Evaluating YOLOv7-tiny for Olive Tree Pruning Detection in Drone Images

Katya Trufanova

***Abstract*—This project aims to detect and classify pruned and unpruned trees in high-resolution drone images of olive tree fields using the YOLOv7-tiny model for object detection. Experimental results demonstrate the impact of hyperparameters on the model's performance, with the best configuration achieving a Precision of 0.47, Recall of 0.66, mAP@.5 of 0.527, and mAP@.5:.95 of 0.349. However, the binary classification of pruning remains a challenge due to the continuous nature of the process.**

***Index Terms*—computer vision, tree pruning detection, object detection, YOLOv7-tiny, drone vision, precision agriculture.**

I. Introduction

Agriculture is an essential aspect of human life, responsible for feeding the global population and contributing to the economy. Precision agriculture, an approach that utilizes advanced technology to optimize crop production, has drawn increasing attention for its potential to increase efficiency and yield. One of the critical tasks in precision agriculture is monitoring the health and growth of trees in orchards. Accurate, real-time information about tree conditions can facilitate timely interventions and ensure optimal growth.

One of the crucial factors in maintaining tree health and productivity is tree pruning. Proper pruning can promote growth, improve tree structure, and optimize fruit yield. In this context, the ability to efficiently and accurately detect pruned and unpruned trees in large-scale orchards can provide valuable insights to farmers and help them make informed decisions about tree management.

This report presents a Computer Vision project that focuses on detecting and classifying pruned and unpruned trees in high-resolution drone images of olive tree fields using the YOLOv7-tiny model for object detection. The YOLOv7-tiny model was chosen because of its lightweight architecture and balance between speed and accuracy, which is crucial for real-time drone-based monitoring systems. This project explores the effects of different hyperparameters on the model's performance and discusses the challenges associated with the binary classification of tree pruning in drone images.

The remainder of the report is organized as follows: Section II reviews relevant studies that have employed various methodologies and models for tree pruning detection and management, highlighting their contributions and differences in the context of precision agriculture. Section III details the materials and methods used in the project, including image collection and labeling, image processing, dataset preparation, and model selection and training. Section IV presents the experimental results and the performance of the model under different hyperparameters. Section V discusses the results, the effect of hyperparameters on model performance, and the inherent challenges in binary tree pruning detection. Finally, Section VI concludes the report and highlights future work directions.

II. Related works

Several studies have been conducted on the application of computer vision and machine learning techniques for tree pruning detection and monitoring in the context of precision agriculture. This section reviews three relevant studies that have employed different methodologies and models for tree pruning detection and management.

Jiménez-Brenes et al. (2017) propose an innovative procedure that combines unmanned aerial vehicle (UAV) technology and advanced object-based image analysis (OBIA) methodology for multi-temporal three-dimensional (3D) monitoring of olive trees pruned with three different strategies: traditional, adapted, and mechanical pruning [1]. The study focuses on quantifying the impacts of each pruning treatment on the projected canopy area, tree height, and crown volume of every tree over time. The results demonstrate the effectiveness of the proposed UAV-based methodology in automatically identifying individual trees and computing their primary 3D dimensions with high accuracy. The study also highlights the potential of UAV- and OBIA-based technology for designing site-specific crop management strategies in the context of precision agriculture.

De Castro et al. (2018) present a novel and robust object-based image analysis (OBIA) procedure based on Digital Surface Model (DSM) for three-dimensional (3D) grapevine characterization [2]. The procedure is designed to process aerial images collected with Unmanned Aerial Vehicles (UAVs) and is fully automatic and self-adaptive to different crop-field conditions. The algorithm classifies grapevines and row gaps, computes vine dimensions, and generates georeferenced information on vine position, projected area, and volume without any user intervention. This research emphasizes the potential of UAV- and OBIA-based technology as a tool for site-specific crop management applications in precision viticulture.

Di Gennaro et al. (2020) investigate the reliability of Remote Sensing (RS) techniques for the estimation of pruning biomass through differences in the volume of canopy trees and evaluates the performance of an unsupervised segmentation methodology for the analysis of large areas [3]. The study uses data acquired by an Unmanned Aerial Vehicle (UAV) equipped with a multispectral camera on four uneven-aged and irregularly spaced chestnut orchards in Central Italy. The results show that UAV monitoring provides good performance in detecting biomass reduction after pruning, despite some differences between the trees' geometric features. The proposed unsupervised methodology demonstrates good performance in tree detection and vegetation cover evaluation, suggesting that it can provide effective strategic support for chestnut orchard management in line with a precision agriculture approach.

In summary, the related works discussed above have employed UAV technology combined with different image analysis methodologies, such as OBIA and unsupervised segmentation, for tree pruning detection and management. While these approaches have demonstrated promising results, this project aims to explore the use of the YOLOv7-tiny object detection model for similar purposes, focusing on the binary classification of pruned and unpruned trees in drone images of olive tree fields.

III. Materials and Methods

This section details the materials and methods used in the tree pruning detection project, covering the process of labeling and processing images, dataset preparation, and model selection and training.

A. Image Collection and Labeling

The project utilized two high-resolution TIF images of olive tree fields in Apulia, Italy, captured using a drone. Each image contained both pruned and unpruned trees. The trees in both images were manually labeled using the Computer Vision Annotation Tool (CVAT), an open-source image annotation tool [4]. Pruned trees were assigned the label 0, and unpruned trees were given the label 1. The resulting annotated images were exported in PASCAL VOC format. A partial view of the result of labeling one of the original images is shown in Figure 1.

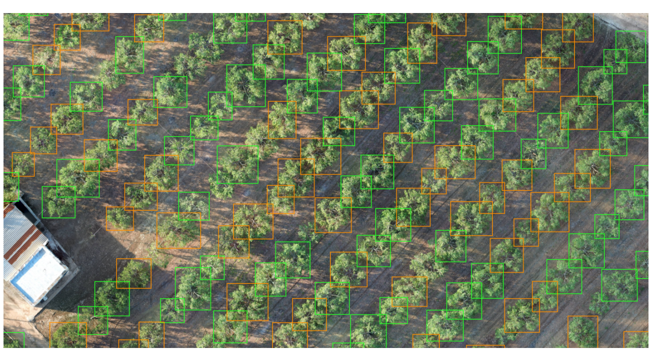

Fig. 1: Partial view of one of the original images after labeling.

B. Image Processing and Dataset Preparation

A custom Python script was developed to divide the two images into smaller 768x768 pixel squares. The size of the squares was decided heuristically, with the aim of each image containing at most 5-6 trees. Each square had a corresponding XML file containing information about the tree bounding boxes and their respective labels.

The resulting dataset had an imbalance between the number of pruned and unpruned trees. To address this issue, a Python script was developed to perform offline data augmentation. Images containing at least one unpruned tree were rotated once by a random angle to balance the dataset. The optimal number of rotations was determined heuristically to be one per image, as many images with unpruned trees also contained pruned trees. Thus, augmenting the former also augmented the latter.

It is important to note that not all images in the dataset contained labeled trees. Some images only displayed terrain, roads, immature trees, or trees that could not be definitively classified as pruned or unpruned. Consequently, the label files for these images did not contain bounding boxes or labels.

C. Model Selection and Training

After reviewing state-of-the-art models for object detection, two main models were considered: YOLOv7 [5] and InternImage-H [6]. Although InternImage-H outperformed YOLOv7 on several datasets, its recent release and ongoing development made YOLOv7 a more reliable choice for this project.

YOLO (You Only Look Once) is a widely recognized real-time object detection algorithm known for its efficiency and performance in various applications, including multi-object tracking, autonomous driving, robotics, and medical image analysis. YOLOv7 is the most recent iteration in the YOLO series, surpassing its predecessors and other contemporary object detectors in both speed and accuracy and improving detection accuracy without increasing the inference cost.

The YOLOv7-tiny version, a lightweight variant of the YOLOv7 model, was designed to meet different service requirements while maintaining high detection accuracy. With reduced parameters and computation, YOLOv7-tiny is ideal for deployment on devices with limited resources. The model outperforms previous YOLO versions, such as YOLOv4-tiny-31, in average precision (AP) while having fewer parameters and lower computation, demonstrating the effectiveness of the optimizations introduced in YOLOv7.

The YOLOv7-tiny model was chosen for this project due to its fewer parameters and faster training and inference times, making it suitable for drone vision applications (edge computing).

To prepare the dataset for training, a Python script was developed to randomly split the dataset into training, validation, and test sets. Labels were also converted from VOC to YOLO format. The model was trained and tested using various hyperparameters, leading to different results, which are discussed in the following sections.

IV. Experimental results

In this section, the results of the three experiments conducted to evaluate the performance of the YOLOv7-tiny model in detecting and labeling pruned and unpruned trees are presented. Table I shows the hyperparameters used for each experiment, and Table II presents the performance metrics obtained for each experiment, including Precision, Recall, mAP@.5, and mAP@.5:.95. Additionally, a grouped bar chart is provided in Figure 2 to showcase the difference in performance metrics.

TABLE I: Hyperparameter Comparison

Exp	Epochs	LR0	Final LR	Optimizer	Weight Decay
1	50	0.01	0.01	SGD	0.0005
2	100	0.001	0.0001	Adam	0.0005
3	200	0.001	0.0001	Adam	0

TABLE II: Experiment Results Comparison

Experiment	Precision	Recall	mAP@.5	mAP@.5:.95
1	0.348	0.542	0.308	0.174
2	0.395	0.711	0.53	0.335
3	0.47	0.66	0.527	0.349

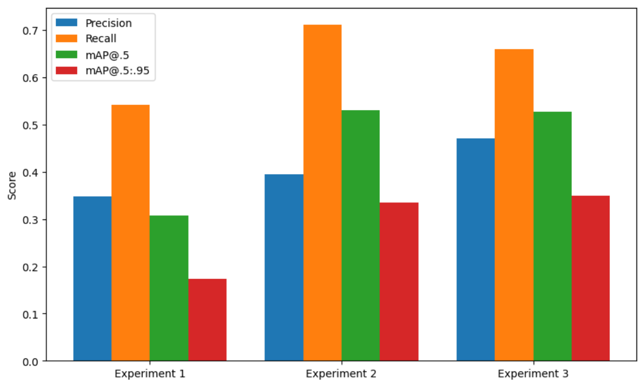

Fig. 2: Grouped bar chart of experimental results.

V. Discussion

This section discusses the results of the experiments and tries to provide insights into the effect of hyperparameters on the model's performance. Additionally, it addresses the inherent challenges of binarizing tree pruning detection.

A. Hyperparameter Effects on Model Performance

Increasing the number of epochs in Experiment 2 and 3 (from 50 to 100 and 200) led to improvements in all performance metrics compared to Experiment 1. This suggests that more training iterations allowed the model to better learn the features of pruned and unpruned trees. However, the increase in performance metrics from Experiment 2 to Experiment 3 was not as significant as from Experiment 1 to Experiment 2, indicating diminishing returns in performance improvement as the number of epochs increases.

The initial learning rate was reduced from 0.01 in Experiment 1 to 0.001 in Experiment 2 and 3. This change, along with the final learning rate reduction in Experiment 2 and 3, may have allowed the model to converge more effectively to an optimal solution. The better convergence can be observed in the higher performance metrics of Experiment 2 and 3 compared to Experiment 1.

Changing the optimizer from SGD in Experiment 1 to Adam in Experiment 2 and 3 contributed to better performance metrics. Adam is known to adapt the learning rate for each parameter, which can lead to faster convergence and improved generalization. This could be one of the reasons for the better performance in Experiment 2 and 3.

Experiment 3, with no weight decay, showed slightly better performance than Experiment 2, which used a weight decay of 0.0005. This suggests that the model might have been more prone to underfitting when using weight decay in the specific problem of tree pruning detection.

B. Challenges in Binarizing Tree Pruning Detection

The performance metrics of the experiments, although improved with hyperparameter tuning, were not exceptionally high. One possible reason for this is the inherent difficulty of binarizing the task of tree pruning detection. The task of tree pruning is a continuous process, making it challenging to label a tree as definitively ”pruned” or ”unpruned.” In some cases, a tree might be in an intermediate state, making it difficult even for a human to accurately categorize it.

This challenge is further exacerbated in drone vision for precision agriculture, as the quality and perspective of the images may vary. Factors such as lighting conditions, tree density, and occlusions may hinder the ability of the model to detect and classify pruned and unpruned trees accurately.

Moreover, the growth monitoring process in precision agriculture might require more granular information about the tree's pruning state rather than a binary categorization. A more appropriate approach could involve estimating the degree of pruning for each tree, which could provide more valuable information for agricultural decision-making.

C. Impact of YOLOv7-tiny Architecture

Another aspect to consider in the discussion of the model's performance is the choice of the YOLOv7-tiny architecture. Although this architecture is known for its balance between speed and accuracy, it may have limited the model's ability to learn more intricate features of pruned and unpruned trees. The tiny version of YOLOv7 has fewer layers and parameters than its full counterpart, which could have impacted the model's performance in our specific problem. Exploring other architectures and their impact on the tree pruning detection task could be beneficial in future studies.

D. Dataset Quality and Labeling Challenges

Moreover, the quality of the dataset could play a significant role in the model's performance. The presence of mislabeled or ambiguous examples in the dataset might have hindered the model's ability to learn the features of pruned and unpruned trees effectively. Improving the dataset by removing ambiguous examples, adding more images with varying conditions, or employing a domain expert's assistance to ensure accurate labeling could lead to better performance in detecting tree pruning.

E. Key Insights

To summarize the discussion, the experiments demonstrate the effects of different hyperparameters on the performance of the YOLOv7-tiny model in the tree pruning detection task. The results suggest that the model benefits from an increased number of epochs, reduced learning rates, and the use of the Adam optimizer. However, the performance metrics, while improved through hyperparameter tuning, were not exceptionally high. This can be attributed to the inherent difficulty of binarizing the tree pruning detection task, which involves categorizing trees as definitively "pruned" or "unpruned," even when trees may be in intermediate states. Additionally, challenges specific to drone vision in precision agriculture, such as varying image quality and environmental conditions, further impact the model's performance. Furthermore, the choice of the YOLOv7-tiny architecture, which sacrifices some complexity for faster inference, and the quality of the dataset, including potential labeling challenges and ambiguities, may have influenced the model's ability to learn intricate features.

VI. Conclusion

This report presented a Computer Vision project aimed at detecting tree pruning in images using the YOLOv7-tiny model. The experiments conducted demonstrated the impact of various hyperparameters on the model's performance, such as the number of epochs, learning rates, optimizer, and weight decay. The results revealed that increasing the number of epochs, reducing the learning rates, and using the Adam optimizer led to improvements in performance metrics. However, the model's performance was not exceptionally high, which could be attributed to the inherent difficulty of binarizing tree pruning detection and the challenges of drone vision in precision agriculture.

Future work for this project could involve exploring alternative approaches that estimate the degree of pruning rather than relying on a binary categorization, which would provide more granular information for agricultural decision-making. Additionally, improving the dataset quality by adding more images, ensuring accurate labeling, or seeking a domain expert's help could lead to better performance in detecting tree pruning. Experimenting with different models and comparing their performance in the tree pruning detection task could also help identify the most suitable architecture for this specific problem. Finally, integrating the developed model with a drone-based monitoring system could further enhance its practical application in precision agriculture, providing valuable insights for tree management and growth monitoring.

References

[1] F. Jiménez-Brenes, F. López-Granados, A. De Castro, J. Torres-Sánchez, N. Serrano, and J. M. Peña-Barragán, "Quantifying pruning impacts on olive tree architecture and annual canopy growth by using uav-based 3d modelling," *Plant Methods*, vol. 13, 07 2017. II

[2] A. De Castro, F. Jiménez-Brenes, J. Torres-Sánchez, J. M. Peña-Barragán, I. Borra-Serrano, and F. López-Granados, "3-d characterization of vineyards using a novel uav imagery-based obia procedure for precision viticulture applications," *Remote Sensing*, vol. 10, p. 584, 04 2018. II

[3] S. Di Gennaro, C. Nati, R. Dainelli, L. Pastonchi, A. Berton, P. Toscano, and A. Matese, "An automatic uav based segmentation approach for pruning biomass estimation in irregularly spaced chestnut orchards," *Forests*, vol. 11, p. 308, 03 2020. II

[4] OpenCV, Intel Corporation, and contributors, "Computer Vision Annotation Tool (CVAT)," https://github.com/openvinotoolkit/cvat, 2018. III-A

[5] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," 2022. III-C

[6] W. Wang, J. Dai, Z. Chen, Z. Huang, Z. Li, X. Zhu, X. Hu, T. Lu, L. Lu, H. Li, X. Wang, and Y. Qiao, "Internimage: Exploring large-scale vision foundation models with deformable convolutions," 2023. III-C