COMPARATIVE ANALYSIS OF YOLOV3–YOLOV12 ARCHITECTURES FOR AUTOMATIC OIL PALM DETECTION IN AGRICULTURAL MONITORING

Matheus Ciscato Arnosti[1]*, Thiago Orlando Costa Barboza[1], Girley Valdes Fernandez[1], Eliseu Amaral[1], Layla Souza Pinto[1], Gustavo Lacerda da Silveira[1], Adão Felipe dos Santos[1],

[1] Department of Agriculture (DAG) Federal University of Lavras (UFLA), Lavras, Minas Gerais, Brazil.

A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil

Abstract.

Oil palm (Elaeis guineensis) is considered the most productive oilseed crop worldwide, and Brazil holds one of the greatest global potentials for palm oil production. Efficient monitoring of cultivated areas is therefore essential for proper crop management, enabling the detection of planting gaps, yield estimation, and decision-making support. In this context, computer vision techniques based on deep learning models, particularly those from the YOLO (You Only Look Once) family, have emerged as promising tools to improve the accuracy and efficiency of agricultural monitoring. This study aimed to evaluate the performance of different YOLO architectures for the automatic detection of oil palm plants, identifying the most suitable model for the evaluated production system.A total of 1,500 manually labeled images were obtained from the Roboflow platform and divided into training, validation, and test sets at proportions of 70%, 20%, and 10%, respectively. The models YOLOv3, YOLOv5, YOLOv8, YOLOv9, YOLOv10, YOLOv11, and YOLOv12 were trained using the medium configuration to maintain balance among hyperparameters. The training hyperparameters were set as follows: 500 epochs, batch size of 8, 2 workers, image size of 640×640 pixels, and early stopping with a patience of 70 epochs. Training was conducted on an RTX 5070TI GPU using the Google Colab environment. Model performance was evaluated using precision, recall, and mean Average Precision (mAP), considering an Intersection over Union (IoU) threshold of 0.50 (mAP50) and from 0.50 to 0.95 (mAP50–95).YOLOv12 achieved the best performance across all evaluated metrics, reaching 0.922 for precision, 0.901 for recall, 0.948 for mAP50, and 0.613 for mAP50–95. In contrast, YOLOv3 showed the lowest performance, with precision of 0.905, recall of 0.887, mAP50 of 0.926, and mAP50–95 of 0.569. The intermediate

---

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture. EXAMPLE: Last Name, A. B. & Coauthor, C. D. (2026). Title of paper. In Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture and Brazilian Congress on Precision and Digital Agriculture.

architectures (YOLOv5, YOLOv8, YOLOv9, YOLOv10, and YOLOv11) achieved mAP50 values ranging from 0.936 to 0.942 and mAP50–95 between 0.592 and 0.605, remaining up to 2.1 percentage points below YOLOv12. Recall values varied between 0.889 and 0.895, while precision ranged from 0.910 to 0.914. Analysis of training and validation loss curves indicated greater stability for YOLOv12, whereas YOLOv3 rapidly converged to overfitting.The superior performance of YOLOv12 may be associated with the incorporation of area-based attention mechanisms, which enhance the modeling of global spatial relationships and improve foreground activation accuracy. Overall, the results demonstrate that YOLOv12 is the most suitable architecture for automatic oil palm detection, presenting superior detection metrics and greater training stability. These findings highlight the continuous advancement of YOLO architecture and reinforce the potential of YOLOv12 as an efficient tool for agricultural monitoring applications.

Keywords. *Elaeis guineensis, deep learning, precision agriculture, computer vision*

INTRODUCTION

Oil palm (*Elaeis guineensis*) is considered one of the most productive oilseed crops worldwide, with high oil production per unit area compared with other crops used for vegetable oil production (Villela et al., 2014). In addition to its economic importance for the food, cosmetic, and energy sectors, oil palm cultivation stands out for its high agricultural yield, allowing greater production in smaller cultivated areas. Currently, palm oil supplies approximately 40% of the global demand for vegetable oils while using only about 5 to 5.5% of the area allocated to oilseed crops, highlighting its high productive efficiency and strategic importance for global food security (Zhao et al., 2024).

In the Brazilian context, although the country still represents a small share of global production, it has favorable edaphoclimatic conditions and large areas suitable for cultivation, especially in the Amazon region. Brazil is therefore frequently considered one of the countries with the greatest potential for the sustainable expansion of oil palm production (Villela et al., 2014). Among crop management activities, plant population estimation plays a fundamental role in agricultural planning, allowing the identification of planting gaps, the assessment of crop establishment, and the support of decisions related to crop management and productivity (Yang et al., 2024). Traditionally, this process is performed through field surveys, which require high operational effort, time, and costs, especially in large areas. In addition, manual methods are subject to observation errors and variations associated with the operator responsible for the assessment.

To enable automated monitoring of agricultural areas, computer vision has been used as an approach capable of extracting quantitative and qualitative information from digital images. When combined with machine learning and deep learning, computer vision allows computational models to learn visual patterns from previously annotated data, characterizing a supervised learning process focused on image classification, segmentation, or object detection (Zheng et al., 2022). In agriculture, these techniques have been applied to reduce dependence on manual evaluations, which are often time-consuming, subjective, or destructive, enabling faster and more standardized analyses of plants, fruits, grains, and disease symptoms (Fernandes et al., 2023). These approaches have shown efficiency in tasks such as image-based phenotyping, oil palm bunch classification, and detection of plant structures under field conditions, highlighting the potential of computer vision to support accurate and non-destructive agricultural monitoring (Pipitsunthonsan et al., 2023).

In recent years, the use of aerial images acquired by unmanned aerial vehicles (UAVs), combined with deep learning techniques, has emerged as a promising alternative to increase the efficiency of agricultural monitoring and contribute to improving global food productivity (Bouguettaya et al., 2022). The application of these technologies allows the automation of processes related to yield estimation, agricultural mapping, phenotyping, fruit quality assessment, and monitoring of plant development (Li et al., 2017; Fernandes et al., 2023). In different agricultural crops, image-based models have shown relevant results in oil palm bunch classification, wheat spike detection, estimation of morphological traits, and identification of disease symptoms, demonstrating the versatility of these approaches across different production systems (Pipitsunthonsan et al., 2023; Ye et al., 2023). When applied to oil palm cultivation, these approaches have the potential to replace or complement conventional methods based on visual inspections and manual counts, which often require long execution times, high operational costs, and greater subjectivity in evaluations, especially in extensive cultivated areas (Thapa et al., 2025).

In this context, computer vision techniques based on deep learning models, especially architectures from the YOLO (You Only Look Once) family, have been widely used in agriculture due to their ability to perform real-time object detection while maintaining high levels of accuracy in target localization and automatic classification (Bouguettaya et al., 2022). Deep learning can autonomously extract complex information from images, generating more accurate results (Shammi et al., 2023). However, these models still present several limitations, such as parameter redundancy, limited robustness, and low computational efficiency (Ding et al., 2024).

In oil palm plantations, recent studies have demonstrated the promising performance of YOLO models for plant detection, canopy classification, and support for the management of cultivated areas, with high values of precision, F1-score, and mAP (Ariyadi et al., 2023; Lee et al., 2024; Shaikh et al., 2024). However, despite these advances, most studies have focused on specific versions of the architecture, such as YOLOv5, YOLOv7, or YOLOv8, and there is still a lack of studies comparing different generations of the YOLO family under the same dataset and experimental conditions for the automatic detection of oil palm plants. This gap makes it difficult to identify the most suitable architecture for operational applications in agricultural monitoring, especially considering the rapid advancement of the most recent versions of the algorithm (Beqqali Hassani et al., 2026).

Thus, the objective of this study was to evaluate the performance of YOLOv3, YOLOv5, YOLOv8, YOLOv9, YOLOv10, YOLOv11, and YOLOv12 architectures for the automatic detection of oil palm plants using aerial images. The study was based on the hypothesis that more recent versions of the YOLO family would present better performance in detection metrics and greater generalization capacity than earlier architectures.

MATERIALS AND METHODS

Dataset

The dataset used in this study was obtained from the public PSSNet repository, available on the GitHub platform (Tong, 2021). Aerial images were cropped into dimensions of 640 × 640 pixels using code developed in Python within Visual Studio Code, resulting in 1,500 cropped images of oil palm plants. These images were manually annotated using the Roboflow platform with bounding boxes for plant identification, considering a single class named “D” (Figure 1). The dataset was divided into training, validation, and test sets at proportions of 70%, 20%, and 10%, respectively.

(A) (B)

Figure 1. Oil palm image annotation: (A) UAV-acquired image cropped to 640 × 640 pixels and (B) corresponding manual labeling using bounding boxes on the Roboflow platform.

Model Training

Seven architectures from the YOLO family were evaluated: YOLOv3, YOLOv5, YOLOv8, YOLOv9, YOLOv10, YOLOv11, and YOLOv12. All models were trained using the medium (m) configuration to maintain greater uniformity among the evaluated architectures and to enable a more consistent comparison of their performance.

Training was conducted using the following hyperparameters: a maximum of 500 epochs, batch size of 8, two workers, input image size of 640 × 640 pixels, and an early stopping mechanism with a patience of 70 epochs. Training was performed in the Google Colab environment using an NVIDIA RTX 5070 Ti graphics processing unit (GPU).

Evaluation Metrics

Model performance was evaluated using the metrics Precision, Recall, and Mean Average Precision (mAP), considering an Intersection over Union (IoU) threshold of 0.50 (mAP50) and the range from 0.50 to 0.95 (mAP50–95).

In addition to these metrics, train class loss and validation class loss curves were analyzed throughout the training process to assess convergence behavior and the generalization capability of the different models. A comparison between the actual and predicted number of detected plants for each model was also performed to evaluate their potential for plant monitoring and counting applications.

The Precision (P) metric represents the proportion of correctly classified positive detections relative to the total number of detections made by the model and was calculated according to Equation 1.

$$P = \frac{TP}{TP+FP} \tag{1}$$

where TP corresponds to the number of True Positives and FP to the number of False Positives.

The Recall (R) metric represents the model's ability to correctly identify the objects present in the images and is defined according to Equation 2.

$$R = \frac{TP}{TP+FN} \tag{2}$$

where FN represents the number of False Negatives.

To determine whether a detection was considered correct, the Intersection over Union (IoU) metric was used, which measures the degree of overlap between the predicted bounding box and the ground truth bounding box. IoU is calculated according to Equation 3.

$$IoU = \frac{Area(pred \cap real)}{Area(pred \cup real)} \tag{3}$$

IoU values closer to 1 indicate better overlap between the bounding boxes predicted by the model and the manually generated ground truth bounding boxes.

Based on the Precision and Recall values obtained at different IoU thresholds, the Average Precision (AP) metric was calculated, representing the area under the Precision–Recall curve. The overall performance of the models was evaluated using the Mean Average Precision (mAP) metric, which is commonly employed in object detection evaluation (Padilla; Netto; Silva, 2020).

In this study, two mAP indicators were considered: mAP50 and mAP50–95. The mAP50 was calculated using a single IoU threshold equal to 0.50, whereas mAP50–95 corresponds to the average of AP values obtained for IoU thresholds ranging from 0.50 to 0.95, with increments of 0.05. Therefore, mAP50 provides a simpler evaluation, whereas mAP50–95 adopts a more stringent criterion, as it requires greater precision in the positioning of the predicted bounding boxes.

In addition to the detection metrics, train class loss and validation class loss curves were analyzed throughout the training epochs. The analysis of these curves enables the evaluation of model convergence behavior and their generalization capability.

Finally, a comparison was performed between the actual and predicted plant counts for each evaluated model, allowing assessment of the models' potential for counting and monitoring applications in oil palm agricultural systems.

RESULTS AND DISCUSSION

Although the evaluated models showed similar performance across the evaluation metrics, a trend toward improved results was observed in the more recent versions of the YOLO family, especially YOLOv11 and YOLOv12, which achieved the best results for all metrics. In contrast, YOLOv3 was the only model that presented values below the overall average for all evaluated metrics, indicating lower performance compared with the more modern architectures.

The results demonstrate that all evaluated models achieved high performance in oil palm plant detection, as evidenced by the high mAP50, precision, and recall values (Figure 2). YOLOv12 obtained the best results across all analyzed metrics, reaching 0.948 for mAP50, 0.613 for mAP50–95, 0.922 for precision, and 0.901 for recall. In comparison, YOLOv3 presented values of 0.926, 0.569, 0.905, and 0.887 for mAP50, mAP50–95, precision, and recall, respectively.

Although the gains observed in mAP50, precision, and recall were relatively modest, the more pronounced increase in mAP50–95 indicates an improvement in the localization quality of the detected plants. Because this metric considers multiple Intersection over Union (IoU) thresholds, it provides a more rigorous assessment of the spatial accuracy of the bounding boxes. Thus, the results suggest that YOLOv12 not only correctly identifies a greater number of plants but also provides greater precision in delimiting their position in the images.

In addition, the high precision and recall values observed across all models indicate a good balance between false positives and false negatives, which is a fundamental characteristic for inventory and agricultural monitoring applications based on images acquired by unmanned aerial vehicles.

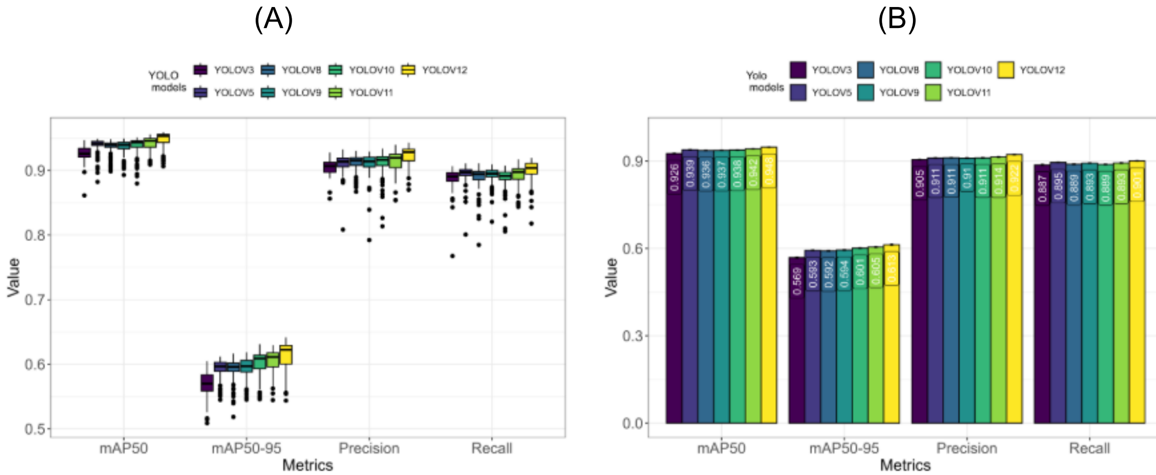

Figure 2. Variability analysis of model performance for precision, recall, mAP50, and mAP50–95 (A), and mean performance values for each model (B).

The improvement observed in the detection metrics with the advancement of YOLO family versions, especially in YOLOv11 and YOLOv12, may be associated with structural advances incorporated throughout the evolution of these models. Studies have shown that more recent YOLO versions include significant improvements in architecture and training methods, offering greater speed, accuracy, and efficiency through optimized training pipelines and designs that balance accuracy and performance while improving feature extraction with more advanced backbone and neck architectures (Beqqali Hassani et al., 2026).

In addition, although YOLOv12 showed slightly superior performance in the detection metrics, YOLOv11 achieved very similar results with a shorter training time, suggesting a better balance between accuracy and computational efficiency. While YOLOv11 required 2.44 hours to run 127 epochs, YOLOv12 required 3.48 hours to complete 150 training epochs.

The longer total training time observed for YOLOv12 is related not only to the greater number of epochs until interruption by the early stopping mechanism, but also to the longer processing time per epoch, suggesting greater computational complexity of the architecture. Recent studies comparing YOLOv11 and YOLOv12 models have also shown similar behavior, indicating that performance gains in more recent versions may occur at the cost of greater computational demand (Putra & Wowor, 2025).

These results can also be observed in Figure 3, which presents the class loss curves. All models showed convergence throughout training, with a reduction in the loss curves followed by stabilization. YOLOv11 and YOLOv12 showed greater stability between the training and validation curves, with lower divergence across epochs, suggesting better generalization capacity. In contrast, YOLOv3 showed greater separation between the curves and less stable convergence, indicating possible model overfitting and lower generalization capacity.

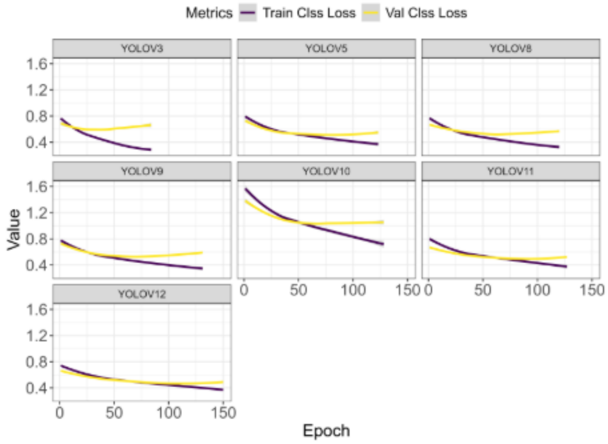

Figure 3. Training class loss and validation class loss curves across epochs for each model.

Another aspect evaluated was the relationship between the number of observed plants, obtained through manual annotation, and the number of plants predicted by the models (Figure 4). For this analysis, the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and the index of agreement (d) were considered. The results showed similar behavior among the models, with approximately 50% accuracy in plant counting considering the validation image set (150 images).

Among the evaluated models, YOLOv11 and YOLOv12 achieved the best performance for plant counting, with R^2 values of 0.530 and 0.521 and RMSE values of 5.104 and 5.153, respectively. In contrast, YOLOv10 showed the lowest performance, with an R^2 of 0.108 and an RMSE of 7.032, suggesting higher prediction errors, possibly associated with discrepant observations (outliers), which contributed to reducing the model's explanatory capacity.

Despite the slight advantage observed for YOLOv11, the differences relative to YOLOv12 were small, indicating similar performance between both models for the plant counting task. The observed differences may be related to architectural distinctions among the models, which influence feature extraction capability and sensitivity to false positives and false negatives during the detection stage.

(A) (B)

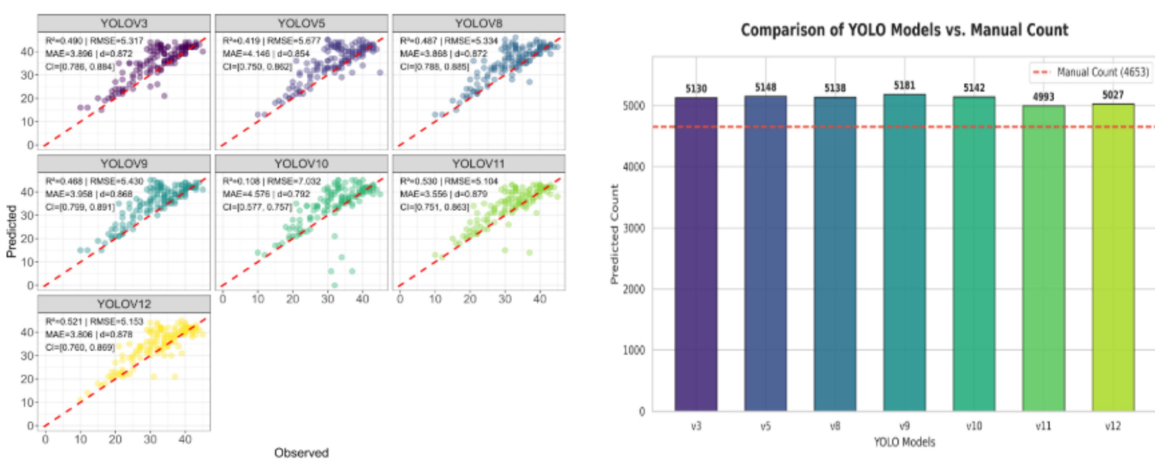

Figure 4. Scatter plots between observed and predicted values per image for each model (A), and comparison between the absolute number of observed and predicted plants for each model (B).

In addition to the architectural differences among the evaluated models, it is important to highlight that the performance of detection algorithms is also directly related to the quality of the annotations used as reference during training and validation. In aerial image datasets, the occurrence of plants that are partially visible at the image borders is common, and these plants are not always included in the manual annotation process. In such cases, even when the model correctly identifies the presence of these plants and generates coherent bounding boxes, these detections are counted as false positives because no corresponding annotation exists in the reference dataset.

This behavior may negatively affect metrics such as mAP50 and mAP50–95, since the evaluation is based exclusively on the comparison between model detections and manual annotations. Therefore, part of the difference observed between actual and predicted plant counts may be associated not only with model limitations but also with inconsistencies or criteria adopted during the annotation process.

The most pronounced difference among the models was observed for the mAP50–95 metric. Because this metric considers multiple Intersection over Union (IoU) thresholds, it provides a more rigorous assessment of the localization quality of detected objects. Thus, the results indicate that the more recent models not only correctly identify the presence of oil palm plants but also achieve greater precision in the spatial delimitation of the predicted bounding boxes (Figure 5). This characteristic may be particularly relevant for applications requiring precise individual localization, such as georeferenced inventories and automated monitoring of cultivated areas.

Figure 5: Comparison between manually annotated plants (green) and plants detected by the model (red) for an image from the dataset.

Another characteristic that facilitates oil palm plant identification and counting is the way plantations are managed. Oil palm cultivation commonly adopts spacing of eight meters between rows and nine meters between plants, significantly reducing canopy overlap. This spatial arrangement facilitates plant identification and counting, as well as the annotation process and the automatic detection performed by object detection models. In crops with higher plant density and greater canopy closure, overlap among plants may hinder individual delimitation and reduce model accuracy.

Similar results were reported by Farhan et al. (2025), who achieved identification performance of up to 95.6% using the YOLOv8 architecture for oil palm plant detection, reinforcing the potential of combining structural crop characteristics with computer vision techniques. Likewise, results obtained using YOLOv8 showed performance nearly equivalent to that of YOLOv10l. This finding highlights YOLOv8 as a suitable model for oil palm detection, enabling the differentiation of healthy, yellowing, small, and dead plants (Thapa et al., 2025).

Despite the excellent results reported in the literature, in the present study YOLOv8 was outperformed by YOLOv11 and YOLOv12.

CONCLUSION

The obtained results demonstrated that models from the YOLO family have strong potential for oil palm plant detection in production systems. Overall, a trend toward improved detection performance was observed as more recent versions of the architecture were evaluated, with YOLOv11 and YOLOv12 standing out among the tested models.

YOLOv12 achieved the best results across the evaluation metrics used, whereas YOLOv11 showed comparable performance with lower computational cost. In addition, the analysis of the training curves indicated convergent behavior and good generalization capability for the more recent models. However, despite the high values obtained for mAP50, precision, and recall, plant population estimation showed only moderate accuracy, with a tendency to overestimate the predicted number of plants. In this context, YOLOv11 presented the smallest difference between observed and predicted plant counts.

The evaluated architectures demonstrate potential for oil palm monitoring and plant counting applications, representing a promising alternative for agricultural monitoring and population inventory tasks. Although further improvements are required to increase the accuracy of population estimation, the results highlight the potential of computer vision techniques to automate processes traditionally performed manually, contributing to greater efficiency, reliability, and speed in agricultural crop monitoring.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Department of Agriculture of the Federal University of Lavras (UFLA), the Coordination for the Improvement of Higher Education Personnel (CAPES), the National Council for Scientific and Technological Development (CNPq) and the Research Support Foundation of the State of Minas Gerais (FAPEMIG) for their support and infrastructure.

REFERÊNCIAS

Ariyadi, M. R. N., Pribadi, M. R., & Widiyanto, E. P. (2023). Unmanned Aerial Vehicle for remote sensing detection of oil palm trees using You Only Look Once and convolutional neural network. In *2023 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)* (pp. 226–230). IEEE. https://doi.org/10.1109/EECSI59885.2023.10295670

Badgujar, C. M., Poulose, A., & Gan, H. (2024). Agricultural object detection with You Only Look Once (YOLO) algorithm: A bibliometric and systematic literature review. Computers and Electronics in Agriculture, 223. https://doi.org/10.1016/j.compag.2024.109090

Beqqali Hassani, I., Benhida, S., Lamii, N., Oqaidi, K., Ouiddad, A., & Ghiadi, S. (2026). From YOLO V1 to YOLO V11: Comparative analysis of YOLO algorithm (review). International Journal of Electrical and Computer Engineering (IJECE), 16(1), 450. https://doi.org/10.11591/ijece.v16i1.pp450-462

Bouguettaya, A., Zarzour, H., Kechida, A., & Taberkit, A. M. (2022). Deep learning techniques to classify agricultural crops through UAV imagery: A review. Neural Computing and Applications, 34(12), 9511–9536. https://doi.org/10.1007/s00521-022-07104-9

Ding, B., Zhang, Y., & Ma, S. (2024). A Lightweight Real-Time Infrared Object Detection Model Based on YOLOv8 for Unmanned Aerial Vehicles. Drones, 8(9). https://doi.org/10.3390/drones8090479

Farhan, M., Akhtar, M. N., & Bakar, E. A. (2025). Efficient real-time palm oil tree detection and counting using YOLOv8 deployed on edge devices. Journal of Umm Al-Qura University for Engineering and Architecture, 16, 1293–1308. https://doi.org/10.1007/S43995-025-00164-7

Fernandes, A. C. G., Figueiredo, J. A., Oliveira, A. C. B., Azevedo, A. M., Andrade Júnior, V. C., & Nascimento, I. R. (2023). Viabilidade da visão computacional no melhoramento genético na produção de raízes de batata-doce. *Horticultura Brasileira, 40*, 378–383.

Lee, S. S., Lim, L. G., Palaiahnakote, S., Cheong, J. X., Lock, S. S. M., & Ayub, M. N. B. (2024). Oil palm tree detection in UAV imagery using an enhanced RetinaNet. *Computers and Electronics in Agriculture, 227*, 109530. https://doi.org/10.1016/j.compag.2024.109530

Li, W., Fu, H., Yu, L., & Cracknell, A. (2017). Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. Remote Sensing, 9(1), 22. https://doi.org/10.3390/rs9010022

Padilla, R., Netto, S. L., & Silva, E. A. B. (2020). A survey on performance metrics for object-detection algorithms. International Conference on Systems, Signals and Image Processing, 237–242. https://doi.org/10.1109/IWSSIP48289.2020.9145130

Pipitsunthonsan, P., Pan, L., Peng, S., Khaorapapong, T., Nakasathien, S., Channumsin, S., & Chongcheawchamnan, M. (2023). Palm bunch grading technique using a multi-input and multi-label convolutional neural network. *Computers and Electronics in Agriculture, 210*, 107864. https://doi.org/10.1016/j.compag.2023.107864

Putra, G. A. A., & Wowor, A. D. (2025). Analisis perbandingan performa YOLO v11 dan v12 menggunakan model N dan S. Innovative: Journal of Social Science Research, 5(4), 8593–8601.

Shammi, S., Sohel, F., Diepeveen, D., Zander, S., & Jones, M. G. K. (2023). A survey of image-based computational learning techniques for frost detection in plants. In Information Processing in Agriculture (Vol. 10, Issue 2, pp. 164–191). China Agricultural University. https://doi.org/10.1016/j.inpa.2022.02.003

Shaikh, I. M., Akhtar, M. N., Aabid, A., & Ahmed, O. S. (2024). Enhancing sustainability in the production of palm oil: Creative monitoring methods using YOLOv7 and YOLOv8 for effective plantation management. *Biotechnology Reports, 44*, e00853. https://doi.org/10.1016/j.btre.2024.e00853

Thapa, A., Horanont, T., Neupane, B., et al. (2025). Classification of oil palm tree conditions from UAV imagery using the YOLO object detector. Big Earth Data, 9, 860–886.

Tong, P. (2021). PSSNet. GitHub. https://github.com/tongpinmo/PSSNet. Last accessed 14 June 2026.

Villela, A. A., Jaccoud, D. B., Rosa, L. P., & Freitas, M. V. (2014). Status and prospects of oil palm in the Brazilian Amazon. Biomass and Bioenergy, 67, 270–278. https://doi.org/10.1016/j.biombioe.2014.05.005

Yang, R., Yuan, D., Zhao, M., Liang, G., & Zhou, Y. (2024). *Camellia oleifera* tree detection and counting based on UAV RGB image and YOLOv8. *Agriculture, 14*(10), 1789. https://doi.org/10.3390/agriculture14101789

Ye, J., Yu, Z., Wang, Y., Lu, D., & Zhou, H. (2023). WheatLFANet: In-field detection and counting of wheat heads with high-real-time global regression network. *Plant Methods, 19*(1), 103. https://doi.org/10.1186/s13007-023-01079-x

Zhao, Q., Yu, L., Li, X., et al. (2024). Expansion and remaining suitable areas of global oil palm plantations. Global Sustainability, 7, e9. https://doi.org/10.1017/sus.2024.8

Zheng, W., Yan, L., Gou, C., Wang, F.-Y., & Fei-Yue Wang. (2022). Computational knowledge vision: Paradigmatic knowledge based prescriptive learning and reasoning for perception and vision. *Artificial Intelligence Review, 55*(8), 5917–5952. https://doi.org/10.1007/s10462-022-10166-9