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

Seyed Iman Saedi[1]

1 Shahrood University of Technology

Potential competing interests: No potential competing interests to declare.

The authors have conducted research comparing YOLOv8 and Mask-RCNN for object segmentation in complex orchard environments. This is a significant and innovative topic, particularly in the context of robotic fruit harvesters and prunners. The hypothesis is clearly stated, and the research provides valuable insights. The study has been conducted scientifically, and the results are open to discussion. The challenges regarding object detection in fully natural orchard conditions have been well addressed. The authors have attempted to evaluate the performance of two state-of-the-art computer vision systems in detecting single-class and multi-class objects in an apple orchard. The models have been assessed, considering challenges such as complex background, variable light, overlaps, and occlusions with other plant parts. Additionally, the authors have addressed challenges regarding the running time of a method, ensuring that a proposed algorithm can recognize objects for on-the-go applications.

To further improve the scientific rigor, clarity, and evaluation of the research, the following points should be addressed:

1. Correct "Mast-RCNN" to "Mask-RCNN" in the abstract.

2. Discuss how the authors claim the generalization property of the models, for example, in terms of fruit species/variety or the trunk and branches in non-dormant periods.

3. Indicate the units of augmentation values in parentheses (page 7) to clarify that these are probabilities of applying the augmentation or manipulation values.

4. Provide the train and validation trend curves for the two models in question to gain insight into the stability of the models.