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

Abdussalam Elhanashi[1]

1 University of Pisa

Potential competing interests: No potential competing interests to declare.

This study compares the performance of one-stage YOLOv8 and two-stage Mask R-CNN machine learning models for instance segmentation in agricultural images. The focus is on delineating individual objects in orchard conditions, specifically dormant apple trees and tree canopies with green foliage and immature apples.

Some important points have to be clarified or justified and few concerns need to be addressed by the authors for the betterment of the manuscript

1. what is the primary objective of the study, and how does instance segmentation play a crucial role in achieving this objective, especially in the context of agriculture automation?
2. What motivated the study to compare the YOLOv8 and Mask R-CNN models for this application?
3. Why did the study specifically choose YOLOv8 and Mask R-CNN for comparison in instance segmentation?
4. Can you elaborate on the key innovations in YOLOv8's architecture?
5. In the results section, what challenges or limitations were identified in Mask R-CNN's performance, especially in scenarios with varying lighting conditions and complex backgrounds? How did YOLOv8 demonstrate better adaptability in these situations?
6. Considering the practical applications in agriculture, how do the computational speeds (inference times) of YOLOv8 and Mask R-CNN influence their suitability for real-time tasks, such as robotic harvesting and pruning?
7. How do the findings of this study contribute to the advancement of precision agriculture and the development of smart and automated orchard operations? are there any recommendations for future research in this field?