# Comparative evaluation of combined and task specific detectors for pomegranate yield and loss detection

Yuval Tenenboim[1,2], Yael Edan[1], Idit Ginzberg[3], Tarin Paz-Kagan[2]

[1] Dept. of Industrial Engineering & Management, Ben-Gurion University of the Negev, Israel

[2] The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Israel

[3] Institute of Plant Sciences, Agricultural Research Organization, Volcani Institute, Israel

## 1. Introduction

Fruit cracking is a major cause of yield loss in pomegranate orchards, including fruit waste, reduced fruit quality, increased susceptibility to pathogens, and decreased marketability. Reported yield losses due to cracking can reach 40-60%, depending on climatic conditions, irrigation and nutrition regimes, and orchard management practices (Ikram & Nafees, 2020). Despite the severity of these losses, most vision-based yield estimation systems in the orchards focus exclusively on detecting and counting healthy fruits on trees, neglecting defective fruits. This limitation constrains the operational relevance of such systems, as they fail to capture the full yield gap between potential and actual production. Recent advances in precision agriculture and computer vision, particularly deep learning-based object detection, have enabled automated fruit detection under complex field conditions. However, these approaches typically treat non-marketable fruits as noise rather than as a meaningful indicator of loss. As a result, there is a need for detection frameworks that can simultaneously quantify both yield and loss. This study addresses this gap by evaluating different object detection strategies for joint fruit yield and loss detection in pomegranate orchards. Specifically, we compare combined multi-class detectors with task-specific single-class detectors under conditions of class imbalance and field variability. The goal is to determine which modeling approach provides more reliable and operationally relevant performance for orchard monitoring.

## 2. Materials and Methods

### 2.1 Dataset and Data Collection

Image acquisition was conducted during two sampling campaigns in 2023 and 2024 in October, during the pomegranate ripening period, in a commercial orchard at Kibbutz Tsor'a, Israel (31°45′32.7″N, 34°57′17.9″E). High-resolution smartphone and DSLR images were collected from 30 sample trees at multiple viewing angles to ensure canopy coverage. Two cameras were used: a Xiaomi 2201116TG smartphone (4000 × 3000 pixels) and a Canon EOS 200D DSLR (6000 × 4000 pixels). Images were acquired under natural conditions, capturing variability in illumination, occlusion, and canopy complexity. The dataset included both healthy (yield) and defective (loss) fruits, including cracked and fallen fruits. A total of 203 annotated images were used, containing 18,863 healthy and 5,323 defective instances, reflecting strong class imbalance that challenges multi-class detection. To improve robustness, data augmentation was applied to the training set only. Spatial transformations included random cropping and rotation. Photometric augmentations included exposure adjustment, Gaussian noise, and Gaussian blur. These steps increased the training dataset sixfold to 852 images. Preprocessing was applied to all images using contrast-limited adaptive histogram equalization (CLAHE) and saturation enhancement to improve contrast and color separability.

### 2.2 Detection Models and Training Strategies

Object detection was performed using the YOLOv10s architecture (Wang et al., 2024), a single-stage detector known for its efficiency and suitability for real-time agricultural applications. Multiple training configurations were evaluated to assess the impact of model design and initialization strategy (Tenenboim, 2026). Four main training paradigms were considered: (i) Joint multi-class detector trained from scratch, (ii) Joint multi-class detector initialized with COCO pretrained weights (Lin et al., 2014), (iii) Task-specific single-class detectors (yield and loss) trained from scratch, and (iv) Task-specific single-class detectors initialized with COCO pretrained weights. The joint models were trained to detect both healthy and defective fruits simultaneously, while the task-specific models were trained separately for each class. This separation allows each model to specialize in a single detection task, potentially improving sensitivity to minority classes. Transfer learning was incorporated through initialization with COCO pretrained weights, which provides general visual feature representations and improves convergence during training.

### 3. Results

Joint multi-class models achieved moderate overall performance but showed reduced sensitivity to the loss class. Pretraining improved performance, increasing overall F1 from 0.62 to 0.69 and mAP@0.5 from 0.72 to 0.74, with notable gains in loss detection.

Task-specific models outperformed joint models. The loss-specific model improved by +0.14 (F1) and +0.18 (recall) compared to the joint model trained from scratch, while the yield model achieved the highest performance (F1 ≈ 0.79, mAP@0.5 ≈ 0.83). Overall, pretraining consistently improved results, and task-specific detectors provided more reliable performance under class imbalance.

**Table 1**. Detection performance metrics for the yield and loss models.

| Model | Class | Precision | Recall | F1 | AP@0.5 |
|---|---|---|---|---|---|
| Joint yield and loss model trained from scratch | Loss | 0.8930 | 0.3253 | 0.4769 | 0.7182 |
| Joint yield and loss model trained from scratch | Yield | 0.8491 | 0.6982 | 0.7663 | 0.7182 |
| Joint yield and loss model pretrained weights on COCO | Loss | 0.7904 | 0.4693 | 0.5889 | 0.7434 |
| Joint yield and loss model pretrained weights on COCO | Yield | 0.8486 | 0.7327 | 0.7864 | 0.7434 |
| Loss model, train from scratch | Loss | 0.8567 | 0.4033 | 0.5484 | 0.6507 |
| Loss model, pretrained weights on COCO | Loss | 0.7957 | 0.5022 | 0.6158 | 0.6604 |
| Yield model, train from scratch | Yield | 0.8532 | 0.7056 | 0.7724 | 0.8193 |
| Yield model, pretrained weights on COCO | Yield | 0.8508 | 0.7366 | 0.7896 | 0.8318 |

### 4. Discussion and Conclusion

Detection of defective fruits is challenging due to class imbalance, visual variability, and low contrast with the ground. These fruits are typically smaller, darker, and desiccated following pericarp cracking, and are often partially occluded or visually similar to background elements, reducing recall. Under these conditions, joint multi-class models are particularly affected, as the dominant yield class biases the learning process and suppresses detection of the minority loss class, whereas task-specific detectors improve performance by focusing on class-specific features. Transfer learning with COCO pretrained weights further enhanced detection accuracy, particularly for the loss class. Together, these findings indicate that task-specific optimization combined with transfer learning is essential for reliable yield and loss detection under class imbalance (Tenenboim, 2026). Future work will integrate these models into a tracking-by-detection framework for joint yield and loss estimation. In this setting, each video frame provides an additional opportunity for detection, allowing missed instances in individual frames to be recovered over time. This approach is expected to improve overall estimation accuracy and support a reliable, operational assessment of orchard productivity in precision agriculture.

### References

Ikram, S., & Nafees, M. (2020). Causes and control of fruit cracking in pomegranate: A review. Journal of Global Innovations in Agricultural and Social Sciences, 8(4), 183–190. https://doi.org/10.22194/jgiass/8.920

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. European Conference on Computer Vision (ECCV), 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Tenenboim, Y. (2026). Tracking-by-detection framework for simultaneous tree-scale pomegranate yield and fruit loss estimation from UAV [M.Sc. thesis]. Ben-Gurion University of the Negev.

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-Time End-to-End Object Detection. http://arxiv.org/abs/2405.14458