Kırklareli Üniversitesi
Mühendislik ve Fen Bilimleri Dergisi
Cilt 11, Sayı 1, 11-22, 2025

Kırklareli University
Journal of Engineering and Science
Volume 11, Issue 1, 11-22, 2025

Research Article DOI: : 10.34186/klujes.1560553

Yield Prediction with Deep Learning on UAV Images: Banana tree application

Furkan SÖNMEZ[1], Polat ASHYROV[1], Hayrettin TOYLAN[1*]

[1]*Department of Mechatronics Engineering, Faculty of Technology, Kırklareli University, Kırklareli, Türkiye*

Received: 03.10.2025, Accepted: 02.01.2025, Published: 04.03.2025

ABSTRACT

Agriculture is developing with the integration of smart imaging technologies into the production, harvesting, and classification of agricultural products. This paves the way for obtaining qualified and quantitative products. The use of imaging technologies and deep learning methods in the agricultural field can increase the success of yield prediction, considering climate change and environmental conditions. This study proposes yield prediction for banana trees based on the YOLO method, using images obtained from unmanned aerial vehicles. Firstly, the performance of YOLOv8 and YOLOv9 models trained using the RoboFlow dataset is analysed. According to the comparison results, it was observed that the YOLOv9 model obtained more successful results with 87.6% mAP, 94% precision, 96% recall, and 94.9% F1-score. Using the YOLOv9 model, the banana yield in the trees was estimated correctly by an average of 78% in the experimental studies conducted on the images obtained by the UAV. This method provides a reliable detection approach for accurately estimating the banana tree yield but needs to be improved.

Keywords: Deep Learning; UAV; YOLO; Yield Prediction

İHA Görüntüleri Üzerinden Derin Ögrenme ile Rekolte Tahmini: Muz ağacı uygulaması

ÖZ

Akıllı görüntüleme teknolojilerin tarım ürünlerinin üretimine, hasadına ve sınıflandırmasına entegre edilmesi ile tarım gelişmektedir. Bunun etkisi ile nitelikli ve nicelikli ürünler elde edilmesinin önü açılmaktadır. Tarım alanında görüntüleme teknolojileri ve derin öğrenme yöntemleri kullanımı ile, iklim değişikliği ve çevresel koşullara bağlı değişen rekolte tahmin başarısı da artırılabilir. Bu çalışma, insansız hava araçlarından elde edilen görüntüler yardımı ile YOLO metodunu temel alarak muz ağaçlarında rekolte tahminini önermektedir. İlk olarak, RoboFlow veri kümesi kullanılarak eğitilen YOLOv8 ve YOLOv9 modellerinin performansı analiz edildi. Karşılaştırma sonuçlarına göre YOLOv9 modeli muz görüntülerini %87.6 mAP, %94 Precision, %96 recall ve %94.9 F1-score ile daha başarılı sonuçlar elde ettiği görülmüştür. YOLOv9 modeli ile, İHA tarafından elde edilen görüntüler üzerinden yapılan deneysel çalışmalarda ağaçlardaki muz rekoltesi ortalama %78 oranında doğru tahmin etmiştir. Bu yöntem, muz ağacı verimini doğru bir şekilde tahmin etmek için güvenilir fakat geliştirilmesi gereken bir tespit yaklaşımı sunar.

Anahtar Kelimeler: Derin Öğrenme; İHA; YOLO; Rekolte Tahmini

**Corresponding Author: Hayrettin TOYLAN*
hayrettintoylan@klu.edu.tr

1. INTRODUCTION

Ensuring food sustainability is of paramount importance for the survival of people around the world. Crop yield forecasts provide important support in ensuring food continuity (Sneha et al., 2024). Nowadays, most farmers manually check the growth of their crops. This method makes it very difficult to determine the status of crops in large agricultural lands. However, image processing techniques offer a solution by allowing for the determination of size characteristics, quality characteristics, disease, and pest detection of agricultural products. This information is crucial for production planning, determining base prices, and minimizing potential issues in this field (Balambar et al., 2021). Unmanned aerial vehicles (UAVs) equipped with cameras can capture high-resolution images that can be processed to obtain valuable information about crops in agricultural lands from various angles.

In addition to traditional methods, image data analysis methods are frequently used in yield estimation problems. The integration of these images with artificial intelligence-based deep learning algorithms significantly enhances the success of the method. In their study, Yıldırım and Ulu employed Single-Shot Multi-Box Detection (SSD) Mobilnet and Faster R-CNN deep learning model architectures to detect apples. They conducted yield prediction using the obtained apple images and reported a 10% increase in accuracy (Yıldırım & Ulu, 2023). In recent years, YOLO, one of the deep learning algorithms, has been widely used by researchers in yield prediction. Koirala et al. compared the performance of 6 deep learning methods (Faster R-CNN (VGG) and Faster R-CNN (ZF), and the single-stage techniques YOLOv3, YOLOv2, YOLOv2 (tiny), and SSD) for the mango fruit detection task from tree images. The new model designed based on YOLOv3 and YOLOv2 (tiny) was found to be more successful than the other six models (Koirala et al., 2019). Bai et al. used an improved YOLO 7 model to accurately identify the flowers and fruits of strawberry seedlings in a greenhouse. By improving the spatial interactions and feature extraction capability of the model, they obtained precision (P), recall (R) and mean accuracy (mAP) values of 92.6%, 89.6% and 92.1% respectively (Bai et al., 2024). Paul et al. compared the performance of various YOLO models to improve capsicum harvesting. From this evaluation, the YOLOv8s model was found to be the most effective for capsicum detection. Throughout the harvest cycle, this model achieved a remarkable counting accuracy of 94.1% (Paul et al., 2024).

In a study conducted by Sneha et al., the potential of the YOLO deep learning network was explored for detecting grape clusters and predicting grape harvests. Three models, YOLOv3, YOLOv4, and YOLOv5, were trained on grape clusters and compared. The recall and F1 scores for YOLOv3, YOLOv4, and YOLOv5 were found to be 90.2, 91.25, 95.63 and 92.21, 92.68, 93.21, respectively. The results indicated that YOLOv5 was the most successful model (Sneha et al., 2024). Tian et al. integrated the DenseNet method into the YOLOv3 model for real-time detection of apples in orchards, evaluation of apple growth stages, and prediction of yield. They emphasized the real-time applicability of this model for analysing images (Tian et al., 2019). Sun et al. used an improved detection method based on Yolov5 to determine rice panicle density, which is an important feature in determining rice yield. Compared to the original Yolov5, the model increased the detection rate by 12.63% and improved the accuracy by 3.82% (Sun et al., 2024).

Deep learning methods such as YOLO which have powerful feature extraction capabilities, are of great interest in predicting agricultural yields (Chakraborty et al., 2022). The close proximity of fruits on a banana tree and their similar colors during ripening pose a challenge in individual recognition. This study aims to tackle the complex issue of estimating yield in banana trees, which currently relies heavily on human vision and is a labour-intensive process. The main goal is to develop deep learning-based object detection models for identifying fruits in aerial images of banana trees captured by UAVs, and to experimentally compare their performance.

2. MATERIAL AND METHODS

The data used in the experimental studies were obtained from a single-crop banana orchard located in the Manavgat district of Antalya province in Turkey. The orchard consists of a total of 3000 banana trees and yields approximately 100,000 tons of bananas per year. The general methodology of this study is shown in **Figure 1**. Initially, the UAV captured images of the banana trees in RGB format, with a resolution of 3840x5120 pixels. The banana images were then resized to 640x640 for training and testing the deep learning algorithms YOLOv8 and YOLOv9 models.

Figure 1: The general methodology of the system

The application for estimating banana fruit yield using image processing method was implemented using the Microsoft Visual Studio Code program and the OpenCv and Ultralytics library. The Ultralytics libraries is an open-source artificial intelligence library commonly used in computer vision projects, particularly for object detection models. One of Ultralytics' most well-known projects is the enhanced version of the YOLO (You Only Look Once) algorithm.

2.1 YOLO: Deep learning model for object detection

In recent years, the YOLO (You Only Look Once) algorithm has gained popularity for its real-time object detection capabilities. It differs from traditional two-stage approaches by using a single-stage object detection method. With each new version, YOLO has improved its performance. However, the latest version, YOLOv5, introduced a different structure compared to the previous four versions. Instead of the Darknet framework, YOLOv5 adopted the PyTorch framework (Bakirci & Bayraktar, 2024). Following YOLOR, YOLOX, YOLOv6, YOLOv7, and DAMA-YOLO versions, Ultralytics, the entity developing YOLOv5, released YOLOv8 (Jocher et al., 2023) in January 2023. YOLOv8 maintains a similar backbone to YOLOv5 but introduces modifications to the CSPLayer, also known as the C2f module. The YOLOv8 architecture is given in **Figure 2**.

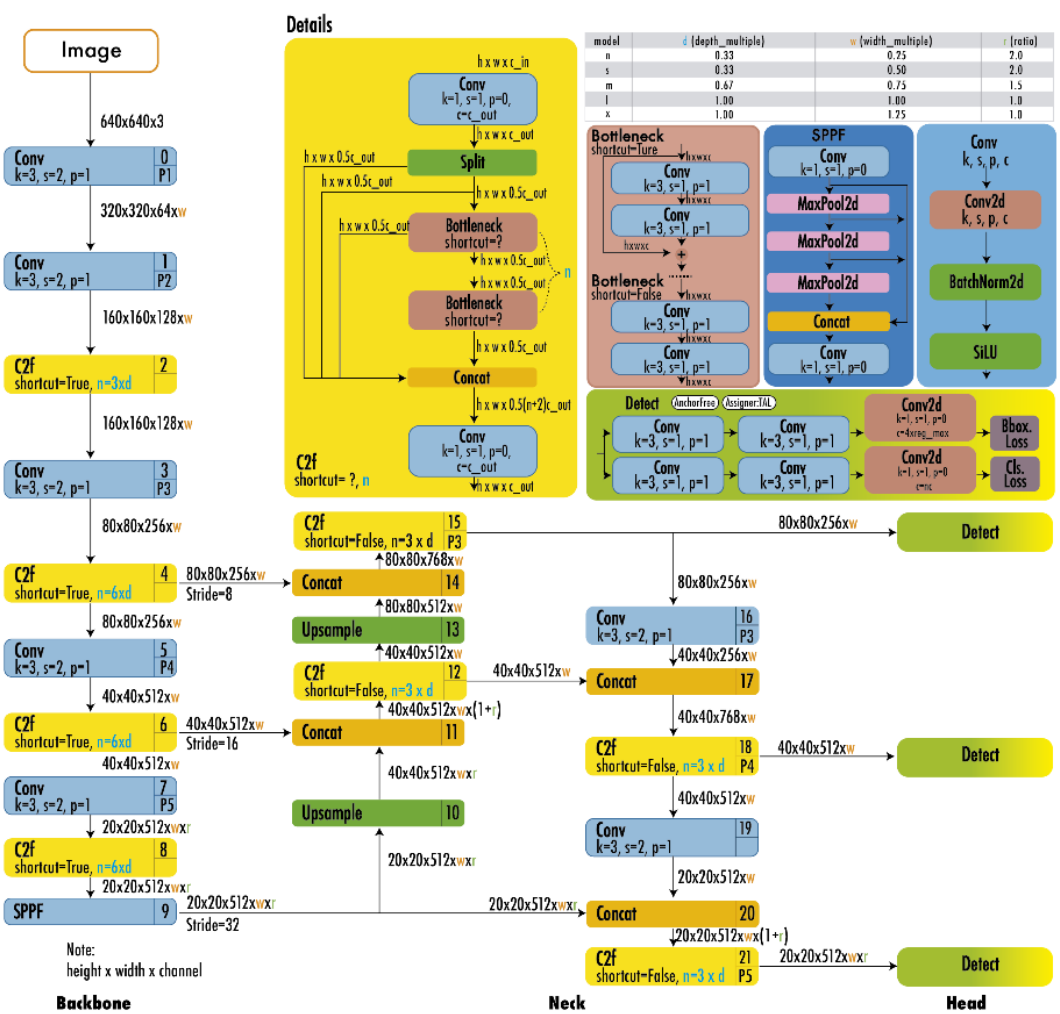


Figure 2: YOLOv8 architecture (Terven et al., 2023)

The YOLOv8 model has demonstrated enhanced performance in a multitude of scenarios, such as complex backgrounds, low-light conditions, and the detection of small objects. However, it's important to consider that real-time videos may experience a decrease in speed. In the YOLOv9 model, introduced in February 2024, two new features were implemented to enhance performance and localization accuracy. These features are programmable gradient information (PGI) and the generalized efficient layer aggregation network (GELAN). Their purpose is to prevent data loss during feed-forward operations. Furthermore, improvements have been made to the computational modules used in the artificial neural network structures, resulting in enhanced flexibility of the network (Yang et al., 2024). The YOLOv9 architecture is given in **Figure 3**.

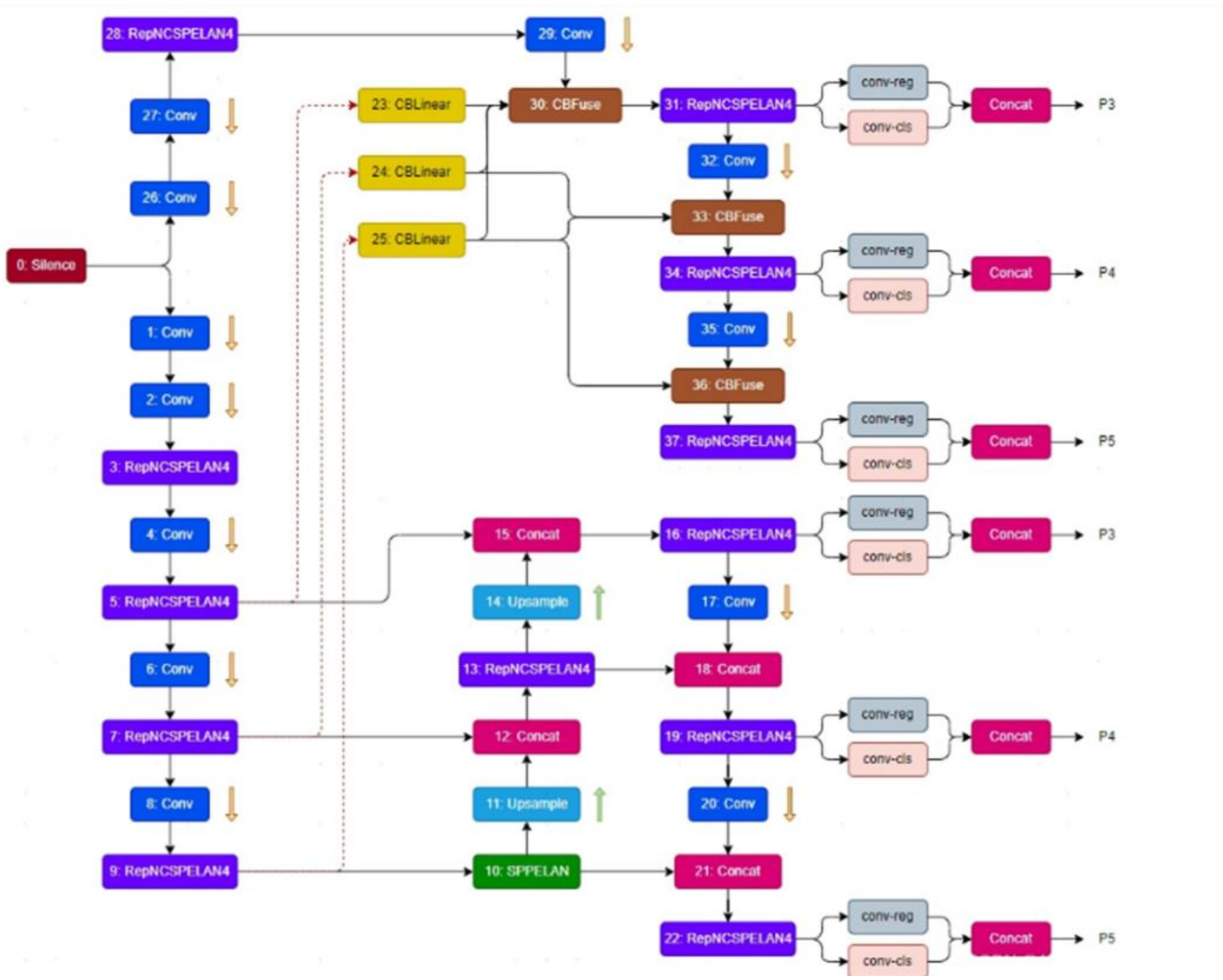


Figure 3: YOLOv9 architecture (Wang, 2024)

It seems reasonable to posit that YOLOv9 will prove particularly adept at meeting the demands of agricultural applications, given its performance. YOLO methods enhance accuracy and efficiency in computer vision applications, establishing a solid foundation for future advancements.

2.2 Metrics

Performance metrics are utilized to assess the accuracy and effectiveness of object recognition models. These metrics provide numerical insights into the model's ability to correctly identify and determine the position of objects. Furthermore, they evaluate the model's performance by taking into account the occurrences of false positives and false negatives. For this study, four metrics were employed to measure the performance of both YOLOv8 and YOLOv9 models.

Mean Average Precision (mAP) is a performance statistic used to predict target locations and categories. Precision (P) measures the proportion of true positives, while recall calculates the proportion of true positives among all the positives. The F1 Score, on the other hand, is the harmonic mean of precision and recall (Subeesh et al., 2024).

$$\mathbf{P} = \frac{Tp}{Tp+Fp} \tag{1}$$

$$\mathbf{R} = \frac{Tp}{Tp+Fn} \tag{2}$$

$$\textbf{F1 score} = \frac{2xPxR}{P+R} \tag{3}$$

$$\boldsymbol{mAP} = \frac{1}{K}\sum_{i=1}^{k} \mathrm{AP}_i \tag{4}$$

Tp represents True Positive, F_N represents False Negative, F_P represents False Positive, K represents the total number of classes, and AP represents average precision (Subeesh et al., 2024).

3. RESULTS and DISCUSSION

3.1. Dataset

In this study, we used Roboflow software to create the dataset. This software effectively organizes object images using artificial intelligence-supported bounding boxes. A total of 70 banana images, collected from a banana greenhouse and captured from different angles, were divided into four classes according to their weights for yield prediction: 1 kg, 3 kg, 5 kg, and 7 kg. The Roboflow software was employed to label the banana images according to the aforementioned four classes. **Figure 4** depicts sample banana images that were prepared for labelling. The training process of the YOLOv8 and YOLOv9 models was conducted by transferring the dataset that was generated following the labelling process to Google Colab. The training process of the two models was completed after 5000 iterations.

Figure 4: Sample banana images

3.2 Performance of YOLO models

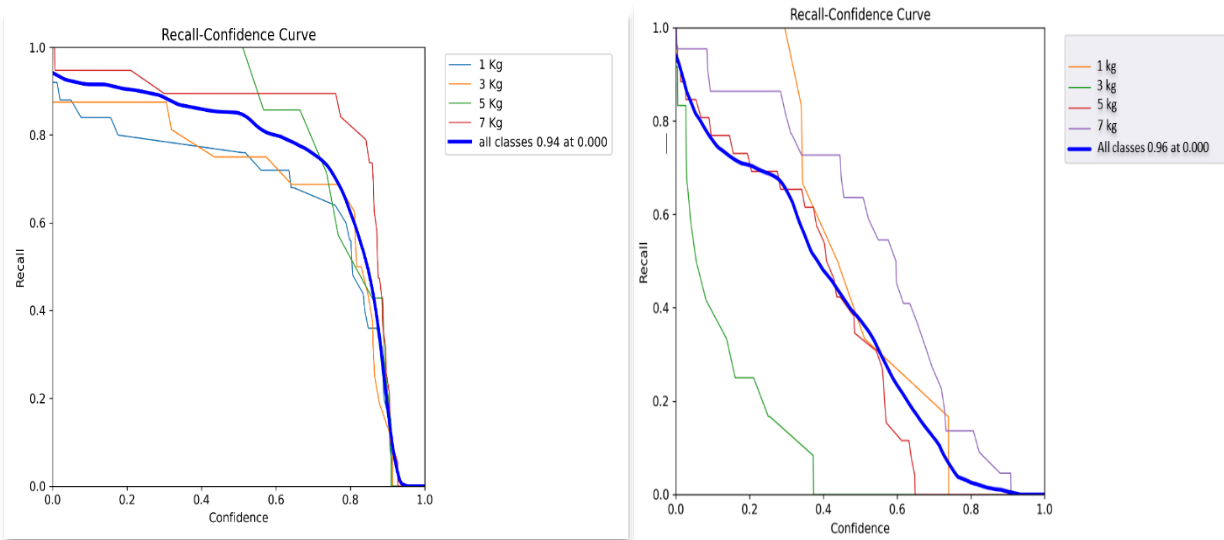
In this study, YOLOv8 and YOLOv9 deep learning models were trained and tested with the dataset described above. Accuracy, F1-score, Precision, and Recall were calculated using Equation (1), Equation (2), Equation (3), and Equation (4) respectively. The comparison results of the two models are shown in **Table 1**.

Table 1: The comparison results of the YOLOv8 and YOLOv9

Models	**Accuracy (%)**	**Precision**	**Recall**	**F1-score**
YOLOv8	84.9	88	94	90.9
YOLOv9	87.6	94	96	94.9

When the results are analyzed in terms of mAP value, it is seen that banana labels are detected with a higher mAP value with YOLOv9. The confidence and recall graphs of YOLOv8 and YOLOv9 models trained with images with four different kilogram labels are given in **Figure 5**.

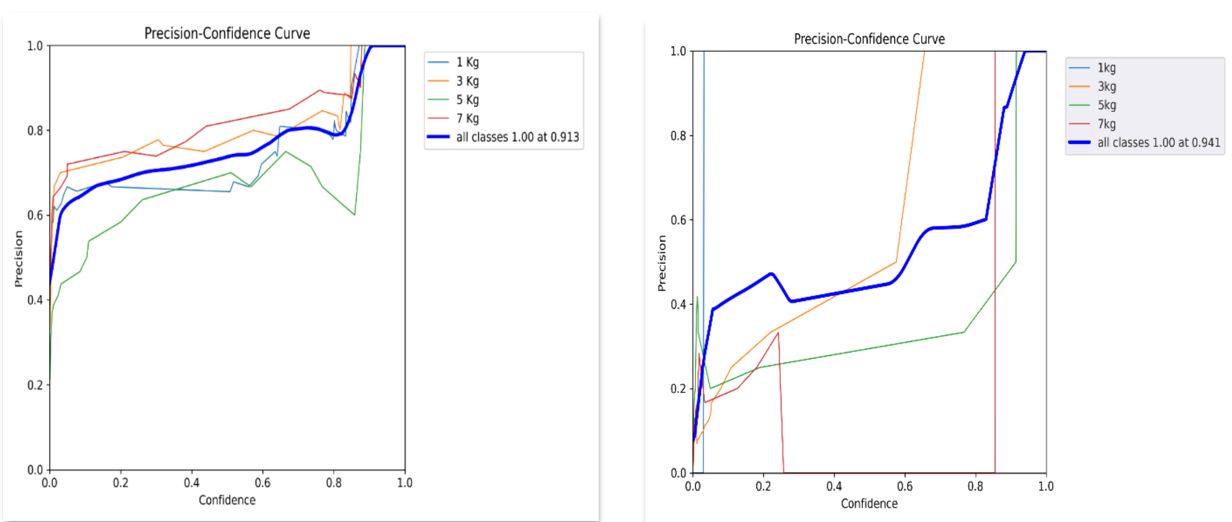
The X-axis of Figure 5 represents the confidence threshold, which determines the minimum confidence score needed for an estimate to be considered valid. On the other hand, the Y-axis represents recall, which is also known as precision or the true positive rate. Recall measures the ratio of a model's true positive predictions to the total true positives and gives an indication of how accurately the model predicts true positive instances. This metric is especially important in classification problems where the cost of false negatives is high. A high recall value suggests that the model is more likely to generate false positives than to miss true positives. As the confidence threshold decreases, more predictions are considered valid, including both true positives and potential false positives.



YOLOv8 YOLOv9

Figure 5: Recall-Confidence Curves of the YOLOv8 and YOLOv9

The confidence-precision curve given in **Figure 6** is a graphical representation that depicts the relationship between the confidence and precision threshold of a deep learning model. As the level of confidence increases, the precision of the model will also increase in proportion.

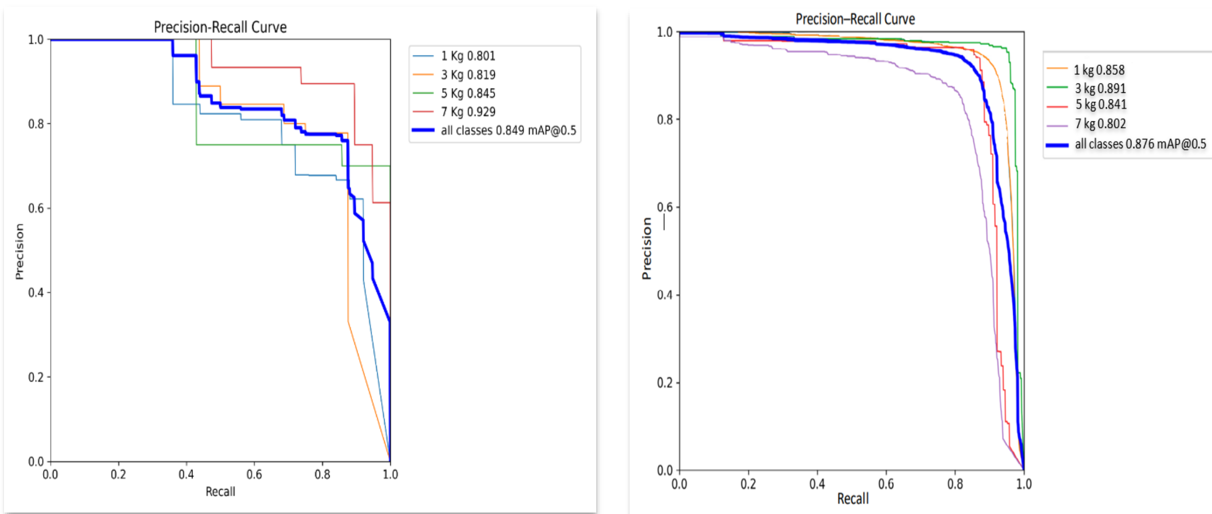


YOLOv8 YOLOv9

Figure 6: Confidence-Precision Curves of the YOLOv8 and YOLOv9

The X-axis represents the confidence threshold, while the Y-axis represents the sensitivity. The curve illustrates how the precision changes as the confidence threshold is altered.

The recall -precision curve is a graphical representation that demonstrates the performance and efficiency of the YOLOv8 and YOLOv9 models in identifying the labels of banana images. This is demonstrated in **Figure 7.**



Yolov8 Yolov9

Figure 7: Recall-Precision Curves of the YOLOv8 and YOLOv9

The X-axis represents the recall threshold, while the Y-axis represents the precision. The curve illustrates how the precision changes with varying recall thresholds. The recall-precision curve is the graph that shows the performance and efficiency of the YOLOv8 and YOLOv9 models for labelling banana images.

YOLOv9 has a deeper and wider architecture compared to YOLOv8, which allows for learning more complex features. Both models used a learning rate of 0.001 in combination with the ADAM optimization algorithm. In this study, it is possible to improve accuracy by using more advanced data augmentation techniques, new loss functions, and more efficient optimization algorithms. However, optimizing the different optimization algorithms and learning rates is not the aim of this study. This study aims to use deep learning models for yield prediction and demonstrate their success numerically. The sample images obtained from the banana tree in **Figure 8** were analysed with the YOLOv9 model. The yield predictions of the model, the actual yield value, and the related success rate are given in **Table 2**.

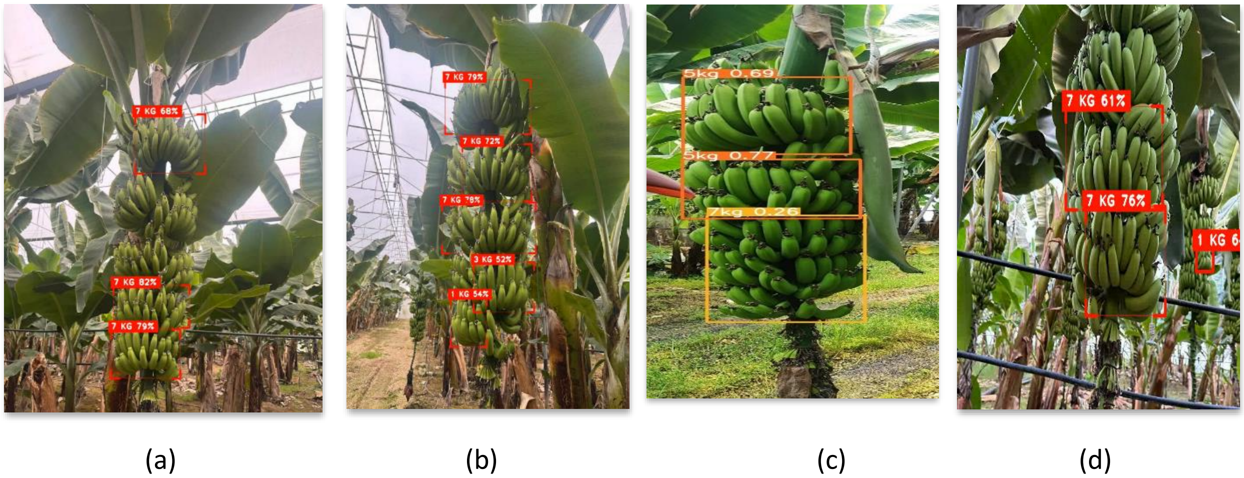


Figure 8: Images of banana tree

Table 2: YOLOv9 yield prediction values

Images	Actual Yield Value (kg)	Estimated Yield Value (kg)	Accuracy Rate (%)
Image (a)	28	21	75
Image (b)	32.5	25	77
Image (c)	19	17	89
Image (d)	21	15	71

The study encountered several unfavourable conditions that affected the accuracy of the predictions. The camera quality, angle of sunlight on the banana fruits, and tree leaves covering the fruits directly impacted the performance of the model. To enhance the accuracy of the predictions, the training dataset can be expanded and the parameters of the deep learning algorithm can be optimized.

4. CONCLUSION

This experimental study demonstrates the performance of YOLO, a widely used deep learning model, on the yield prediction of banana trees. Based on results from experiments with two different versions of the model, YOLOv8 and YOLOv9, we concluded that the YOLOv9 approach, which performs high-precision processing, is more suitable for this application, where accuracy in yield predictions is crucial. The YOLOv9 model, trained on a database created using images obtained from an unmanned aerial vehicle, achieved an accuracy rate of 87.6%. In tests conducted on images captured from four distinct banana trees, the model exhibited an average of 76% accuracy in predicting banana yield. Additionally, it may be possible to enhance its performance further by optimizing the basic parameters of the YOLOv9 architecture.

As a result, it is an important fact that yield prediction using deep learning on images obtained from UAVs will play a pivotal role in the planning of food production for human consumption. This method allows for the investigation of plant protection and pest control. In this regard, it is inevitable that the rapid development of technology, the introduction of image processing methods due to the progress of unmanned aerial vehicles and camera systems, and the introduction of unmanned aerial vehicles into our lives in various branches, including agriculture, will occur. It can be said that technological developments will continue in the coming processes and that artificial intelligence will enter our lives more with UAVs, so that the use of unmanned aerial vehicles in agriculture will gain more speed.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

ACKNOWLEDGEMENT

This study supported by TUBITAK (The Scientific and Technological Research Council of Turkey) with 2209-A - Research Project Support Programme

CONTRIBUTIONS OF AUTHORS

F.T.: Methodology, investigation, research, application, writing-original draft preparation,

P.A.: Methodology, investigation, research, application, writing-original draft preparation,

H.T.: Methodology, supervision, validation, writing-original draft preparation, writing-review and editing

REFERENCES

Bakirci, M., & Bayraktar, I. (2024, April). Boosting aircraft monitoring and security through ground surveillance optimization with YOLOv9. In *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, 1-6

Bai, Y., Yu, J., Yang, S., & Ning, J. (2024). An improved YOLO algorithm for detecting flowers and fruits on strawberry seedlings. Biosystems Engineering, 237, 1-12.

Balambar, Ş., Karimi, Z. K., Öztürk, F., Acet, Ş. B., & Pekkan, Ö. I. (2021). Uzaktan algılama tekniklerinden yararlanarak tarımsal faliyetlerin izlenmesi. *GSI Journals Serie C: Advancements in Information Sciences and Technologies*, *4*(2), 58-79.

Chakraborty, S. K., Chandel, N. S., Jat, D., Tiwari, M. K., Rajwade, Y. A., & Subeesh, A. (2022). Deep learning approaches and interventions for futuristic engineering in agriculture. *Neural Computing and Applications*, *34*(23), 20539-20573.

Jocher, G.; Chaurasia, A.; Qiu, J. YOLO by Ultralytics. (2023). Available online: **https://github.com/ultralytics/ultralytics** (accessed on 04 September 2024).

Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. (2019). Deep learning for real-time fruit detection and orchard fruit load prediction: Benchmarking of 'MangoYOLO'. Precis. Agric. 20, 1107–1135.

Paul, A., Machavaram, R., Kumar, D., & Nagar, H. (2024). Smart solutions for capsicum Harvesting: Unleashing the power of YOLO for Detection, Segmentation, growth stage Classification, Counting, and real-time mobile identification. *Computers and Electronics in Agriculture*, *219*, 108832.

Sneha, N., Sundaram, M., & Ranjan, R. (2024). Acre-Scale Grape Bunch Detection and Predict Grape Harvest Using YOLO Deep Learning Network. *SN Computer Science*, *5*(2), 250.

Subeesh, A., Kumar, S. P., Chakraborty, S. K., Upendar, K., Chandel, N. S., Jat, D., Dubey, K., Modi, R.U., Khan, M. M. (2024). UAV imagery coupled deep learning approach for the development of an adaptive in-house web-based application for yield estimation in citrus orchard. *Measurement*, *234*, 114786.

Sun, J., Zhou, J., He, Y., Jia, H., & Rottok, L. T. (2024). Detection of rice panicle density for unmanned harvesters via RP-YOLO. *Computers and Electronics in Agriculture*, *226*, 109371.

Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-nas. *Machine Learning and Knowledge Extraction*, *5*(4), 1680-1716.

Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., & Liang, Z. (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and electronics in agriculture*, *157*, 417-426.

Wang, Z. (2024, May). Enhanced Fire Detection Algorithm for Chemical Plants Using Modified YOLOv9 Architecture. In *2024 9th International Conference on Electronic Technology and Information Science (ICETIS)*, 22-25.

Yang, S., Cao, Z., Liu, N., Sun, Y., & Wang, Z. (2024). Maritime Electro-Optical Image Object Matching Based on Improved YOLOv9. *Electronics*, *13*(14), 2774.

Yıldırım, Ş., & Ulu, B. (2023). Deep learning-based apples counting for yield forecast using proposed flying robotic system. *Sensors*, *23*(13), 6171.