das métricas de validação e sugere que o modelo atingiu um ponto de maturidade no processo de aprendizagem sem necessidade de completar as 1.000 épocas inicialmente definidas. O tempo total de processamento para o treinamento das 1.570 imagens anotadas foi de aproximadamente 7 horas e 3 minutos. Ademais, as métricas finais obtidas ao término do treinamento são apresentadas na Tabela 1.

TABELA 1. Métricas de desempenho do YOLOv12n para detecção de objetos (*Box*) e segmentação de instâncias (*Mask*).

Tarefa	Precisão	Recall	mAP@50	mAP@50–95
Detecção de objetos	97.78%	98.46%	99.24%	93.18%
Segmentação	97.53%	95.93%	98.45%	81.45%

Fonte: Elaborado pelo autor (2026).

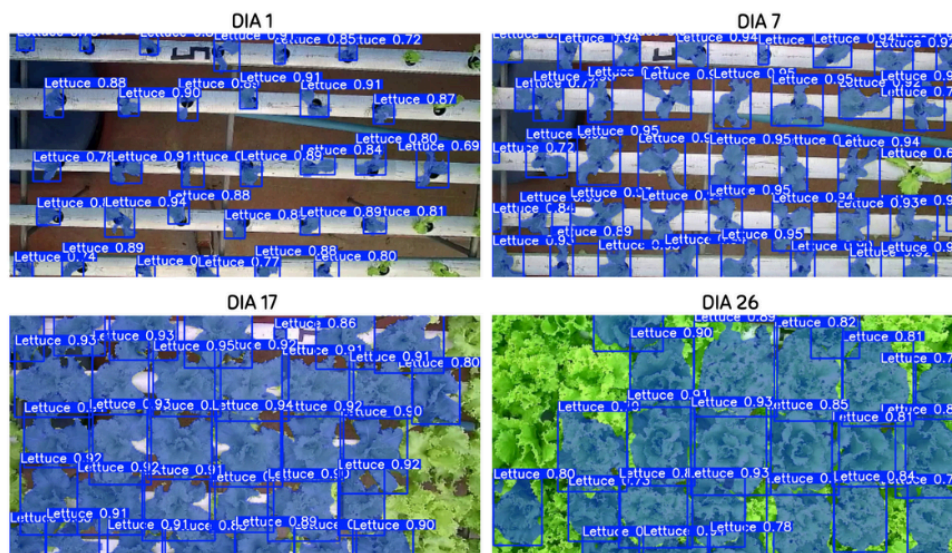
Conforme apresentado na Tabela 1, o desempenho do modelo mostrou-se elevado tanto na tarefa de detecção quanto na segmentação de instâncias. Na detecção de objetos, os valores de Precisão (97,78%) e Recall (98,46%) evidenciam equilíbrio entre baixa taxa de falsos positivos e reduzida ocorrência de omissões, refletindo capacidade consistente de identificação das plantas nas diferentes condições avaliadas. O mAP@50 de 99,24% confirma alta taxa de correspondência entre as caixas preditas e as anotações de referência sob limiar de IoU igual a 0,5, enquanto o mAP@50–95 de 93,18% demonstra manutenção do desempenho mesmo sob critérios mais restritivos de sobreposição.

Na tarefa de segmentação, embora os valores de Precisão (97,53%) permaneçam próximos aos observados na detecção, o Recall (95,93%) apresenta leve redução, sugerindo maior sensibilidade da segmentação às variações de contorno e às oclusões parciais entre folhas adjacentes. O mAP@50 de 98,45% confirma elevada qualidade das máscaras sob limiar moderado de IoU; contudo, a redução para 81,45% no mAP@50–95 evidencia maior impacto quando considerados limiares progressivamente mais rigorosos. Esse comportamento é esperado em cenários com sobreposição foliar e transições suaves de textura, nos quais pequenas imprecisões nas bordas segmentadas influenciam diretamente o cálculo do IoU. No que se refere à etapa de inferência aplicada a um ciclo completo de crescimento, os resultados podem ser observados na Figura 1.

A análise qualitativa das predições ao longo do ciclo produtivo evidencia que o modelo manteve comportamento consistente nas fases iniciais, caracterizadas por menor área foliar e maior espaçamento entre plantas. Nesses estágios, a maioria das hortaliças foi corretamente detectada e segmentada, com delimitação coerente das regiões de interesse. À medida que o ciclo evolui, verifica-se aumento significativo da massa foliar, redução do espaçamento entre indivíduos e intensificação da sobreposição entre folhas, fatores que ampliam a complexidade visual da cena. Nesse contexto, pequenas imprecisões tornam-se mais evidentes, como falhas pontuais de detecção em regiões periféricas da imagem,

comportamento que se mantém até os estágios finais do ciclo. Ainda assim, mesmo diante do adensamento vegetal e da dificuldade em distinguir visualmente os limites entre plantas adjacentes, o modelo preservou capacidade de separar e segmentar parte das instâncias de forma individualizada.

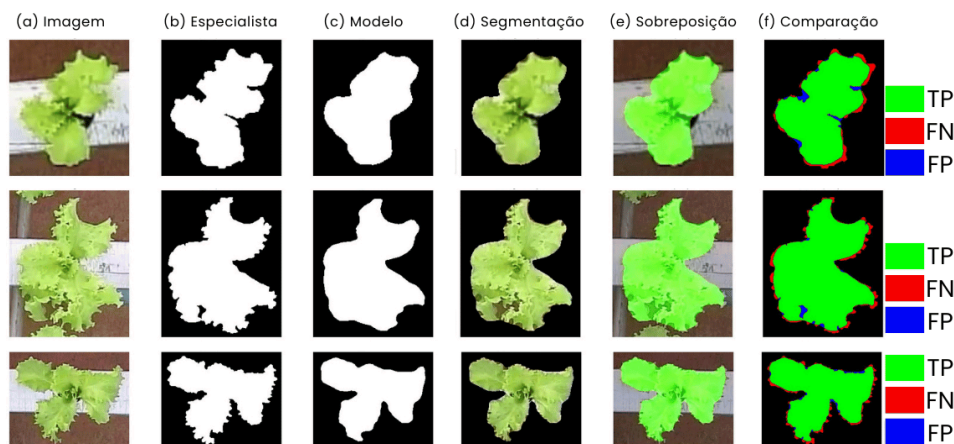
FIGURA 1. Inferência do modelo YOLOv12 com treinamento supervisionado.



Fonte: Elaborado pelo autor (2026).

Esse desempenho demonstra robustez frente à oclusão parcial e às transições graduais de contorno, características típicas de sistemas hidropônicos em fase avançada de crescimento. Tal comportamento pode ser observado na Figura 2, que apresenta, para diferentes exemplos, a imagem original (a), a segmentação de referência elaborada pelo especialista (b), a máscara predita pelo modelo (c), a aplicação da segmentação sobre a imagem (d), a sobreposição entre predição e referência (e) e, por fim, o mapa comparativo com a discriminação de Verdadeiros Positivos (TP), Falsos Negativos (FN) e Falsos Positivos (FP) (f). Na representação final, as regiões em verde correspondem aos pixels corretamente segmentados (TP), enquanto as áreas em vermelho indicam regiões do objeto não capturadas pelo modelo (FN) e as áreas em azul representam pixels incorretamente atribuídos à classe alface (FP). Observa-se predominância de TP em todas as amostras apresentadas, com discrepâncias localizadas principalmente nas bordas foliares, no qual pequenas variações de contorno e transições graduais de textura influenciam o alinhamento pixel a pixel entre a máscara predita e a anotação de referência.

FIGURA 2. Comparação entre o modelo e a máscara verdadeira (especialista).



Fonte: Elaborado pelo autor (2026).

4. CONSIDERAÇÕES FINAIS

Os resultados obtidos apresentaram elevado desempenho tanto na tarefa de detecção quanto na segmentação de instâncias, com valores expressivos de Precisão, *Recall* e $mAP@50$. A diferença observada entre $mAP@50$ e $mAP@50-95$, especialmente na segmentação, evidencia maior sensibilidade da delimitação pixel a pixel em cenários de sobreposição foliar e aumento da complexidade morfológica. Ainda assim, a análise da matriz de confusão confirmou baixa incidência de erros de omissão e reduzida taxa de confusão entre alface e *background*.

Em relação a avaliação qualitativa ao longo de um ciclo completo de crescimento, este mostrou que o modelo manteve estabilidade nas fases iniciais e apresentou desempenho consistente mesmo diante do adensamento vegetal e da intensificação das oclusões nos estágios mais avançados. Embora pequenas imprecisões tenham se tornado mais perceptíveis à medida que a massa foliar aumentava, o modelo preservou capacidade de individualização de instâncias em cenários de elevada complexidade visual.

De modo geral, os resultados obtidos apresentam a viabilidade do uso de arquiteturas baseadas na família YOLO para aplicações agrícolas em ambiente hidropônico, contribuindo para o avanço de soluções automatizadas voltadas ao monitoramento e à análise do desenvolvimento de hortaliças. Como perspectivas futuras, sugere-se a ampliação do conjunto de dados para diferentes condições sazonais, bem como a investigação de estratégias de refinamento de bordas e técnicas de pós-processamento que possam mitigar imprecisões observadas nos estágios de maior sobreposição foliar.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ALIF, M. A. R.; HUSSAIN, M. YOLOv12: A Breakdown of the Key Architectural Features. 2025. Disponível em: <https://arxiv.org/abs/2502.14740>.

AMARO, G. B. et al. Embrapa Hortaliças: recomendações técnicas para o cultivo de hortaliças em agricultura familiar. Brasília, DF: Embrapa Hortaliças, 2007.

ANDRADE, J. S. Segmentação de imagens para análise do crescimento de alface em sistema hidropônico por aprendizado profundo. 2026. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) – Universidade Tecnológica Federal do Paraná, Câmpus Medianeira, Medianeira, 2026. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/40527>.

ANDRADE, J. S.; CORDEIRO, A. F.; AIKES JUNIOR, J.; PAULA FILHO, P. L. Reconhecimento de alfaces em sistema hidropônico utilizando técnicas de processamento digital de imagens. In: CONGRESSO LATINO-AMERICANO DE SOFTWARE LIVRE E TECNOLOGIAS ABERTAS (LATINOWARE), 19., 2022, Evento Híbrido. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2022. p. 137-140. DOI: 10.5753/latinoware.2022.228035.

DIDONET, A. D. Estresse abiótico. Brasília, DF: Embrapa, 2009. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/697072/1/doc-239-p25.pdf>.

GÓIS, H.; OLIVEIRA, F.; OLIVEIRA, R.; PINTO, F.; AROUCHA, E.; QUEIROZ, G. C. M.; ALMEIDA, J.; OLIVEIRA, C. Lettuce growing in different hydroponic systems and nutrient concentrations of the nutrient solution. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 28, 2024. DOI: 10.1590/1807-1929/agriambi.v28n8e279686.

HERRERA-ARROYO, R.; MARTÍNEZ-NOLASCO, J.; BOTELLO-ÁLVAREZ, E.; SÁMANO-ORTEGA, V.; MARTÍNEZ-NOLASCO, C.; MORENO-AGUILERA, C. Smart Hydroponic Cultivation System for Lettuce (*Lactuca sativa* L.) Growth Under Different Nutrient Solution Concentrations in a Controlled Environment. *Applied System Innovation*, v. 8, p. 110, 2025. DOI: 10.3390/asi8040110.

MEIRELLES, A. F. M.; BALDOTTO, M. A.; BALDOTTO, L. E. B. Produtividade da alface (*Lactuca sativa* L.) em resposta à aplicação de ácidos húmicos e bactérias diazotróficas, em condições de campo. *Revista Ceres*, v. 64, p. 553-556, 2017. LUZ, J. M. Q.; GUIMARÃES, I. T. M. R.; KORNDORFER, G. H. Produção hidropônica de alface em solução nutritiva com e sem silício. *Horticultura Brasileira*, v. 24, n. 3, 2006.