Indian Journal of Agricultural Sciences **95** (6): 678–683, June 2025/Article
https://doi.org/10.56093/ijas.v95i6.161451

Mango (*Mangifera indica*) tree detection and counting in mango orchard with satellite images using deep learning model YOLO: A comparative analysis

LALIT BIRLA[1], ANSHU BHARADWAJ[1]*, RAJNI JAIN[2], CHANDAN KUMAR DEB[1], VINAY KUMAR SEHGAL[1] and RAMASUBRAMANIAN V[3]

ICAR-Indian Agricultural Research Institute, New Delhi 110 012, India

Received: 29 November 2024; Accepted: 2 March 2025

ABSTRACT

Mango (*Mangifera indica* L.) is a widely cultivated horticultural cash crop in tropical and subtropical regions, valued for its exceptional taste, aroma, nutritional benefits, and medicinal properties. Tree counting is a crucial aspect of orchard inventory management, enabling efficient resource allocation, yield estimation, and precision agriculture applications. However, traditional methods often rely on manual efforts or expensive feature engineering, leading to errors, inefficiencies, and limited scalability. Recent advancements in deep learning-based approaches have demonstrated state-of-the-art performance in automated tree counting, offering improved accuracy, robustness, and computational efficiency. The study was carried out during 2023–24 presenting a comparative evaluation of YOLO architectures for mango tree detection and counting. The research analyzes YOLOv5, YOLOv6, YOLOv7, and YOLOv8 using satellite remote sensing imagery from the Bulandshahr district of Uttar Pradesh. Performance evaluation is conducted using precision, Recall, F1-score, and mean average precision (mAP). Experimental results reveal that YOLOv8 exhibits superior performance, achieving a well-balanced trade-off between detection accuracy, processing speed, and generalization. These findings highlight the potential of deep learning models for scalable orchard monitoring, precision agriculture, and sustainable fruit production.

Keywords: Computer vision, Deep learning, Mango, Mango tree detection, Mango tree counting, YOLO

Mango (*Mangifera indica* L.) is important and widely cultivated fruit crops in India. With an annual production of 20.77 Mt from an area of 2.35 million hectares, India is the world's largest producer of mangos (National Horticulture Board 2023). Mango is India's most important commercial fruit crop, accounting for more than 54% of global mango production (Srividhya *et al.* 2024). A crucial tool for managing orchards is orchard inventory. Orchard inventory supports a wide range of management activities, including efficient resource allocation, disease control, harvest planning, and targeted treatments. By maintaining accurate and comprehensive orchard inventory, orchard managers may promote sustainability. The orchard tree inventory may benefit government insurance schemes in addition to orchard management. In India, mango orchard inventory includes number of trees to assess the production potential of mango orchards. For large mango orchard sites, tree counting in mango orchard is costly and time consuming. Combination of remote sensing and deep learning tools is game changing approach for tree counting in mango orchard. This approach can also help in estimating crop yields, planning harvesting operations, and making informed decisions regarding fertilizer, irrigation, and pest control.

Computer algorithms have proven highly effective in autonomously detecting and pinpointing trees as distinct objects within images (Putra and Wijayanto 2023). Among these algorithms, the You Only Look Once (YOLO) (Liu *et al.* 2024) architecture has established itself as a highly efficient and precise framework for performing object detection tasks (Redmon *et al.* 2015).

This study provides in-depth performance assessment of multiple YOLO architectures YOLOv5, YOLOv6, YOLOv7, and YOLOv8 for the tasks of mango tree detection and tree counting utilizing remote sensing imagery. Notably, YOLOv8 is highlighted as a key focus of this evaluation. This study uses a variety of measures, including as recall, precision, F1-score, loss, and mean average precision (mAP), to assess the performance of the examined architectures. Additionally, this research utilizes satellite images of Bulandshahr district, Uttar Pradesh, India. These images were collected from Google Earth Engine. In this study all models showed great accuracy.

[1]ICAR-Indian Agricultural Research Institute, New Delhi; [2]ICAR-National Institute of Agricultural Economics and Policy Research, New Delhi; [3]ICAR-National Academy of Agricultural Research Management, Hyderabad. *Corresponding author email: ansdix@gmail.com

However, YOLOv8 outperformed the earlier YOLO versions by achieving the highest scores in precision, recall, F1-score, and mAP@50.

MATERIALS AND METHODS

Study area: The study was carried out during 2023–24. Uttar Pradesh, Andhra Pradesh, Karnataka, Maharashtra, Bihar, Gujarat, Tamil Nadu, Odisha, West Bengal, and Jharkhand are the major Indian states that produce mangoes. Western Uttar Pradesh, with its favourable agro-climatic conditions, serves as a significant region for mango production, offering a diverse range of mango varieties. The major mango producing districts in Uttar Pradesh are Lucknow, Sahranpur, Bareilly, Bulandshahr, Unnao, Muzzaffarnagar, Sultanpur, Meerut and Firozabad (Supplementary Fig. 1), these districts are known for their favourable climatic conditions and suitable soil for mango cultivation. Bulandshahr (28.000’– 28.040’ N latitude and 77.030’– 78.030’ E longitude) is one of the districts which was selected for this study. Total mango orchard area in Bulandshahr district is 15–16 thousand hectares with 2,65,000 Mt of mango production. Different-aged mango orchards (Cultivar: Daseri, Langra, Chausa) of 2–50 years old planted in Bulandshahr.

The selection of Bulandshahr as the study area is based on several factors, including the ease of access and convenience for the investigator. The region is known for its large-scale mango orchards, which supply mangoes to the nearby Delhi NCR market. The straight-line distance between Delhi and Bulandshahr is 80.1 km. Additionally, the investigator's familiarity with the local language and close association with the area's people and officials, who are directly or indirectly involved in mango cultivation, played a significant role in choosing this location.

Data acquisition and preparation: High Resolution remote sensing images of study area by Google Earth Engine (GEE) were used for this study. Data was collected from 14 location of Bulandshahr district (Gyaspur, Nimkhera, Chewali, Bahalimpur, Kazapur, Ranapur, Waira Firozpu, Sahanpur, Siyana, Bulandshahr, Rampura, Chandpur, Bigaur and Ghansoorpur) through GEE. The images collected were within the visible spectrum and had a resolution of 1162 × 632 pixels. These images, which depicted various sections of the orchard, feature 4 different varieties of mango trees (Alphonso, Dasheri, Langra, and Chausa), included both young and mature trees. The images were carefully selected to create a varied dataset. Any images with ambiguous tree appearances, identified through visual inspection, were cropped to eliminate the uncertainty. In total, 750 images were gathered from 14 different locations across the Bulandshahr district in Uttar Pradesh.

Ground truth labelling was done manually while inspecting images of the corresponding study area. Mango tree crown pixels were marked with a green box in the ground truth images for individual tree detection, while all other pixels belonged to the background class. The JPG files were used to hold the ground truth photographs. Fig. 1 displays the green regions in the ground truth, which represent individual mango tree crowns, each enclosed within a bounding box. These bounding boxes, known as annotation boxes, are used to identify and assess the detection and counting of individual tree crowns. Each annotation box corresponds to one distinct mango tree crown, serving as a key element for evaluation. The green areas in Fig. 1 are the ground truth's individual mango tree crowns, each of which is bounded by a box. Individual tree crowns are detected and counted using these bounding boxes, often referred to as annotation boxes. An important component for evaluation is the annotation box, which represents a single, unique mango tree crown.

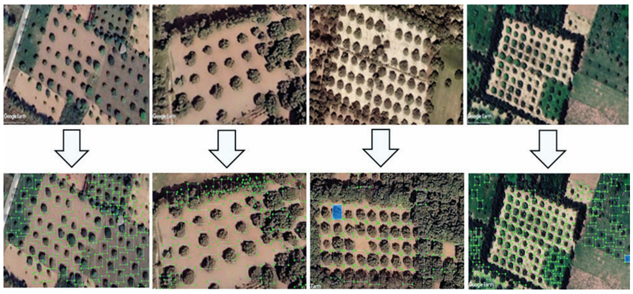

Fig. 1 Mango orchard image with bounding box.

Architectures under study

YOLOv5: “You Only Look Once version 5”, YOLOv5 (Zhao *et al.* 2019) has rapidly established itself as a leading solution in the object detection field due to its continuous performance and speed optimization. The YOLOv5 model (Supplementary Fig. 2) is made up of three essential components; the head, neck, and backbone. The backbone of YOLOv5 utilizes the CSP-Darknet53 convolutional network and the Cross Stage Partial (CSP) network structure (Itakura and Hosoi 2020, Mekhalfi *et al.* 2021). This approach successfully addresses problems with duplicate gradients and vanishing gradients while guaranteeing a strong information flow, particularly in deep layers (Mathew and Mahesh 2022). The YOLOv5 model's neck implements Path Aggregation Network (PANet) and uses a variant of the Spatial Pyramid Pooling (SPP). As with its predecessors, the YOLOv5 model's head, which consists of three convolution layers, is the last component. Similar to earlier versions, the YOLOv5 model's head is made up of three convolution layers that predict object classifications, scores, and bounding box coordinates (Mathew and Mahesh 2022).

YOLOv6: YOLOv6 (Jiang *et al.* 2022) model, which consists of the Backbone, Neck, and Head (Supplementary Fig. 3). To further increase its uniqueness in the field, YOLOv6 sets itself apart by presenting an anchor free model with a reparameterized backbone (Jiang *et al.* 2022). YOLOv6 introduces reparameterized backbones to address the competing demands of speed and accuracy that are frequently present in linear networks like VGG and classic multi-branch networks like ResNets (Gupta *et al.* 2023). YOLOv6 has a reparametrized backbone called EfficientRep, while medium and large models utilize

CSPStackRep and nano and tiny models have RepVGG. The neck structure resembles that of YOLOv5, but it has a separated classification and detection head and bi-directional concatenation for improved localization accuracy (Jiang *et al.* 2022).

YOLOv7: YOLOv7 (Wang *et al.* 2023) is also composed of three primary parts: the head, neck, and backbone (Supplementary Fig. 4). The ultimate objective of YOLOv7's development was to develop a network architecture that could predict bounding boxes more accurately than other models of a similar kind while still retaining a similar inference speed (Wang *et al.* 2023).

One notable advancement in YOLOv7 is the incorporation of the Extended Efficient Layer Aggregation Network (E-ELAN), which is an optimized version of the ELAN computational block (Patel *et al.* 2022, Wang *et al.* 2023). This enhancement improves the efficiency of the convolutional layers within the YOLO network's backbone by organizing computational blocks without altering the transition layers. It aims to shorten the gradient's backpropagation distance in order to increase the network's learning efficiency, while also accounting for the memory requirements for layer retention (Gupta *et al.* 2023).

YOLOv8: YOLOv8 (Sohan *et al.* 2024) retains the core architectural components of the YOLO framework, namely the Backbone, Neck, and Head, which work synergistically to perform various tasks (Supplementary Fig. 5). A notable feature of YOLOv8 is its adoption of an anchor-free model, which distinguishes it from previous YOLO versions that utilized anchor boxes for object detection. This anchor-free approach significantly speeds up the Non-Maximum Suppression (NMS) phase, a crucial post-processing step that filters through potential detections after the model makes its predictions, by reducing the total number of box predictions.

YOLOv8 also introduces key advancements in its convolutional operations, which are the foundational elements of its neural network structure. The architecture has been made more efficient and adaptable with the replacement of C3 with C2f, as well as replacing of the original 6×6 convolution in the core with a more efficient 3×3 convolution. To improve the model's capacity to identify objects of different sizes, YOLOv8 integrates the Spatial Pyramid Pooling Feature (SPPF), improving its multi-scale feature extraction. Furthermore, the model employs an image augmentation approach, such as mosaic augmentation, during training. This method combines four images into one, allowing the model to learn objects in various locations with variable surrounding pixel arrangements, and under partial occlusion (Tamang *et al.* 2023).

Performance evaluation: Precision, recall, accuracy, loss, and mean average precision (mAP) are among the frequently used metrics that we have chosen from the literature to assess model performance and provide useful comparisons. We then give a thorough explanation of every metric that was chosen.

Precision: Precision, as discussed by Padilla *et al.* (2020), is a key metric in evaluating the accuracy of positive predictions made by object detection models.

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}} \tag{1}$$

Where TP, Count of true positive predictions; and FP, Count of false positives, or instances incorrectly classified as positive.

Recall: Recall (Padilla *et al.* 2021) in equation 2, also known as the true positive rate or sensitivity, measures the proportion of actual positive instances correctly identified by the model, ensuring that true detections are maximized.

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}} \tag{2}$$

Where TP, Count of true positive predictions; and FN, Count of false negatives, or instances incorrectly classified as negative.

F1-Score: A balanced statistic that integrates accuracy and recall into a single number is the F1-score (Zhao and Li 2020). It gives a general indication of how accurate the model is in detecting objects.

$$\text{F1–Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision + Recall}} \tag{3}$$

Accuracy: It shows what percentage of all instances both positive and negative were accurately anticipated (Equation 4). The model's overall performance across all classes is gauged by accuracy.

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \tag{4}$$

Where TP, Count of true positive predictions; TN, Count of true negative predictions; FP, Count of false positives or instances incorrectly classified as positive; FN, Count of false negatives, or instances incorrectly classified as negative.

Mean average precision(mAP@50): mAP (Zhu *et al.* 2020) computes the average of the average precision (Equation 5) values over different recall thresholds with a 50% IoU (Intersection Over Union) for each class in order to provide a comprehensive assessment of the model's performance across a number of classes.

$$\text{mAP} = \frac{1}{N}\sum_{i=1}^{N} \text{AP}_i \tag{5}$$

Loss: Loss, as described by Casas *et al.* (2023), is a metric that quantifies the difference between a model's predicted outcomes and the actual or expected values. A smaller loss value signifies closer alignment between the model's predictions and the actual data. The primary objective during training is to minimize this loss to enhance the model's performance. Different YOLO versions incorporate distinct loss functions to optimize various aspects of detection, including box loss, objectness (obj) loss, distributional focal loss (dfl), and classification (cls) loss.

To determine whether observed performance differences between YOLO models are statistically significant, we conducted a pairwise t-test along with 95% Confidence

Intervals (CIs). This statistical analysis provides an in-depth validation of performance comparisons.

RESULTS AND DISCUSSION

This research evaluated the performance of YOLO models in object detection and counting tasks by comparing key metrics, including precision, recall, F1-score, mAP@50, and training time. According to Table 1, training time increases from YOLOv5 to YOLOv8, with YOLOv8 requiring 48 min and 58 sec. YOLOv5 has the shortest training time, while YOLOv7 takes the longest. These results align with expectations, as YOLOv7 is as resource-intensive as YOLOv8, both models showing similar training durations. The study concluded that YOLOv5 is the most efficient model for detecting mango trees in satellite imagery and for training on large datasets for tree counting. These findings highlight the computational cost of every model variant. When selecting a model, training time is an essential factor, but it should be weighed against other performance indicators like inference speed.

Table 1 Training time of YOLO's variant

Variants of YOLO	Training Time (min)
YOLOV5	37.54
YOLOv6	45.55
YOLOv7	49.08
YOLOv8	48.58

Each YOLO model's performance metrics on the validation dataset are shown in Table 2. The ability of YOLOv8 to lower false positive detections is demonstrated by its greatest precision of 93.1% when precision values are compared. Fig. 2(b), which visualizes the precision values for each model at 50 epochs, shows YOLOv8 leading with the highest precision, followed by YOLOv7 at 86.5%. Both YOLOv5 and YOLOv6 have a precision value of 85.7%.

Turning to recall, Fig. 2 (a) demonstrates that YOLOv8 outperforms the others with a recall value of 94.2%, underscoring its ability to identify real-world mango tree occurrences. With a recall of 80.3%, Table 2 shows that YOLOv5 comes in second, demonstrating its capacity to identify occurrences. Additionally, YOLOv6 and YOLOv7 perform well, with recall scores of 79.3% and 79.5%, respectively.

With the best score of 93.6%, YOLOv8 once again distinguishes out when examining the F1-score (Fig. 2c), demonstrating a solid balance between recall and accuracy. Notably, the F1-scores of the other models are almost the same, with YOLOv5, YOLOv6, and YOLOv7 obtaining scores of 82.9%, 82.3%, and 82.8%, respectively, indicating their competitive ability to strike a balance between recall and precision.

YOLOv8 achieves the greatest score of 97% in terms of mAP@50, followed by YOLOv6 with a score of 86% (Table 2). The equivalent scores for YOLOv5 and YOLOv7 are 85.9% and 85%. The trade-off in mAP@50 scores as the number of epochs grows is seen in Fig. 2(d).

Table 2 Comparative performance of different YOLO models in terms of precision, recall, F1-score, and mAP@50

YOLO variant	Precision (%) (95% CI)	Recall (%) (95% CI)	F1-score	mAP @50 (95% CI)
YOLOv5	85.7	80.3	82.9	85.9
YOLOv6	85.6	79.3	82.3	86.0
YOLOv7	86.5	79.5	82.8	85.0
YOLOv8	93.1	94.2	93.6	97.0

CI, Confidence intervals.

YOLOv8 achieved the highest performance across all key metrics, demonstrating significantly superior precision (93.1%), recall (94.2%), F1-score (93.6%), and mAP@50 (97.0%) (Table 2). The confidence intervals (CIs) for YOLOv8 do not overlap with those of YOLOv5, YOLOv6, and YOLOv7, confirming that its improvements in detection accuracy are statistically significant. Conversely, the overlapping CIs among YOLOv5, YOLOv6, and YOLOv7 indicated that their performance differences were minor and not statistically significant. While YOLOv8 is the most accurate model, practical considerations such as computational efficiency and processing time may influence model selection, making YOLOv5, YOLOv6, or YOLOv7 viable alternatives for applications where a balance between accuracy and efficiency is required.

These findings suggested that YOLOv8 has outstanding accuracy, recall, F1-score, and mAP@50 performance, which makes it a viable choice for tasks involving the recognition and counting of mango trees. However, other models, including YOLOv5, YOLOv6, and YOLOv7, also showed competitive performance, particularly in terms of training time.

Fig. 2 displays the performance of every measurement for variants of YOLO on the validation dataset. It is evident that YOLOv8 experiences slower convergence across all metrics when compared to the other models. YOLOv8 showed stability for all measures beyond the 50th epoch, suggesting that model's performance has reached its peak. In contrast, YOLOv5, YOLOv6, and YOLOv7 exhibit faster convergence during the initial epochs, especially with regard to precision, F1-score, and mAP@50. Additionally, after 50th epoch, YOLOv8 model exhibits reduced variability and improved stability across all measures. Furthermore, YOLOv8 maintains consistent performance across all epochs without significant degradation in results.

YOLOv8 is a promising choice for mango tree recognition and tree counting in mango orchards since it finds an equilibrium between model complexity and performance measures. A key observation is the balance between accuracy and computational efficiency. While YOLOv8 exhibits the highest detection accuracy, its training time was relatively longer than other models. YOLOv7 also demands considerable computational resources but does not

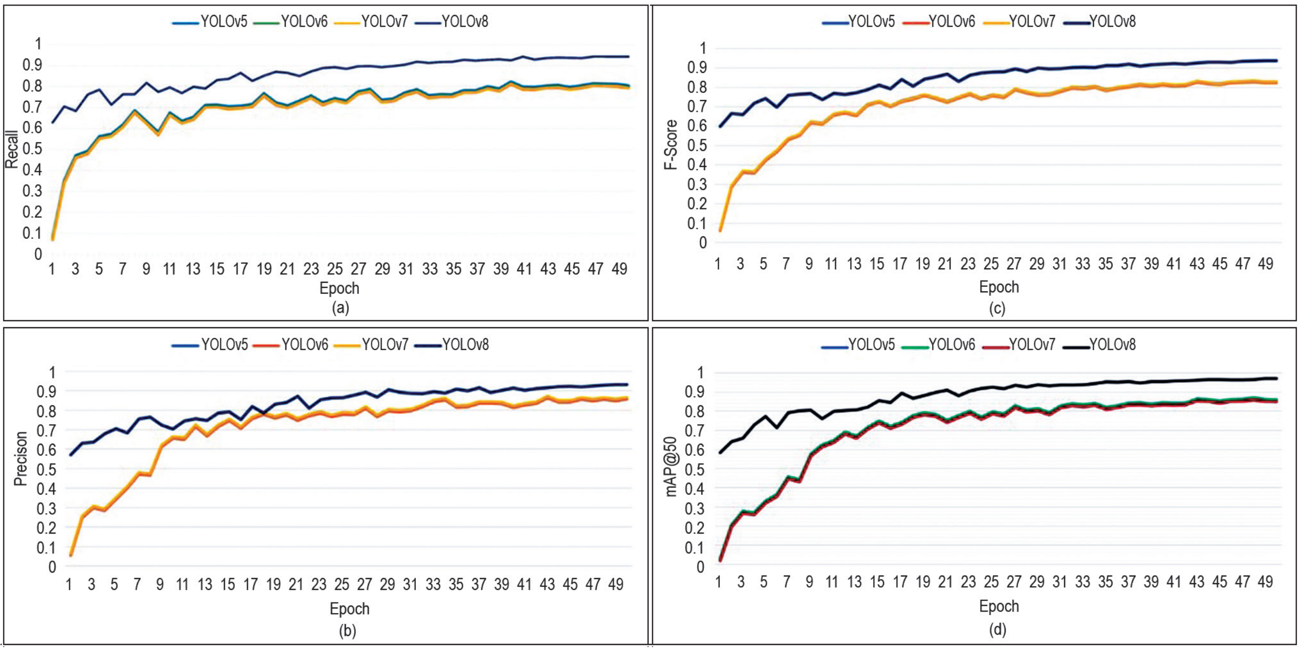

Fig. 2 Performance of all metrics for validated YOLO models.

match the accuracy of YOLOv8. In contrast, YOLOv5 and YOLOv6 offered a more computationally efficient alternative while maintaining competitive accuracy levels, making them viable choices for real-time applications where processing power is a limiting factor. These findings emphasize that model selection should depend on the specific use case, whether prioritizing accuracy (YOLOv8) or computational efficiency (YOLOv5/YOLOv6).

Moving on to the loss analysis, Fig. 3 presents average loss values for YOLO models. Beginning with box loss, both YOLOv8 and YOLOv5 exhibit consistent trend, with YOLOv8 having the lowest training and validation box losses compared to all previous YOLO versions. This suggested that the YOLOv8 model learned to predict bounding boxes more well, which improved generalization of bounding box predictions. As the box loss, the classification (cls) loss in Fig. 3 has a somewhat predictable pattern. The declining trend in training cls loss suggested that YOLOv5 and YOLOv7 performed better during training. Notably, YOLOv6 displayed a higher validation cls loss than the other models, yet it doesn't outperform them overall. This suggested that YOLOv6's training in cls prediction does not necessarily result in better generalization. When considering the training and validation cls losses, YOLOv5 and YOLOv7 have nearly identical loss values, while YOLOv8 showed higher but still competitive losses. Regarding obj loss, both YOLOv5 and YOLOv7 exhibited a small difference between training and validation losses in Fig. 3(a) and 3(c), suggested that they are not overfitting the training dataset. Despite significant overfitting during training, YOLOv5 had a smaller validation obj loss than YOLOv7, indicating that it may generalize better in identifying objects in the validation data.

The deployment of YOLO-based models in agricultural applications extends beyond mango tree detection, playing a significant role in precision agriculture, automated orchard

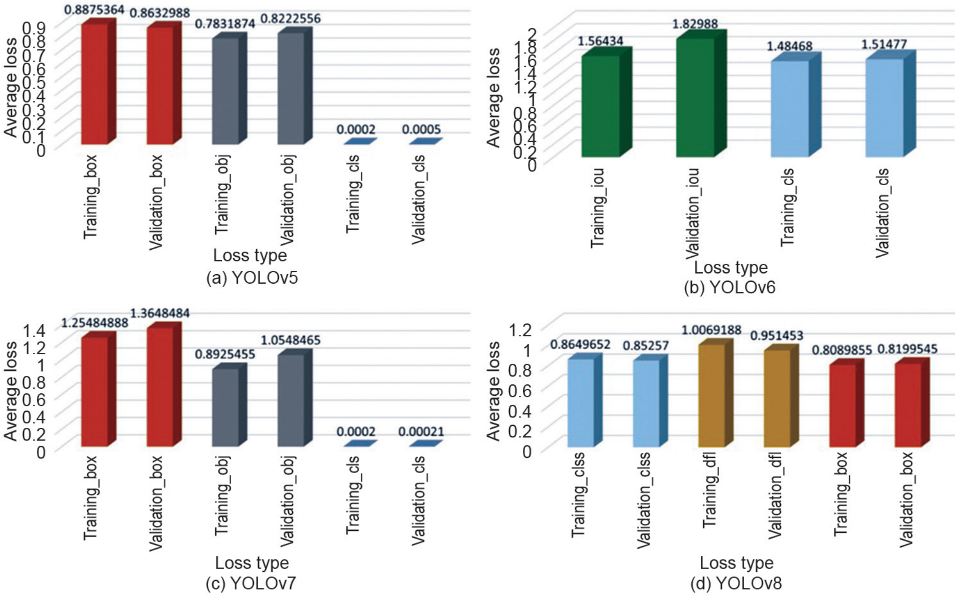

Fig. 3 The average loss values for all the YOLO's variants.

monitoring, and yield estimation. Accurate tree detection facilitates resource optimization, irrigation planning, and harvesting strategies, reducing the dependence on manual surveying techniques. However, practical deployment of YOLO based mango tree detection is limited by hardware constraints, high computational demands, and model generalizability. While YOLOv8 delivers superior accuracy, its resource intensive nature makes it unsuitable for low-power devices, whereas YOLOv5, though less accurate, offers better efficiency for real-time applications, with potential improvements through model optimization techniques. To address these challenges, future research should focus on model optimization techniques, such as quantization, pruning, and edge computing acceleration using TensorRT, EdgeTPU, or cloud-based processing. Additionally, hybrid models combining YOLO with traditional remote sensing techniques could further enhance precision and detection robustness, particularly in dense orchard settings.

In order to recognize mango trees and count the number of trees in mango orchards, we thoroughly evaluated the performance of several YOLO architectures in this work, including YOLOv5, YOLOv6, YOLOv7, and YOLOv8. The findings showed that, in both testing and validation, YOLOv8 had the best overall balance across all metrics. The YOLOv8 model demonstrated outstanding performance in accurately detecting trees, with an overall Precision of 93.14%, Recall of 94.25%, F1-score of 93.69%, and a mean average precision (mAP) of 97.03%. These results highlight the efficiency of the YOLOv8 architecture in mango tree detection and counting tasks. It is important to note, nevertheless, that throughout testing and validation, the YOLOv6 variation performed much worse across all metrics. In conclusion, the particular requirements of the application need to direct the model selection, taking into factors like accuracy, recall, inference time, and the trade-offs between these characteristics.

REFERENCES

Casas E, Ramos L, Bendek E and Rivas-Echeverría F. 2023. Assessing the effectiveness of YOLO architectures for smoke and wildfire detection. *IEEE Access*.

Gupta C, Gill N S, Gulia P and Chatterjee J M. 2023. A novel finetuned YOLOv6 transfer learning model for real-time object detection. *Journal of Real-Time Image Processing* **20**(3): 42.

Itakura K and Hosoi F. 2020. Automatic tree detection from three-dimensional images reconstructed from 360 spherical camera using YOLO v2. *Remote Sensing* **12**(6): 988.

Liu J, Zhang J, Ni Y, Chi W and Qi Z. 2024. Small-object detection in remote sensing images with super resolution perception. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Mathew M P and Mahesh T Y. 2022. Leaf-based disease detection in bell pepper plant using YOLO v5. *Signal, Image and Video Processing*, pp. 1–7.

Mekhalfi M L, Nicolo, C Bazi Y, Al Rahhal M M, Alsharif N A and Al Maghayreh E. 2021. Contrasting YOLOv5, transformer, and EfficientDet detectors for crop circle detection in desert. *IEEE Geoscience and Remote Sensing Letters* **19**: 1–5.

National Horticulture Board, Indian Horticulture Database. Ministry of Agriculture, Government of India. 2023. www.agricoop.nic.in

Padilla R, Netto S L and Da Silva E A. 2020. A survey on performance metrics for object-detection algorithms. *(In) 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 237–42.

Padilla R, Passos W L, Dias T L, Netto S L and Da Silva E A. 2021. A comparative analysis of object detection metrics with a companion open-source toolkit. *Electronics* **10**(3): 279.

Patel K. Bhatt C and Mazzeo P L. 2022. Improved ship detection algorithm from satellite images using YOLOv7 and graph neural network. *Algorithms* **15**(12): 473.

Putra Y C and Wijayanto A W. 2023. Automatic detection and counting of oil palm trees using remote sensing and object-based deep learning. *Remote Sensing Applications: Society and Environment* **29**: 100914.

Redmon J, Divvala S, Girshick R. and Farhadi A. 2015. You only look once: Unified, real-time object detection (2015). *arXiv preprint arXiv:1506.02640*, *825*.

Sohan M, Sai Ram T, Reddy R and Venkata C. 2024. A review on yolov8 and its advancements. *(In) International Conference on Data Intelligence and Cognitive Informatics,* Springer, Singapore, pp. 529–45.

Srividhya S, Thilagam P, Venilla M A, Geetha K and Rani M A 2024. Efficacy of foliar versus soil application of micronutrients on the production of mango (*Mangifera indica* L.). *Plant Science Today* **11**(3).

Tamang S, Sen B, Pradhan A, Sharma K and Singh V K. 2023. Enhancing covid-19 safety: Exploring yolov8 object detection for accurate face mask classification. *International Journal of Intelligent Systems and Applications in Engineering* **11**(2): 892–97.

Wang C Y, Bochkovskiy A and Liao H Y M. 2023. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *(In) Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7464–75.

Zhao L and Li S. 2020. Object detection algorithm based on improved YOLOv3. *Electronics* **9**(3): 537.

Zhao Z Q, Zheng P, Xu, S T and Wu X. 2019. Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems* **30**(11): 3212–32.

Zhu H, Wei H, Li B, Yuan X and Kehtarnavaz N. 2020. A review of video object detection: Datasets, metrics and methods. *Applied Sciences* **10**(21): 7834.