Two-stage approach for detection and classification of lettuce plants

Veronika Jadrná, Minor I'M AI – Artificial Intelligence for Image Data
Supervisor: Willem Dijkstra

Winter 2025

Introduction

- Monitoring the crop development process is one of the key aspects in farming, as environmental conditions greatly influence plant growth and overall productivity, especially for sensitive crops like lettuce [1].
- Traditional crop monitoring techniques rely on manual inspection, which is labour-intensive, time-consuming and prone to errors.
- The use of computer vision addresses these issues by automating the process, providing real-time data and reducing false results.

Abstract

In this research we focus on applying computer vision to automate lettuce crop monitoring. The primary objective is to detect lettuce plants at various growth stages and classify them based on the diagonal sizes of their bounding boxes as small, medium or large. The methodology consists of two-stage approach for detection and classification. The study employed YOLOV5s model for a single class object detection and tested and evaluated three different classification methods. The best-performing was a distance-based classification, which evaluates the likelihood of each class relative to the statistical metrics of the data.

Materials and Methods

- A video dataset was obtained in an open agricultural field as shown in Figure 1.
- Individual frames were extracted (Figure 2) from the videos with a resolution of 1080x1920 and subsequently annotated (Figure 3) in CVAT [2].
- In total, 140 images were annotated and split into training set (70%), validation set (15%) and testing set (15%).
- YOLOv5 object detection model [3] was trained using image tile size of 512x512. Testing images were displayed in their original dimensions.



Figure 1. A farming vehicle capturing video footage of the lettuce plants in the field.



Figure 2. Example extracted frames depicting different growth stages.

- To classify lettuce plants as small, medium or large, three methods were evaluated:
- **Dynamic Range Classification** - uses standard deviation and mean of each growth stage to define ranges.
- **Fixed Thresholding** - fixed pixel ranges are defined for each growth stage.
- **Distance-Based Classification** - compares the distance of a diagonal value to the statistical metrics of each growth stage.

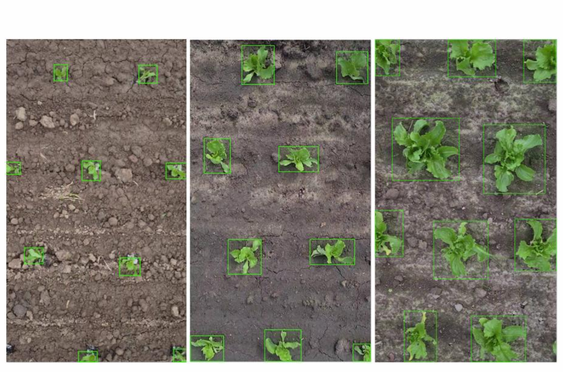


Figure 3. Example annotated frames depicting different growth stages.

Experiments and Results

- Tuning of the hyperparameterers and data augmentation played an important role in optimizing the model's performance.

F1 score	98%
Precision	98%
Recall	99%
MAp	90%

Table 1. Results of YOLOv5s in object detection.



Figure 4. Output images of YOLOv5s with predicted white bounding boxes.

- **Dynamic range classification**: overlapping ranges and gaps, depicted in Figure 5, would lead to inconsistent results.
- **Fixed pixel thresholding:** reliance on static thresholds makes it unable to adjust to different grown patterns and overlaps in the dataset.
- **Distance-based classification:** achieved best results with margin exclusion technique, that excluded edge bounding boxes from the metrics and resulted in reduced overlapping data. The results of this method can be seen in Table 2 and example true and false classifications in Figure 6.

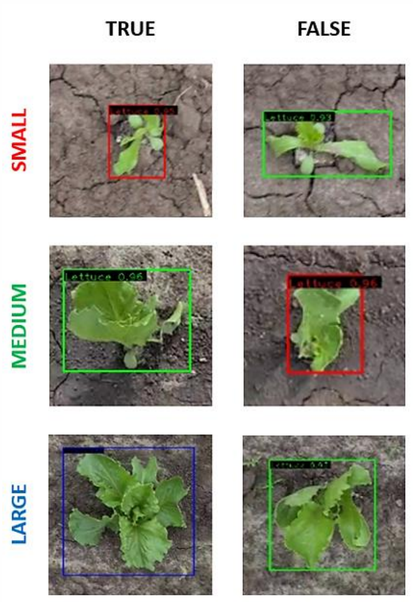


Figure 6. Visual comparison of true and false classification examples for each growth stage, where red color of bounding box is used for small lettuce, green for medium and blue for large.

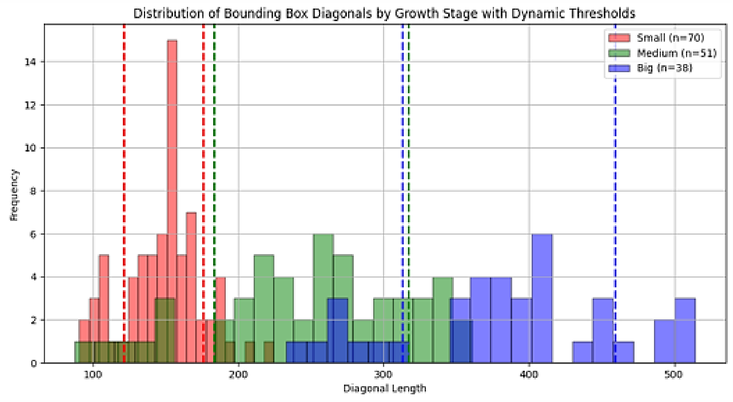


Figure 5. Distribution of diagonals for each growth stage with predicted ranges (dashed lines) for each growth stage.

	Before Margin		After Margin	
Stage	**Precision**	**F1 score**	**Precision**	**F1 score**
Small	90%	88%	94%	96%
Medium	67%	67%	88%	86%
Large	75%	77%	91%	89%

Table 2. Precision and F1 score before, and after applying margin technique for distance-based classification.

Conclusions

- The model achieved a high accuracy in detecting individual lettuce plants and therefore provided a solid base for classification process.
- The comparison of distances across all stages ensured, that the most appropriate class is assigned to the lettuce plants, even in ambiguous scarious. Therefore, the distance-based classification handles the overlaps between data more accurately, compared to the other methods tested.
- Misclassifications occurred when some plants differed significantly in size from other plants within their group, resulting in being categorized into a different stage. However, in most cases, such classifications may not truly be errors.

Acknowledgements

- This project is financially supported by the European Union within the JTF-call project DigiAgro3.

NHL STENDEN computer vision & data science

TRACK 12

SNN Samenwerkingsverband Noord-Nederland • Stimuleert • Faciliteert • Verbindt

Medegefinancierd door de Europese Unie

Ministerie van Economische Zaken

References

- [1] A. Bauer et al., Combining computer vision and deep learning for aerial phenotyping: A case study of lettuce production, *Horticulture Research*, vol. 6, p. 70, 2019.
- [2] CVAT: Computer Vision Annotation Tool, CVAT.ai, 2025. URL: https://www.cvat.ai/.
- [3] G. Jocher, R. Munawar, et al., "Ultralytics YOLOv5 Documentation," Ultralytics Inc., 2025. URL: https://docs.ultralytics.com/models/yolov5/.