Review of: "Comparing YOLOv8 and Mask RCNN for object segmentation in complex orchard environments"

Teerapong Panboonyuen

Potential competing interests: No potential competing interests to declare.

This paper compares the one-stage YOLOv8 and two-stage Mask R-CNN models, for instance, segmentation in agricultural image processing. It assesses their performance in delineating objects within orchard images, focusing on apple tree structures across dormant and early growing seasons. YOLOv8 exhibited superior accuracy, achieving higher precision and recall metrics and faster inference rates, suggesting its suitability for real-time agricultural applications like robotic tasks.

Strengths:
- The paper is well-written, structured, and easy to read.
- This paper conducts a meticulous comparison between YOLOv8 and Mask R-CNN models, evaluating their performance across varying orchard conditions and datasets, providing a comprehensive assessment for agricultural applications.
- It effectively highlights YOLOv8's superiority in accuracy and efficiency, emphasizing its potential for immediate implementation in time-sensitive agricultural tasks like automated pruning, thereby offering practical implications for the farming industry.

Weaknesses:
- In Table 2. Scene interpretation accuracy of segmented image patches was written in a confusing way. Way -it needs to be more apparent to me what dataset splits are used for which datasets or what data was validation vs. testing set. Recommend rewording for clarity.
- It would be helpful to list the performance of each labelling function and each label model (even if only on a subset of the validation or test dataset) instead of only the final classification score.
- The evaluation primarily focuses on apple tree-related imagery, potentially limiting the generalizability of findings across diverse crop types or agricultural contexts.
- While the paper outlines performance metrics, a deeper technical analysis comparing model architectures or algorithmic intricacies could enhance the study's credibility and usefulness.
- The paper lacks an in-depth discussion regarding the limitations or specific scenarios where each model might face challenges, which could provide a more nuanced understanding of their applicability.
- A deeper exploration of dataset characteristics and their influence on model performance, especially in addressing varying light conditions, could enrich the paper's insights.
- While emphasizing real-time applications, the paper could benefit from discussing practical challenges in implementing these models in real agricultural settings, considering environmental variability and hardware constraints.