Broccoli head segmentation: How well does it actually work?

Pieter van den Haspel | p.b.a.vandenhaspel@student.tudelft.nl

1. Introduction

- Smart Farming: Using technology to estimate crop growth for optimal harvesting
- Head segmentation: Algorithms extract the broccoli head from images

Question

How do you know which algorithm works best?

- answer: use metrics to evaluate the algorithms and use the results to compare them
- IoU: Intersection of union, used to see how close the algorithm's guess of the broccoli head is to the ground truth (literally)
- MAE: Mean Absolute error, the mean of all the differences between the measured size and the actual size.
- PA: Pixel accuracy, how many of the pixels are labelled correctly, broccoli, or no broccoli
- F1: The harmonic mean of precision (how many correct guesses) and recall (how many of the total were guessed correctly)

Question

Do the papers presenting these algorithms use the same metrics?

2. Research Question

NO! Therefore the aim is to:

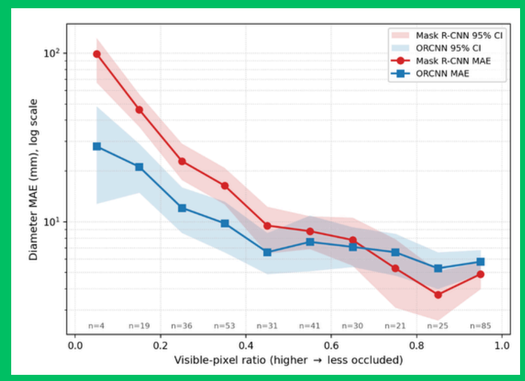
1. Make a taxonomy of all metrics used
2. Dive into how different variables change the evaluation of these algorithms
3. Look at the full pipeline to discover how much impact a change in these metrics actually has

And this is how we will discover the answers...

3. Experiments

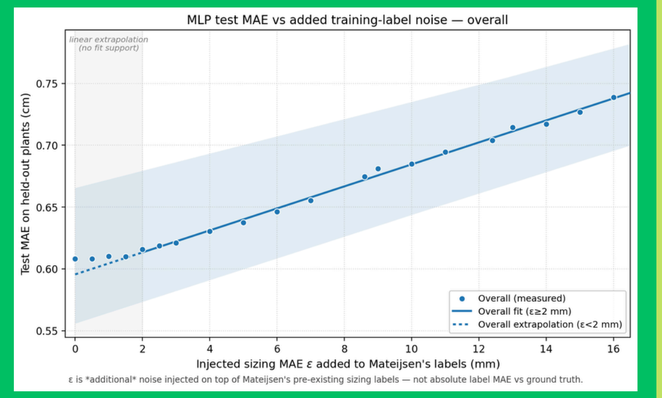
Stratification: Run three models on evaluation data stratified by: Leaf Occlusion, size, and time of day
Robustness: create perturbations in existing data and evaluate models on this data to see if it shows distinctive results between models, variables: blur, sunlight + moisture, camera model
Predictions: Set up Matheijsens pipeline and inject sizing errors to see the downstream errors.

4. Results



Comparison of two models, the accuracy lead switches at a visibility of 70% between Mask R-CNN (a model trained on visible masks) and ORCNN (trained on amodal masks

The added error onto the already present error in the data shows a linear correlation of 0.0089 cm error in prediction per mm error in training data



5. Conclusion

1. The occlusion and size are variables that, when used to stratify evaluation, show operationally meaningful differences between models
2. Blur and sunlight + moisture are valid variables for exposing the robustness of a model in a synthetic way, but camera models are not
3. There is only a gain of 2 hours in growing accuracy from reducing the MAE of head size estimation models to 0, this is an upper bound

6. Discussion

The synthetic data generated for the robustness might not reflect the real world perfectly, leaving room for another experiment with data gathered from the actual fields to test how reliable the synthetic creation is in replicating the real world