Fine-Grained Object Detection for Precision Beekeeping

Gabriela VDOVIAK, Tomyslav SLEDEVIČ

E-mail: {gabriela.vdoviak|tomyslav.sledevic}@vilniustech.lt,

Department of Electronic Systems, Vilnius Gediminas Technical University (VILNIUS TECH)

Introduction

Real-time and non-invasive honeybee monitoring at the hive entrance is crucial for the advancement of precision beekeeping and the maintenance of colony health. Conventional monitoring methods are characterized by their labor-intensiveness and disruptive nature, which motivates the development of autonomous vision-based systems. In this setting, detection models must handle challenging real-world scenarios, including lighting variations, occlusions, motion blur, and fine-grained distinctions between insect classes and small objects such as pollen grains.

Aims and Goals

In this study, we investigate YOLO models for hive entrance monitoring, with an emphasis on enhancing fine-grained class detection. This work aims to:

- modify a YOLO model for fine-grained detection of small insect targets;
- evaluate detection performance in challenging field conditions;
- compare two annotation strategies for pollen-carrying honeybees.

Methods

For the experiments, we used a dataset of 11008 high-definition frames (1920×1080 px) captured by cameras positioned above hive landing boards across multiple apiaries. The frames represent natural bee traffic under diverse real-world conditions, including varying illumination, motion blur, occlusions and complex backgrounds. All images were manually annotated with bounding boxes for three target classes: 116853 honeybees, 14062 pollen loads, and 1405 wasps. The samples from the dataset are shown in Figure 1.

To address the class imbalance between honeybees and pollen, two annotation strategies were applied:

- pollen merged with honeybees into a single pollen-bee class;
- pollen kept as a separate object class.

Figure 1: Samples of labeled images from the dataset used for model training and evaluation.

YOLOv8 s was used as the baseline detector and adapted into two variants to enhance the detection of small objects with reduced computational cost. The dataset was split into 80% for training and 20% for testing, and all three models were trained for 500 epochs at an input resolution of 1024×576 pixels using geometric and photometric data augmentations. Performance was evaluated using per-class true positive rate and mAP@0.5.

Results

Detection performance of the YOLOv8 s baseline and two modified models is summarized in Table 1. The results indicate that modifications improved pollen detection and reduced model size, while the combined pollen-bee labeling strategy achieved the highest overall mAP@0.5. Figure 2 shows YOLOv8 s detection results with separate honeybee and pollen classes.

Table 1: Comparison of model performance using two annotation strategies: separate pollen class and combined pollen-bee class.

Separate Pollen Class					
Model	Params, M	Bee, TPR %	Wasp, TPR %	Pollen, TPR %	mAP@0.5, %
YOLOv8 s	9.8	98	97	59	94.1
YOLOv8 s-1	4.7	97	97	64	96.6
YOLOv8 s-2	4.5	97	97	72	96.3
Combined Pollen-bee Class					
Model	Params, M	Bee, TPR %	Wasp, TPR %	Pollen-bee, TPR %	mAP@0.5, %
YOLOv8 s	9.8	96	96	84	96.9

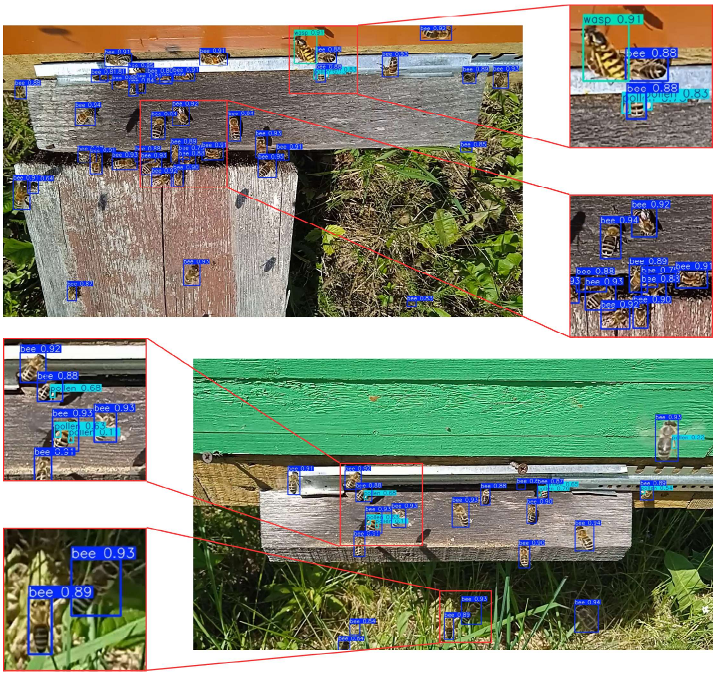

Figure 2: Detection results with separate honeybee and pollen classes (YOLOv8 s). Blue, cyan and green boxes represent honeybees, pollen loads and wasps, respectively.

Conclusions

- The results highlight a trade-off between model size and detection accuracy, and show that small-object-oriented architectural changes can enhance fine-grained detection without sacrificing overall performance.
- The choice between separate pollen and combined pollen-bee labels depends on the application. The combined class improves detection of pollen-carrying honeybees, while a separate pollen class is preferable for tasks such as honeybee tracking and detailed pollen analysis.
- These results provide a basis for extending hive entrance monitoring to include motion-based multi-object tracking and behavioral analysis.