SMART FARMING

A Systematic Comparison of Image Preprocessing Techniques for YOLOv8n under Dutch Field Conditions

Author

Huan Yi Ma

Email: H.Y.Ma-3@student.tudelft.nl

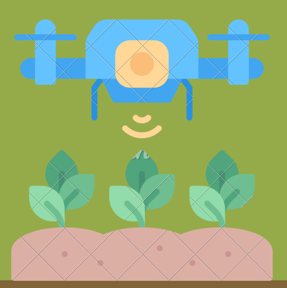
TU Delft

inholland university of applied sciences

Supervisors

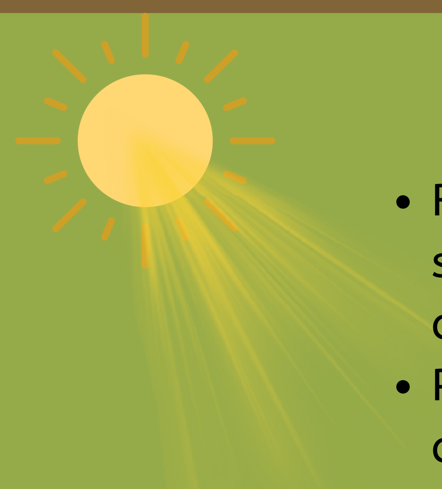
Cynthia Liem, Jeroen Wildenbeest, Surya Giri

1. Introduction



- Currently in Dutch broccoli farming, each broccoli is assessed and harvested by hand.
- Manual labor is intensive, expensive, and scarce.
- Whole-field harvesters cut entire fields at once, leading to crop waste (uneven growth).
- Individual broccoli sizes can be estimated with the use of computer vision.
- The first step is to accurately detect broccoli in raw field imagery.

2. Challenge



- Raw field imagery is often not perfect: glaring sun, deep shadows, leaf occlusion, water droplets, camera blur, etc.
- Preprocessing can improve the input for the detection model.
- **Main Question:** *How can image enhancement techniques improve automated broccoli head segmentation and size estimation in raw field imagery?*

3. Methodology

01 **Dataset:** We use 394 annotated field images of Verdonk broccoli that were captured over six measurement days in July 2025. The dataset contains the target conditions: blur, harsh lighting, and occlusion.

02 **Preprocessing:** We preprocess the dataset using 5 common preprocessing techniques (CLAHE, unsharp masking, wavelet transform, median filtering, bilateral filtering) and 5 combinations using these techniques.

03 **Training:** We partition each variant with 5-fold cross-validation with a fixed seed and a separate YOLOv8n detection model is trained on each variant using identical hyperparameters.

04 **Evaluation:** We use precision, recall, mAP@0.50, mAP@0.50-0.95 to evaluate the performance of each variant. Performance of each variant is compared to the performance on the raw data (baseline).

05 **Difficult subsets:** We inspect the worst detections from the baseline and assign them to the *blur*, *lighting*, *occlusion*, or *mix* subset. Then compare the results of the variants to the baseline for each subset.

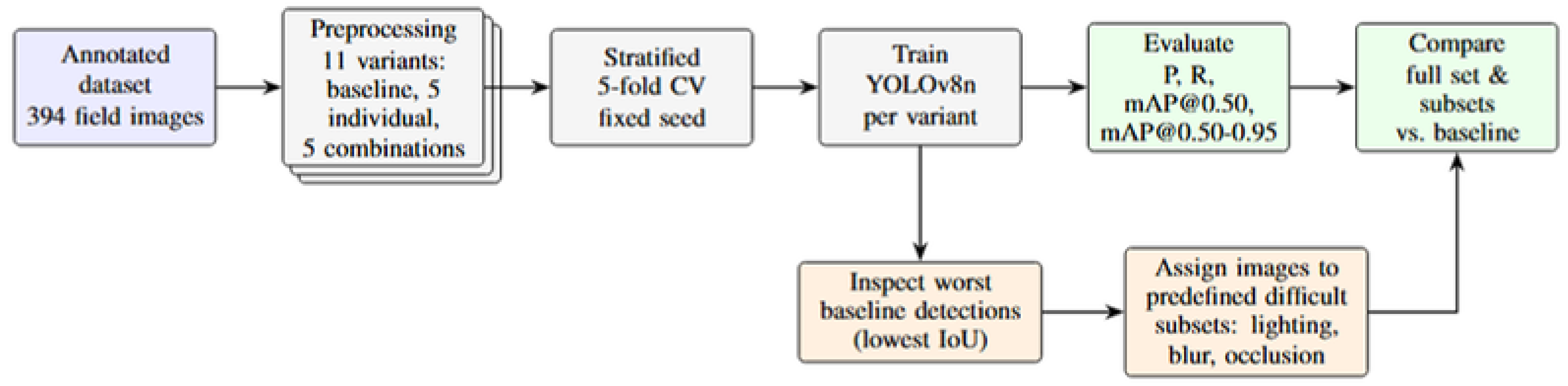


Figure 1: Overview of the experimental pipeline.

4. Results on full dataset

- **Q: Does preprocessing improve performance on the full dataset?**
- All 11 variants are within ~1 percentage point on mAP@0.50 and half percentage on mAP@0.50-0.95.
- Differences between variants are of same order as the standard deviation between folds.
- **A: No single preprocessing variant produces a decisive change in overall detection accuracy compared to the baseline.**

Variant	Precision	Recall	mAP@0.50	mAP@0.50-0.95
Raw (baseline)	**0.969** ± 0.006	0.963 ± 0.016	0.980 ± 0.007	0.748 ± 0.012
CLAHE	0.967 ± 0.015	0.974 ± 0.014	0.986 ± 0.006	**0.751** ± 0.019
Bilateral	0.963 ± 0.012	0.959 ± 0.012	0.984 ± 0.007	0.748 ± 0.015
Unsharp	0.964 ± 0.015	**0.976** ± 0.008	0.988 ± 0.004	0.749 ± 0.016
Wavelet	0.964 ± 0.007	0.965 ± 0.012	0.979 ± 0.009	0.746 ± 0.016
Median	0.968 ± 0.014	0.969 ± 0.014	0.987 ± 0.005	0.745 ± 0.011
Bilateral + Unsharp	0.956 ± 0.032	0.967 ± 0.014	0.982 ± 0.008	0.746 ± 0.014
CLAHE + Unsharp	0.951 ± 0.017	0.969 ± 0.018	0.983 ± 0.009	0.746 ± 0.012
Median + CLAHE	0.966 ± 0.016	0.966 ± 0.019	0.986 ± 0.003	0.747 ± 0.011
Wavelet + Unsharp	0.963 ± 0.018	0.972 ± 0.023	**0.989** ± 0.001	**0.751** ± 0.019
Bilateral + CLAHE + Unsharp	0.957 ± 0.017	0.957 ± 0.021	0.980 ± 0.009	0.746 ± 0.014

5. Results on the subsets

- **Q: Does preprocessing improve performance on difficult subsets?**
- **Blur subset:** wavelet+unsharp mAP@0.50 0.973 → 0.995.
- **Lighting subset:** Multiple techniques increase recall (more broccoli are found).
- **Occlusion subset:** Edge preserving denoising and contrast enhancement increase mAP@0.50-0.95 (better accuracy)
- **Mix subset:** Worst performing baseline. Every technique increases recall (more broccoli are found).
- **A: Different preprocessing techniques help in different situations.**

Table 5: Performance on the *mix* subset.

Variant	Precision	Recall	mAP@0.50	mAP@0.50–0.95
Raw (baseline)	0.972	0.825	0.844	0.624
CLAHE	0.927	0.950	0.936	0.697
Bilateral	0.921	0.875	0.869	0.659
Unsharp	0.951	**0.971**	**0.971**	0.702
Wavelet	0.972	0.875	0.870	0.649
Median	0.923	0.950	0.933	0.682
Bilateral + Unsharp	0.902	0.924	0.915	0.658
CLAHE + Unsharp	0.969	0.950	0.942	0.683
Median + CLAHE	**1.000**	0.933	0.945	**0.705**
Wavelet + Unsharp	0.881	0.925	0.912	0.703
Bilateral + CLAHE + Unsharp	0.920	0.864	0.858	0.651

6. Takeaways

- Preprocessing is a tool for specific difficult field conditions rather than as a general improvement for detection.
- Often trades a little precision for more recall.
- Combining techniques is not automatically beneficial. They need to complement eachother and target a specific goal.

7. Limitations

- The subsets are small (36-44 broccoli heads) causing small changes to move the metrics considerably.
- No significance testing was performed, making the results indicative rather than decisive measurements.
- Subsets were handpicked by non-expert.
- Only one dataset was used, which also contained some inaccurate annotations.