

Application of you only look once algorithms to crop production management using unmanned aerial vehicles and computer vision systems

Amjed R. Al-Abbas¹, Łukasz Gierz^{2*}, Bashar S. Falih³,
Mustafa A.J. Al-Sammarraie⁴

¹ Network Engineering and Cybersecurity Department, College of Engineering, Al-Iraqia University, Baghdad, Iraq

² Institute of Machine Design, Faculty of Mechanical Engineering, Poznan University of Technology, Piotrowo 3, 60-965 Poznan, Poland

³ Department of Computer Techniques Engineering, Al Salam University College, Baghdad, Iraq

⁴ College of Agricultural Engineering Sciences, University of Baghdad, Baghdad, Iraq

* Corresponding author's e-mail: lukasz.gierz@put.poznan.pl

ABSTRACT

Global date palm production is steadily increasing and adopting technologies such as unmanned aerial vehicles (UAVs) and deep learning can reduce costs, save time, and improve productivity. To address this issue, the authors have proposed an innovative approach that uses UAVs for high-resolution aerial imaging. These images, collected by the Department of Computer Engineering at Al-Salam University in Baghdad and the Institute of Machine Design, Faculty of Mechanical Engineering, Poznan University of Technology, support improved orchard management, palm counting, and yield estimation. Precise spraying and pollination are also facilitated and accelerated, reducing overall cultivation costs. The proposed methodology involves processing captured images and applying three versions of the you only look once (YOLO) object detection algorithm, v11, v12, and YOLO-NAS – to determine the most effective model. The YOLOv12 model achieved the highest mAP@50 at 99.12%, which validates its superior performance in this application. The main innovation is the integration of deep learning-based palm crown detection with UAV imagery, enabling automated and scalable monitoring of palm plantations. The proposed methodology enables rapid, cost-effective, and scalable palm tree enumeration and management. A mobile application based on the trained model is planned to support real-time palm detection, yield estimation, and resource optimisation for farmers and stakeholders.

Keywords: aerial imaging, deep learning, palm Crowns, unmanned aerial vehicles, YOLO, yield estimation.

INTRODUCTION

The date palm tree (*Phoenix dactylifera* L.) is of profound cultural, economic, and environmental importance to the Middle East and North Africa (MENA) region. It is one of the oldest fruiting trees in the world and has been a staple crop in its base form. Egypt, Iran, Iraq, Saudi Arabia and the United Arab Emirates are global leaders in date production, producing around 67% of the world's production. The versatility of date palm is notable – not only are its fruits a key nutritional

resource for humans and livestock, but the tree also supports various industries, including those producing oil, wine, and natural fibres [1, 2].

In addition to its economic value, the date palm plays an important ecological role. It helps to sustain arid environments by providing shade and serving as a refuge for desert fauna. Beyond the fruit, other parts of the tree, such as leaves and bark, are utilized in the manufacture of cosmetics, paper, and construction materials. This combination of practical utility and cultural heritage highlights the enduring relevance and its multifaceted

contribution to the agricultural and ecological systems of the region [3].

In Iraqi trade, dates are one of the most important staple foods and form a major part of the local economy, as many farmers and exporters depend on this fruit as a major source of income. The total production value of agricultural production in 2003 was approximately \$139,688. However, it decreased in 2004 to \$30,859, which may be due to the impact of conflicts and socio-economic shocks in the region during that year. The value gradually increased in the following years to reach \$41,172 in 2005 and \$43,969 in 2006. A significant increase is observed by 2020, where the value reaches \$559,535 and then slightly decreases to \$487,388 in 2021. However, the total production value in 2022 reached about \$617,125, indicating good recovery and an increase in the value of date agricultural production of dates, which showed a good sign of recovery [4].

In recent years, modern technologies, particularly those based on computing and AI, have demonstrated remarkable success in various fields. For example, in the healthcare sector, artificial intelligence has been applied effectively for diagnostics and patient monitoring [5, 6].

In agriculture, unmanned aerial vehicles (UAVs) equipped with deep learning models have been used to detect and count olive trees with high precision [7].

Similarly, AI-based systems have achieved impressive accuracy in recognizing handwritten English and Arabic numerals, a task traditionally considered highly challenging for machines [8].

These achievements, many of which involve time-consuming tasks for humans, highlight AI's growing potential to solve complex problems efficiently. Based on these advances, this study proposes the application of AI tools to help detect and counting of date palm trees using UAVs [9, 10].

The objectives are tridimensional: (1) to promote the management of agricultural stock and resources; (2) to optimize the effectiveness of pest control and pollination; and (3) to develop the monitoring and early detection of disease or pest infestation. Furthermore, through continuous training and development, such systems should ideally be able to distinguish between male and female palm trees or even identify specific date palm species using crown morphology or canopy features.

Therefore, the correct identification and detection of palm trees are of utmost importance

for proper management in the context of enumeration and distribution analysis. Against this, the prediction of production volume and proper plantation management [10]. However, obtaining the correct statistics may be difficult and requires modern agricultural technology. Designing an efficient research method for personal palm tree studies is necessary for intelligent palm tree management [11]. They are time consuming, resource intensive and expensive [12, 13]. Remote sensing techniques, for example, satellite and drone imagery, are useful in monitoring palm plantations [14, 15]. Satellite remote sensing techniques have been increasingly used in the past two decades in a wide range of applications, such as land cover classification [16], palm plantation mapping [17], soil classification [18], counting [19], yield estimation [20], age estimation [21] and pest and disease detection [22]. Where satellite images are used, there are also difficulties caused by the cloud base [23], whereby palm trees are difficult to identify because the images obtained are of poor quality [24]. Drones are the best choice, as they are lightweight, compact, and not expensive. The drone typically has a high-resolution camera that can capture medium- to high-quality wide-angle images, depending on the flight altitude. Consequently, aerial photography has emerged as a preferred data source for monitoring palm plantations, offering cost savings, wide coverage, and access to remote areas [14, 25].

During the past decade, revolutionary advances in computer hardware and developments in AI technology have made novel ways of detection and feature extraction from an image possible [26]. One of the primary technologies in AI, convolutional neural networks (CNNs), has been at the forefront of object detection and image interpretation in computer vision. Traditional machine learning object detection has been used to identify palm trees and other objects from images collected from drones. It is a three-step procedure: image preprocessing, feature extraction, and classification. In [25] Bazi et al. employed scale-invariant feature transform (SIFT) features from key points of palm trees. Feature vectors were then drawn at each key point, which were provided as inputs to an extreme learning machine classifier for the identification of palm trees. The precision of the palm tree detection method was 91.11%. Manandhar et al. in [27] used the Circular Autocorrelation matrix for Polar Shape (CAPS) as a feature extractor and an support

vector machine (SVM) classifier for oil palm tree detection using shape features. The mean accuracy of the method was 84%. However, feature extractors in traditional object detection methods are hand-designed features defined by humans. It is difficult to make such a representation robust. The classification accuracy relies heavily on the set of features and classifier, which relies heavily on the set of data, and hence developing such systems manually is extremely difficult.

Another alternative to solving the above problem is deep learning. CNNs are deep learning-based algorithms that could extract millions of object's higher-level features and then use those features efficiently in object detection and object classification. Great progress has been made in the last few years to develop techniques for detecting geolocated objects on high-resolution remotely detected images, such as palm tree crowns [28]. These methods can be classified into two groups: two-stage object detection methods such as the Fast Region-based Convolutional Network method (Fast R-CNN) [29] and Mask R-CNN [30], and one-stage object detection methods such as you only look once (YOLO) v2/v3/v4 [31, 32], RetinaNet [33], and VGG-SSD [32]. Comparing the accuracy and efficiency of several tree crown detection techniques [32], some have found that two-stage object detection algorithms are superior to one-stage algorithms in accuracy, while the one-stage algorithm can significantly speed up tree crown detection [29]. One of the most well-known object detection techniques is the YOLO algorithm, presented by Redmon et al. [34]. YOLO is a single-stage technique that has CNNs as the foundation. The primary motivation behind YOLO was to get beyond the limitation of the then-existent two-stage object detection algorithms' detection time. YOLO variations have been largely utilized and implemented in various applications such as medicine [33], remote sensing [35], transportation [36], and agriculture [37].

Previous studies have shown that the YOLO-based approaches to object detection are effective in different types of images. Traditional methods of remote sensing in agriculture such as SIFT with extreme learning machine (ELM) or CAPS with SVM are less flexible to variations in illumination, occlusion, and scale in aerial images because they are based on handcrafted features. Although two-stage detectors, such as Mask R-CNN, have high accuracy, they are associated with high computational requirements. On the

contrary, YOLOv12 integrates speed and accuracy in a one-stage model, which is suitable for real-time UAV-based surveillance, particularly in edge computing and wide-area coverage.

The proposed research will improve palm tree detection by reducing the computational time in UAV imagery. We offer a quick and stable method of accurate palm location using YOLOv12, which combines the advantages of previous versions. The results are part of the work that helps to restore the palm industry and consequently the sustainability of its agriculture and culture.

DATASET DESCRIPTION

Aerial images and videos of a date palm plantation were captured using a camera mounted on a UAV that flew over the farm. These data were collected by the Department of Computer Techniques Engineering at Al-Salam University, Baghdad, Iraq, and the Institute of Machine Design, Faculty of Mechanical Engineering, Poznan University of Technology, Poland. The study area is a private orchard called Fadak, named after a historic oasis rich in dates. The plantation covers approximately 500 hectares (1.235 acres) and is managed by the Imam Hussein Shrine in the nearby holy city of Karbala. The geographical coordinates of the plantation are 32°43'19.5"N, 43°52'35.4"E. Palm trees are planted in a 5 × 5 m grid, with a total of 30,000 trees recorded in 2024, a number that continues to increase annually.

Image acquisition took place on November 5, 2024, between 11:30 a.m. and 1:00 p.m. under clear sky conditions. A high-resolution RGB camera specifically designed to monitor agricultural conditions was mounted on the UAV platform. As shown in Figure 1, the UAV was used to capture top-view images of palm crowns. Table 1 provides a complete overview of the technical specifications of the UAV used.

The characteristics listed in Table 1 enable high-resolution image acquisition suitable for precision farming.

The resolution of the captured images was 1080 × 1920 and 540 × 960 pixels with 8-bit depth per channel and spatial resolution of about 5–6 cm/pixel at altitudes of 100–310 metres (see Figure 2). Palm tree counting did not require geometric correction. The data set consisted of 195 images and 5 videos recorded by UAV, and the videos were divided into still frames with one

Table 1. Technical specifications of the UAV and camera system used for image acquisition

Component	Specification
UAV model	DJI Mini V3 Fly 4K
Flight altitude range	100–310 meters
Camera sensor	1/2.3" CMOS, 12.35 effective megapixels
Image resolution	3000 × 4000 pixels
Lens field of view	78.8°
Focal length	26 mm
Aperture	f/2.2
ISO range	100–1600
Distortion	<1.5%
Autofocus range	0.5 m to ∞
Image format	JPEG

frame per 2 seconds (total 495 images), assuming an average UAV speed of 6.1 m/s (22 km/h). To improve the classification, data augmentation increased the number of images to 1059. Data were saved on a Micro-SD card and uploaded to a PC to be processed. MATLAB[®] was used to analyse the images, train machine learning, and count them on an HP laptop with an Intel Core i7-7.5K U CPU (2.70–2.90 GHz), 12 GB RAM, and Windows 10 Pro.

DATA PREPARATION

To improve model resilience and generalization across varying visual conditions, an

end-to-end data preparation pipeline was used that included both preprocessing and enhancement. Operations were performed uniformly during training to simulate real-world object appearance and orientation variations, especially in agriculture where lighting and viewpoints can dramatically vary.

Preprocessing

The preprocessing step guaranteed uniform input quality and format. Each image was automatically oriented (Auto-Orient: Applied) from the embedded metadata, placing it in the correct orientation. Images were resized to a uniform resolution of $640 \times 640 \times 3$ pixels by using a stretch-based scaling method. This guaranteed uniform input size across the dataset and satisfied the architectural needs of the model.

Data augmentation

Data augmentation played an important role in augmenting the training set and reducing overfitting by generating diverse image versions. Each training image was subjected to a maximum of seven outputs of augmentation, including spatial, color, and geometric transformations:

- Flipping – horizontal and vertical flips were applied randomly to simulate mirrored views of the orchard.
- Rotations – images were rotated at 90° increments (clockwise, counterclockwise, and upside down) as well as at random angles



Figure 1. The DJI Mini V3 Fly 4K UAV was used to capture top-view images of the palm farm

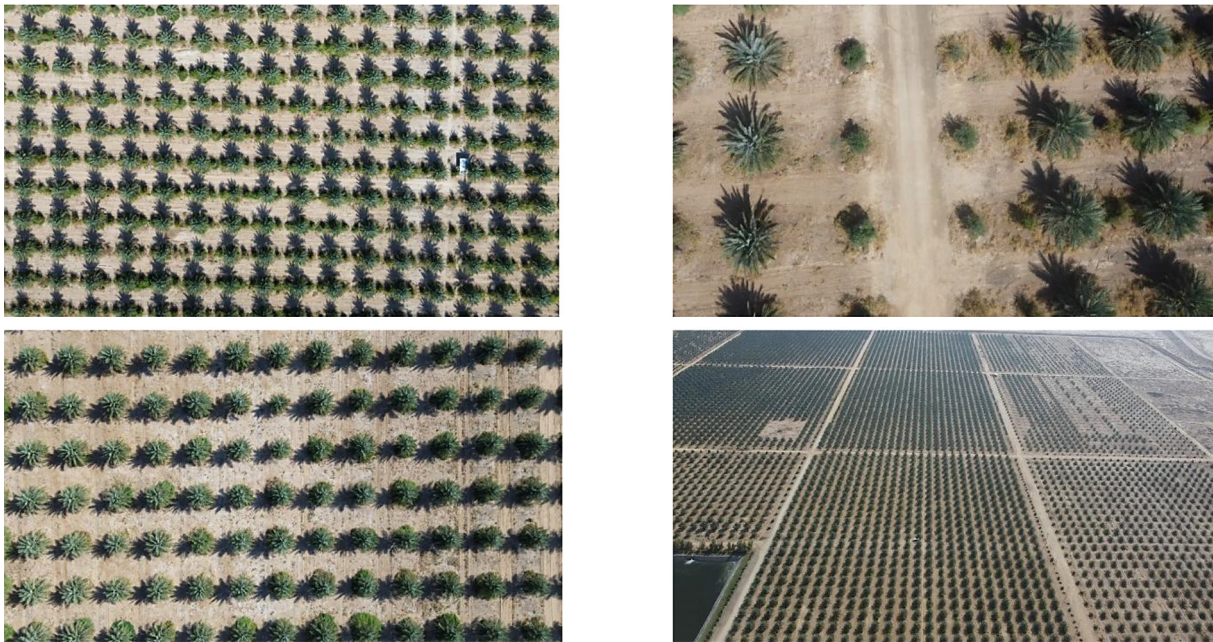


Figure 2. Four examples of UAV images of date palm crowns captured at different altitudes and angles at Fadak Farm, Karbala, Iraq

between -25° and $+25^\circ$, mimicking arbitrary UAV camera orientations.

- Grayscale conversion – to simulate poor lighting or sensor quality, 15% of the images were converted to grayscale.
- Bounding box transformations:
 - rotation – applied a mild bounding box rotation of $\pm 2^\circ$, accounting for subtle angular deviations;
 - shearing – introduced controlled shear transformations of $\pm 6^\circ$ horizontally and $\pm 7^\circ$ vertically, emulating perspective distortions;
 - brightness variation – the boundaries were subjected to brightness adjustments ranging from -15% and +15%, which relate to natural varying conditions of field illumination.

This synergy in augmentation greatly increased the diversity of datasets, through which the YOLO-based model was able to perform better against occlusions, lighting invariances, and object pose and scale variations.

METHOD

You only look once

The recent years have witnessed a paradigm change in object detection models with the YOLO series as a central pillar for real-time

visual perception-based applications. Originally designed by [34], the YOLO architecture initiated the single stage end-to-end detection stage through the combination of localization and classification into a single streamlined process. This revolution brought with it a new benchmark for speed and accuracy, which is crucial for applications based on robotics, surveillance, and agriculture.

With every repetition of the sequence, each occurrence had encountered upcoming issues in detection processes. YOLOv8 experienced the shift towards anchor-free architecture with decoupled detection head and improved multiscale feature extraction through parsed CSP backbone and Feature Pyramid Network. It was an expression of a consistent trade-off between accuracy and speed and was applicable for low-weight real-time applications [38].

YOLOv9 introduced the Generalized Efficient Layer Aggregation Network (GELAN) architecture and the programmable gradient information (PGI), which helped the model converge faster and generalize better, especially in difficult conditions. YOLOv10 also improved performance with a dual-assignment scheme, light heads, and decoupled down sampling – eliminating post-processing steps such as non-max suppression [39].

Following this trajectory, YOLOv11 employed cutting-edge modules such as spatial pyramid pooling-fast (SPPF) and Convolutional block

with Parallel Spatial Attention (C2PSA), improving detection accuracy, particularly in occlusion and varying illumination. It was also the fastest of all its versions, having an inference time of only 2.4 ms, and thus very compatible for real-time application [38].

YOLOv12, the most recent version, represents a significant leap forward in attention-driven design. It incorporates the Area Attention (A²) module for dynamic receptive field adjustment, along with Residual Efficient Layer Aggregation Network (R-ELAN) blocks that ensure robust gradient flow and efficient feature fusion. Additionally, flash attention mechanisms and adaptive multilayer perceptron (MLP) ratios were added to accelerate inference and enhance model precision [40].

The YOLOv12 model was implemented on the Python platform. The image labelling software was used to draw the outer rectangle over the palm trees in every image in the training set to complement hand-marking. The images were marked using the smallest rectangle that could fit around the palm trees to ensure that the rectangle touched the least amount of background. The Rectified Linear Unit (ReLU) activation function was used to optimize model performance can be calculated by Equation 1, and a gradient descent algorithm was used to adjust the neural network weights by Equation 2:

$$f(x) = \max(0, x) \quad (1)$$

$$w_{t+1} = w_t - \eta \cdot \nabla L(w_t) \quad (2)$$

where: w_t is the weight vector at iteration t ; η is the learning rate; $\nabla L(w_t)$ is the gradient of the loss function with respect to the weights.

Figure 3 shows the structure of the YOLOv12 model, where the final output of the neural network is the palm tree detection box.

As shown in Figure 3, the YOLOv12 model comprises four major modules: the input layer, the backbone network, the neck module and the head unit. The input layer processes the input images, which are normalized to 640×640 in RGB and are represented in the algorithm model by the orange background. The backbone network is a feature extractor that depends on the CSPDarknet53 architecture. It has two branching components for pixel feature extraction horizontally and vertically, as in the blue area. The neck module pools and processes features from the backbone. It performs multiscale feature fusion to robustly enhance the representation of objects at different scales, depicted in the green area. And finally, the head unit or the detection head holds the loss function and optimization method. This module utilizes the positive-negative sample matching method and auxiliary head training, enhancing performance and speeding up training through multidirectional branch, highlighted in the yellow area.

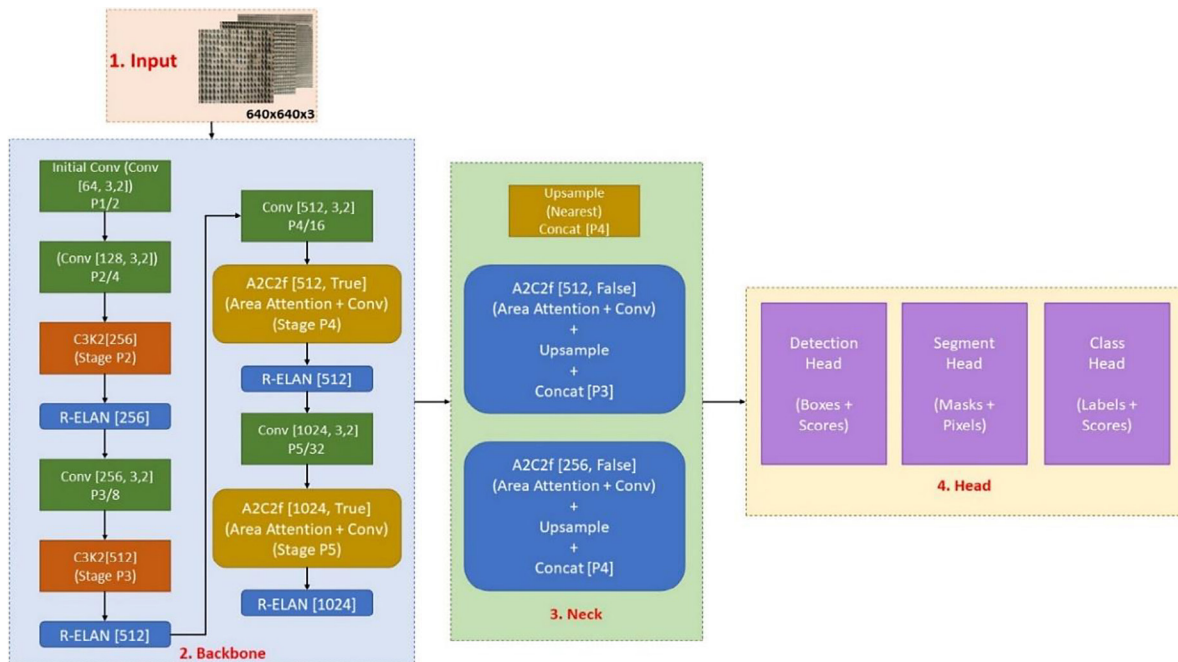


Figure 3. YOLOv12 network architecture

Dataset summary and model evaluation strategy

The experimental data set had a median image resolution of 1280×720 pixels, 8-bit RGB, and an average image resolution was about 0.92 megapixels. Based on these visual sources, 2784 annotated images, which are individual palm tree crowns, were manually annotated after augmentation. 66% for training (2.616 images, generated from 327 originals using 7 augmentations each), 17% for validation (84 images), and 17% for testing (84 images). In particular, augmentation was applied only to the training set after data split to prevent data leakage and ensure reliable evaluation. This methodology enabled consistent performance measures and prevented duplicates of the same image in different subsets. The proposed model for palm crown detection is displayed in Figure 4.

This dataset preparation and augmentation pipeline was uniformly applied to three different object detection architectures: YOLOv11, YOLOv12, and YOLO-NAS, to allow a fair and consistent comparative analysis. The goal of this evaluation was to identify the most accurate and efficient model for the specific task of automating the detection and counting of palms tree using UAV images. Through the preservation of consistent training environments and input preparation across all models, the study aimed to eliminate extraneous confounding variables that may exist under actual agricultural field conditions and

distill the architectural benefits of distinct versions of YOLO. Such a rigorous setup guarantees a valid platform for the subsequent performance results' interpretation with assurance.

RESULTS AND DISCUSSIONS

To compare the performance of object detection models in palm tree detection and counting from UAV images, three state-of-the-art architectures YOLOv11, YOLOv12, and YOLO-NAS were trained and tested under the same conditions. They were compared according to three standard metrics: mean average precision at 0.5 Intersection over union (mAP@50), precision, and recall as presented in Table 2.

Performance analysis

Among the three models, YOLOv12 demonstrated the best overall balance across all evaluation criteria. It achieved the highest mAP@50 value (99.12%), indicating exceptional ability to accurately localizing and classify palm trees across the test dataset. The precision of YOLOv12 (96.8%) suggests that the model consistently minimized false positives, which is particularly crucial in tree-counting applications where over-estimation can mislead agricultural planning. Although its recall (96.7%) was slightly lower than YOLO-NAS, the difference is marginal and compensated by higher localization accuracy.

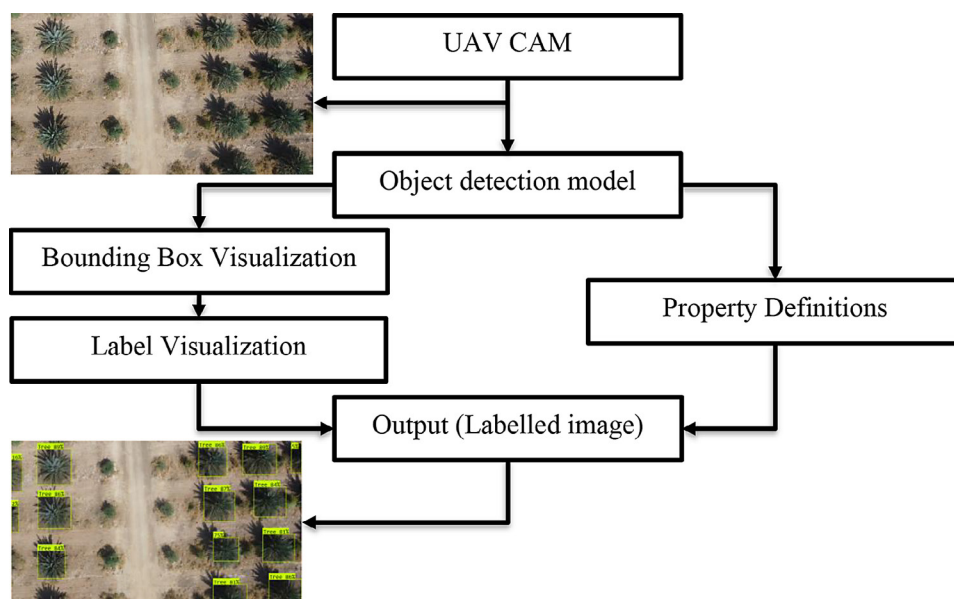


Figure 4. The proposed model for the detection of palm crowns

Table 2. The results of each YOLO model

Model	mAP@50 (%)	Precision (%)	Recall (%)
YOLOv11	99.01	95.7	97.3
YOLOv12	99.12	96.8	96.7
YOLO-NAS	98.5	90.4	99.2

The mathematical equations for precision, recall and mean average precision (mAP@0.5), as written in Equation 3.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

where: TP (true positives) – correctly predicted positive instances, FP (false positives) – incorrectly predicted as positive.

Precision is the proportion of correctly classified positive results among all predicted positive instances.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

where: FN (false negatives) – actual positives incorrectly predicted as negative, Recall – is the ability of the model to identify all instances pertaining to the dataset

Mean average precision (mAP@0.5)

For object detection, mAP@0.5 is typically calculated by first computing average precision (AP) for each class at IoU threshold = 0.5, then averaging over all classes:

- intersection over union (IOU):

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (5)$$


- average precision (AP) is the area under the Precision–Recall curve:

$$AP = \int_0^1 P(r) dr \quad (6)$$

where: $P(r)$ is the Precision as a function of Recall.

- mean average precision (mAP@0.5):

$$mAP@0.5 = \frac{1}{N} \sum_{i=1}^N AP_i \quad (7)$$

where: N – number of object classes.

- AP_i – average precision for class i , at IoU threshold 0.5.

In our study for single-class cases (eg, just palm crowns), mAP@0.5 is equal to AP for that class. Furthermore, the results revealed that detection accuracy improved significantly as the altitude decreased, allowing the camera to capture clearer and more detailed images of individual trees. As illustrated in Figure 9, lower flight altitudes resulted in minimal or no crown overlap between crown of palms, enhancing model performance. On the contrary, Figure 7 shows images captured from a higher altitude, where overlapping crowns were more frequent, which poses a notable challenge for detection models. This observation underscores a critical insight: Image resolution and viewing angle have a direct influence on detection reliability.

This finding also presents a practical challenge: In large-scale field surveys, flying at lower altitudes may not always be feasible due to time and coverage constraints. A promising solution lies in deploying higher-resolution cameras or multisensor payloads to compensate for altitude-related detail loss. Enhancing the optical quality of UAV-mounted sensors could therefore mitigate crown overlap issues and improve detection consistency, especially in densely planted orchards.

YOLOv11, although slightly behind YOLOv12 in overall precision, performed robustly with a mAP@50 of 99.01% and recall of 97.3%, demonstrating a superb capability of identifying most palm trees in the image accurately. It also demonstrated a high precision value (95.7%), reflecting its effectiveness for use in the field by real-time systems with little overhead computation.

On the other hand, YOLO-NAS presented an interesting trade-off. It achieved the highest recall score of 99.2%, reflecting its strength in minimizing missed detections, even under occlusion or complex background interference. However, its lower precision score (90.4%) reveals a tendency toward false positives, which may require further post-processing or threshold adjustments before operational deployment. Its mAP@50 of 98.5%, while still high, places it third in overall performance.

Visual interpretation

Figure 5 shows Precision–Recall curve for the YOLOv12 model. This curve helps to

visualize the trade-off between true positive rate and false positive rate.

In Figure 6 literates bar chart comparing mAP@50, precision, and recall for the three models. The chart clearly shows that YOLOv12 is the most balanced model, while YOLO-NAS leads in recall but falls behind in precision.

In Figure 7–9. Sample inference output on challenging images, showing detection bounding boxes from the YOLOv12 model on the output

image frame. The proposed model was used to detect individual palm tree crowns, and in scenarios with dense tree clusters or varying shadow conditions, YOLOv12 demonstrated precise and consistent bounding box placement.

Model selection consideration

From a practical deployment perspective, especially in UAV-based palm crowns census

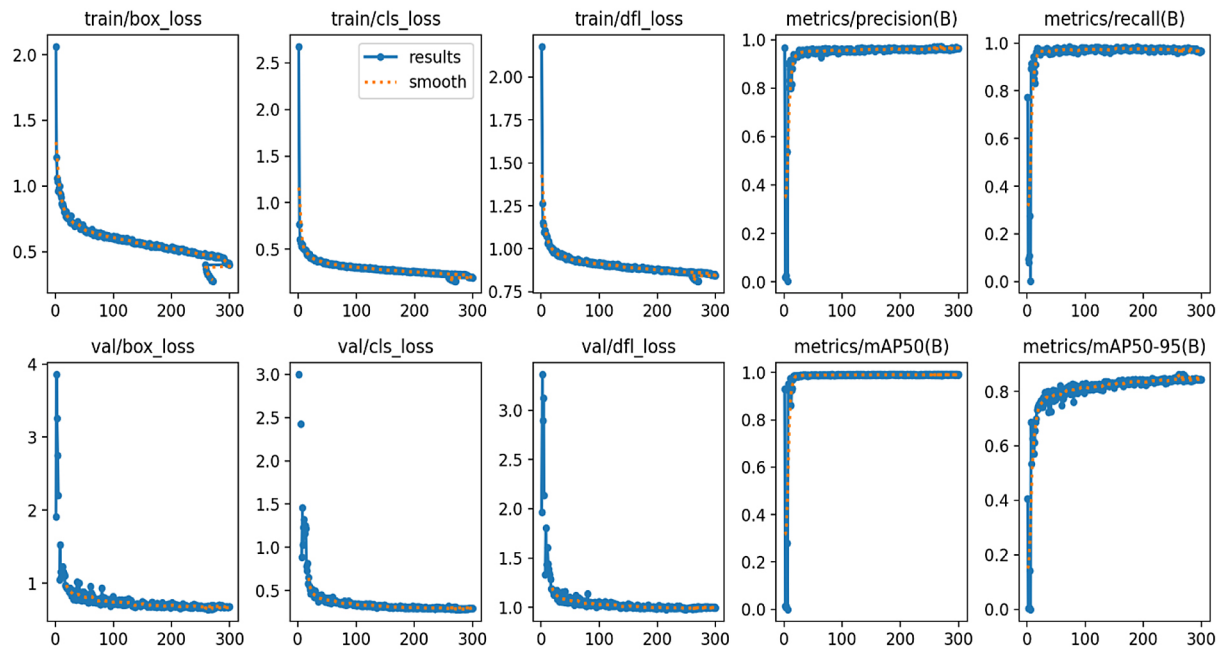


Figure 5. Training and validation performance of the YOLOv12 model for palm crown detection, illustrating steadily decreasing loss curves along with high Precision, Recall, and mAP scores – demonstrating effective learning and strong detection accuracy

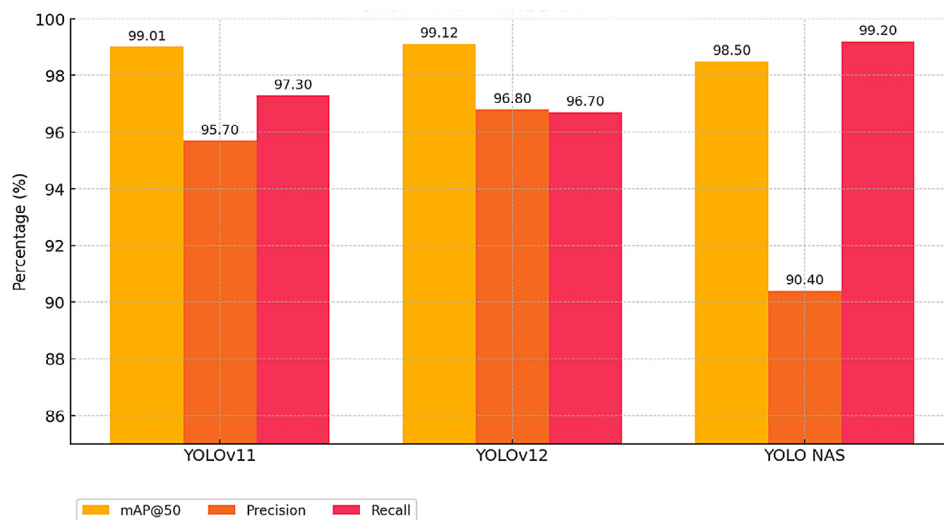


Figure 6. The performance comparison of YOLOv11, YOLOv12, and YOLO-NAS based on mAP@50, Precision, and Recall

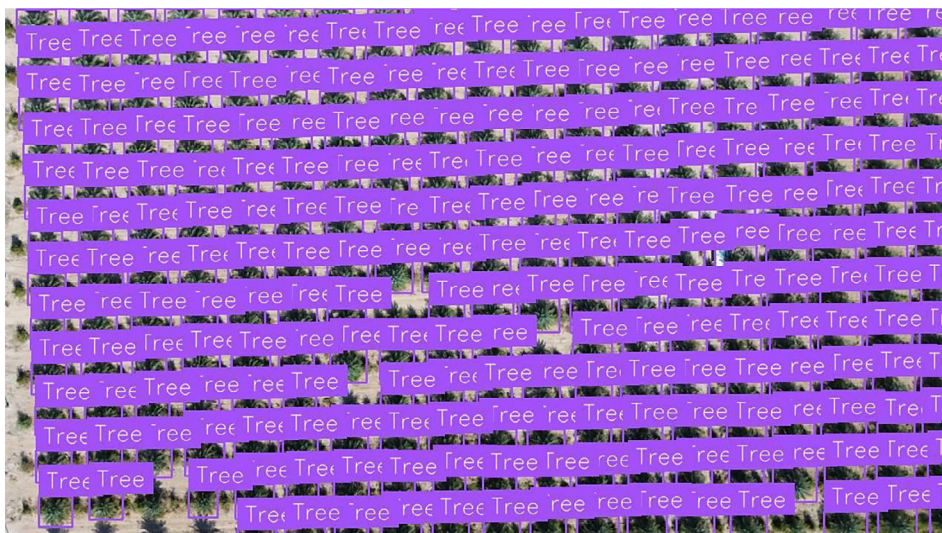


Figure 7. The example of the result of palm crowns detection from altitude 310 meters



Figure 8. The example of a result of palm crowns detection from altitude 160 meters



Figure 9. The example of the result of palm crowns detection from altitude 100 meters

operations, the selection of a model depends on application priorities.

- If Precision is paramount (eg, inventory accuracy, yield forecasting), YOLOv12 is the most reliable.
- To maximize detection coverage, particularly in 21 diverse lighting or occlusion scenarios, YOLO-NAS offers strong recall, but may require downstream filtering.
- YOLOv11 provides a favorable balance of speed and accuracy and is highly recommended for edge devices with constrained resources.

Ultimately, YOLOv12 emerges as the most suitable candidate for the palm crowns detection and counting task, given its superior precision-recall balance and leading detection accuracy. Its performance indicates robust generalization and strong potential for integration into agricultural monitoring systems.

CONCLUSIONS

This paper compared YOLOv11, YOLOv12, and YOLO-NAS in automated detection of date palm crowns based on UAV images. The overall performance of YOLOv12 (mAP@50: 99.12%, precision: 96.8%, recall: 96.7%) was the best, which is why it can be used in precision agriculture in real life. YOLOv11 provided a trade-off between accuracy and efficiency, and YOLO-NAS achieved the best recall (99.2%) but poor precision due to false positives in crowded scenes.

The study presented an innovative application of YOLOv12 enhanced with Area Attention and R-ELAN blocks and augmented to simulate orchard conditions. The model will be implemented in real time on mobile devices to facilitate field inventory.

The data did not have enough variety in environmental settings, including cloudy skies, partly sunny landscapes, different shadows, and orchard types. This diversity should be increased to improve the accuracy of detection and applicability to other tree species. Although some of these challenges were simulated using data augmentation to improve the performance of the model, it is important to collect more diverse images to improve the model further.

Future work will improve the robustness of the system in adverse conditions, increase the

size of data sets with seasonal and structural diversity, add depth sensing to increase accuracy, and optimise models to run in real-time on UAV edge devices.

Acknowledgments

The authors extend their sincere gratitude to the undergraduate students Dhyiaa H. Jeber, Rasha N. Sabri, A'laa H. Obaid, Zahra S. Jasim and Rasha H. Subhi for their valuable efforts and dedication to collecting and organizing the dataset used in this study. Special thanks are also due to Al-Salam University College and Poznan University of Technology for facilitating the administrative process and providing institutional support. The authors are particularly grateful to the management of the Fadak date palm plantation, affiliated with the Imam Hussein Holy Shrine, for their continued cooperation, logistical assistance, and support in granting access to the plantation site. Their collaboration was instrumental in the successful execution of the current study.

REFERENCES

1. Ali A. and Waly M. I., Nutritional and medicinal value of date fruit, 2012. [Online]. Available: <https://www.researchgate.net/publication/257416165>
2. Hajjaji Y., Boulila W., Farah I. R., and Koubaa A., Enhancing palm precision agriculture: An approach based on deep learning and UAVs for efficient palm tree detection, *Ecol Inform*, Mar. 2025, 85, 102952, <https://doi.org/10.1016/J.ECOINF.2024.102952>
3. Fattah A., Abd N., Eqbal R., and Radwan S., The state of the date palm (*Phoenix dactylifera*) in the Gaza Strip, Palestine: A Questionnaire-based Study.
4. World Food and Agriculture – Statistical Yearbook 2024. FAO, 2024. <https://doi.org/10.4060/cd2971en>
5. Al-Saffar B., Ali Y. H., Muslim A. M., and Ali H. A., ECG Signal Classification Based on Neural Network, 2023; 3–11. https://doi.org/10.1007/978-3-031-20429-6_1
6. Falih B. S., Sabir M. K., and Aydın A., Impact of sliding window overlap ratio on EEG-Based ASD diagnosis using brain hemisphere energy and machine learning, *Applied Sciences (Switzerland)*, Dec. 2024; 14(24), <https://doi.org/10.3390/app142411702>
7. Al-Tikriti O. A. A., Al-Saffar B. S. F., Bozdoğan A. M., and Arica S., Detection and Counting of Olive Trees in Images Taken by an Unmanned Aerial Vehicle, in 2022 4th International Conference on Current

- Research in Engineering and Science Applications (ICCRESA), IEEE, Dec. 2022; 75–79. <https://doi.org/10.1109/ICCRESA57091.2022.10352472>
8. Al-Saffar B., Al-Abbas A. R., and Özel S. A., A comparative study on the recognition of English and Arabic handwritten digits based on the combination of transfer learning and classifier, 2023; 95–107. https://doi.org/10.1007/978-3-031-20429-6_10
9. Mizher H. S., Taher M. A., Falih B. S., Gierz Ł. A., Warguła Ł., and Wieczorek B., Detecting and classifying media images of athletes using convolutional neural networks – case study: Individual sports images, *Advances in Science and Technology Research Journal*, Jun. 2025; 19(6), 152–166, <https://doi.org/10.12913/22998624/199798>
10. Falih B. S., Ali Y. H., Alabbas A. R., and Arica S., Optimising yield estimation for grapes: 61(2), 525–533, <https://doi.org/10.21162/PAKJAS/24.46>
11. Falih B. S., Gierz Ł., and Al-Sammarraie M. A. J., Fruit classification by assessing slice hardness based on RGB imaging. Case study: apple slices, *Journal of Applied Mathematics and Computational Mechanics*, Sep. 2024; 23(3), 7–18, <https://doi.org/10.17512/jamcm.2024.3.01>
12. Kipli K. et al., Deep learning applications for oil palm tree detection and counting, *Smart Agricultural Technology*, Oct. 2023; 5, 100241, <https://doi.org/10.1016/j.atech.2023.100241>
13. Istiak Md. A. et al., Adoption of Unmanned Aerial Vehicle (UAV) imagery in agricultural management: A systematic literature review, *Ecol Inform*, Dec. 2023; 78, 102305, <https://doi.org/10.1016/j.ecoinf.2023.102305>
14. Putra Y. C., Wijayanto A. W., and Chulafak G. A., Oil palm trees detection and counting on Microsoft Bing Maps Very High Resolution (VHR) satellite imagery and Unmanned Aerial Vehicles (UAV) data using image processing thresholding approach, *Ecol Inform*, Dec. 2022; 72, 101878, <https://doi.org/10.1016/j.ecoinf.2022.101878>
15. You H., Liu Y., Lei P., Qin Z., and You Q., Segmentation of individual mangrove trees using UAV-based LiDAR data, *Ecol Inform*, Nov. 2023; 77, 102200, <https://doi.org/10.1016/j.ecoinf.2023.102200>
16. Jean Bosco M., Jean Pierre R., Saleh Ali Muthanna M., Jean Pierre K., Muthanna A., and Abd El-Latif A. A., Neural Computing And Applications 1 Mgfeen: A Multi-Granularity Feature Encoding Ensemble Network for Remote Sensing Image Classification.
17. Gibril M. B. A., Shafri H. Z. M., Shanableh A., Al-Ruzouq R., Wayayok A., and Hashim S. J., Deep convolutional neural network for large-scale date palm tree mapping from UAV-based images, *Remote Sens (Basel)*, Jul. 2021; 13(14), 2787, <https://doi.org/10.3390/rs13142787>
18. Al-Sammarraie M. A. J., Al-Aani F., and Al-Mashhadany S. A., Determine, Predict and Map Soil pH Level by Fiber Optic Sensor, *IOP Conf Ser Earth Environ Sci*, Aug. 2023; 1225(1), 012104, <https://doi.org/10.1088/1755-1315/1225/1/012104>
19. Koon Cheang E., Koon Cheang T., and Haur Tay Y., Using Convolutional Neural Networks to Count Palm Trees in Satellite Images. [Online]. Available: <http://www.aarsb.com.my/wp->
20. Li W., Dong R., Fu H., and Yu L., Large-scale oil palm tree detection from high-resolution satellite images using two-stage convolutional neural networks, *Remote Sens (Basel)*, Dec. 2018; 11(1), 11, <https://doi.org/10.3390/rs11010011>
21. Mubin N. A., Nadarajoo E., Shafri H. Z. M., and Hamedianfar A., Young and mature oil palm tree detection and counting using convolutional neural network deep learning method, *Int J Remote Sens*, Oct. 2019; (19), 7500–7515, <https://doi.org/10.1080/01431161.2019.1569282>
22. Zheng J. et al., Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images, *ISPRS Journal of Photogrammetry and Remote Sensing*, Mar. 2021; 173, 95–121, <https://doi.org/10.1016/j.isprsjprs.2021.01.008>
23. Adam F., Mönks M., Esch T., and Datcu M., Cloud Removal in High Resolution Multispectral Satellite Imagery: Comparing Three Approaches, in *The 2nd International Electronic Conference on Remote Sensing*, Basel Switzerland: MDPI, Mar. 2018; 353. <https://doi.org/10.3390/ecrs-2-05166>
24. Colefax A. P., Butcher P. A., and Kelaher B. P., The potential for unmanned aerial vehicles (UAVs) to conduct marine fauna surveys in place of manned aircraft, *ICES Journal of Marine Science*, Jan. 2018; 75(1), 1–8, <https://doi.org/10.1093/icesjms/fsx100>
25. Bazi Y., Malek S., Alajlan N., and AlHichri H., An automatic approach for palm tree counting in UAV images, in *2014 IEEE Geoscience and Remote Sensing Symposium*, IEEE, Jul. 2014; 537–540. <https://doi.org/10.1109/IGARSS.2014.6946478>
26. Gierz Ł., Al-Sammarraie M. A. J., Özbek O., and Markowski P., The use of image analysis to study the effect of moisture content on the physical properties of grains, *Sci Rep*, Dec. 2024; 14(1), <https://doi.org/10.1038/s41598-024-60852-7>
27. Manandhar, Hoegner L., and Stilla U., Palm Tree Detection Using Circular Autocorrelation of Polar Shape Matrix, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Jun. 2016; III–3, 465–472, <https://doi.org/10.5194/isprs-annals-III-3-465-2016>
28. Li K., Wan G., Cheng G., Meng L., and Han J., Object detection in optical remote sensing images: A survey and a new benchmark, *ISPRS Journal of Photogrammetry and Remote Sensing*, Jan.

- 2020; 159, 296–307, <https://doi.org/10.1016/j.isprsjprs.2019.11.023>
29. Chen Z. Y. and Liao I. Y., Improved Fast R-CNN with Fusion of Optical and 3D Data for Robust Palm Tree Detection in High Resolution UAV Images, *Int J Mach Learn Comput*, Jan. 2020; 10(1), 122–127, <https://doi.org/10.18178/ijmlc.2020.10.1.908>
30. Ocer N. E., Kaplan G., Erdem F., Kucuk Matci D., and Avdan U., Tree extraction from multi-scale UAV images using Mask R-CNN with FPN, *Remote Sensing Letters*, Sep. 2020; 11(9), 847–856, <https://doi.org/10.1080/2150704X.2020.1784491>
31. Ammar A., Koubaa A., and Benjdira B., Deep-learning-based automated palm tree counting and geolocation in large farms from aerial geotagged images, *Agronomy*, Jul. 2021; 11(8), 1458, <https://doi.org/10.3390/agronomy11081458>
32. Xia M., Li W., Fu H., Yu L., Dong R., and Zheng J., Fast and robust detection of oil palm trees using high-resolution remote sensing images, in *Automatic Target Recognition XXIX*, T. L. Overman and R. I. Hammoud, Eds., SPIE, May 2019; 11. <https://doi.org/10.1117/12.2518352>
33. Culman M., Delalieux S., and Van Tricht K., Individual palm tree detection using deep learning on RGB imagery to support tree inventory, *Remote Sens (Basel)*, Oct. 2020; 12(21), 3476, <https://doi.org/10.3390/rs12213476>
34. Redmon J., Divvala S., Girshick R., and Farhadi A., You Only Look Once: Unified, Real-Time Object Detection, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Jun. 2016; 779–788. <https://doi.org/10.1109/CVPR.2016.91>
35. Ma H., Liu Y., Ren Y., and Yu J., Detection of collapsed buildings in post-earthquake remote sensing images based on the improved YOLOv3, *Remote Sens (Basel)*, Dec. 2019; 12(1), 44, <https://doi.org/10.3390/rs12010044>
36. Bekele B. and Kekeba K., Developing traffic congestion detection model using deep learning approach: a case study of Addis Ababa City Road, Nov. 25, 2020. <https://doi.org/10.21203/rs.3.rs-113234/v1>
37. Liu G., Nouaze J. C., Touko Mbouembe P. L., and Kim J. H., YOLO-Tomato: A Robust Algorithm for Tomato Detection Based on YOLOv3, *Sensors*, Apr. 2020; 20(7), 2145, <https://doi.org/10.3390/s20072145>
38. Sapkota R., Meng Z., Churuvija M., Du X., Ma Z., and Karkee M., Comprehensive Performance Evaluation of YOLOv12, YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments, Jul. 2024, [Online]. Available: <http://arxiv.org/abs/2407.12040>
39. Zhang T., Luo B., Sharda A., and Wang G., Dynamic label assignment for object detection by combining predicted IoUs and anchor IoUs, *J Imaging*, Jul. 2022; 8(7), 193, <https://doi.org/10.3390/jimaging8070193>
40. Tian Y., Ye Q., and Doermann D., YOLOv12: Attention-Centric Real-Time Object Detectors, Feb. 2025, [Online]. Available: <http://arxiv.org/abs/2502.12524>