Detection of different bee types at the beehive entrance

Kreete Kuusk, Danni Zhang, Jasper Luik and Rasmus Mirma

Course project in Machine Learning - MTAT.03.227
2025

Artjom Kurapov

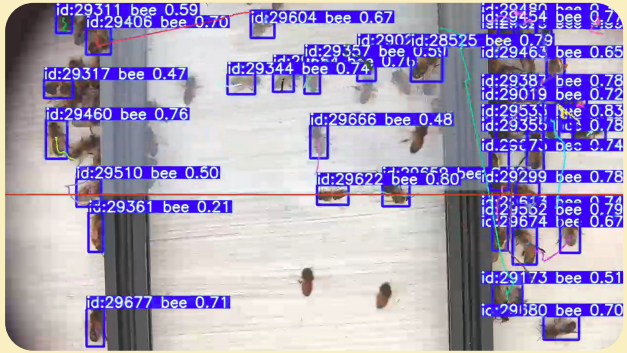
Creating a beehive that **inspects bees** automatically

GRATHEON

Artjom Kurapov

Creating a beehive that **inspects bees** automatically

Only bee detection



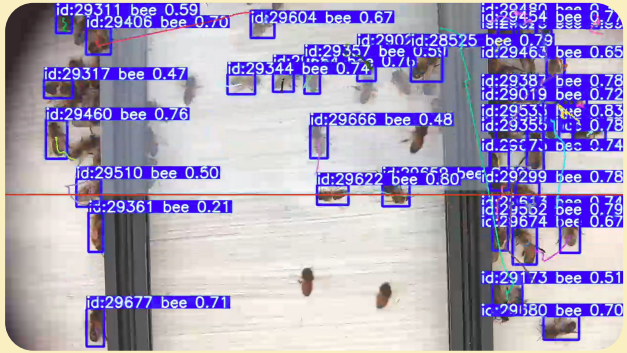
https://github.com/Gratheon/entrance-observer/

GRATHEON

Artjom Kurapov

Creating a beehive that **inspects bees** automatically

Only bee detection

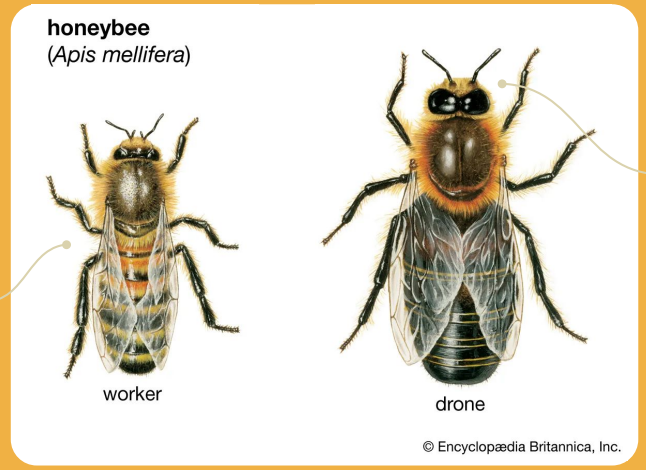


https://github.com/Gratheon/entrance-observer/

"Worker bees represent 80-99% of the colony population."

- Bruce Rutter, bestbees.com

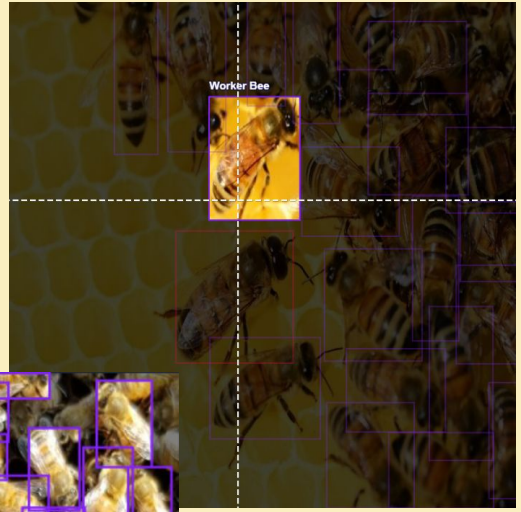
Drone and **worker** bee detection!

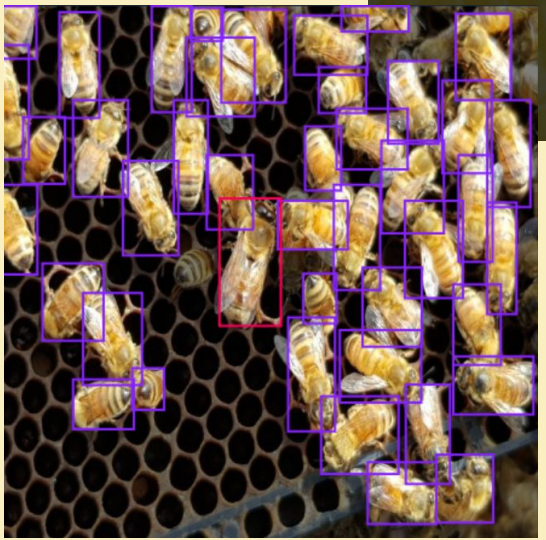


♀ ♂

Initial dataset from Roboflow

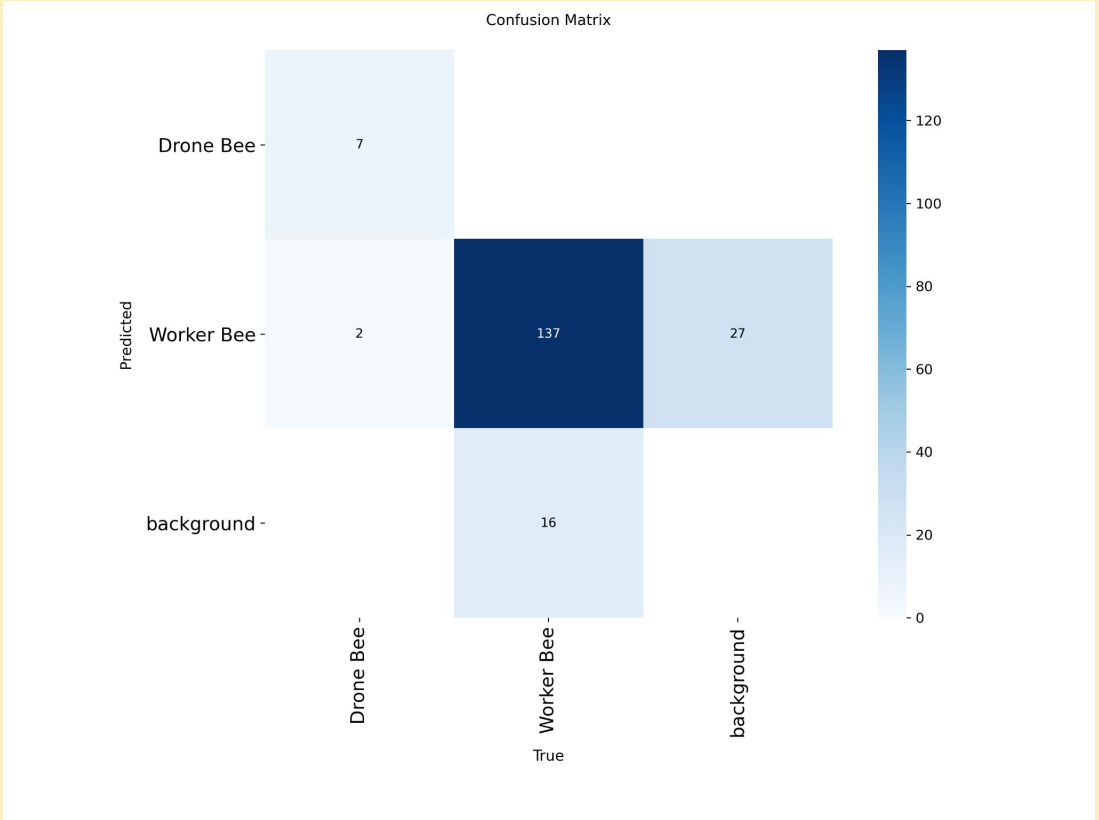
- **Augmentations**
- **Train/Val/Test:** 40/5/5
- **Drone : Worker Ratio –** 34.3 : 1





Initial Result

- Trained 4 models
- YOLOv10



Dataset from Mississippi State University

Image preprocessing:

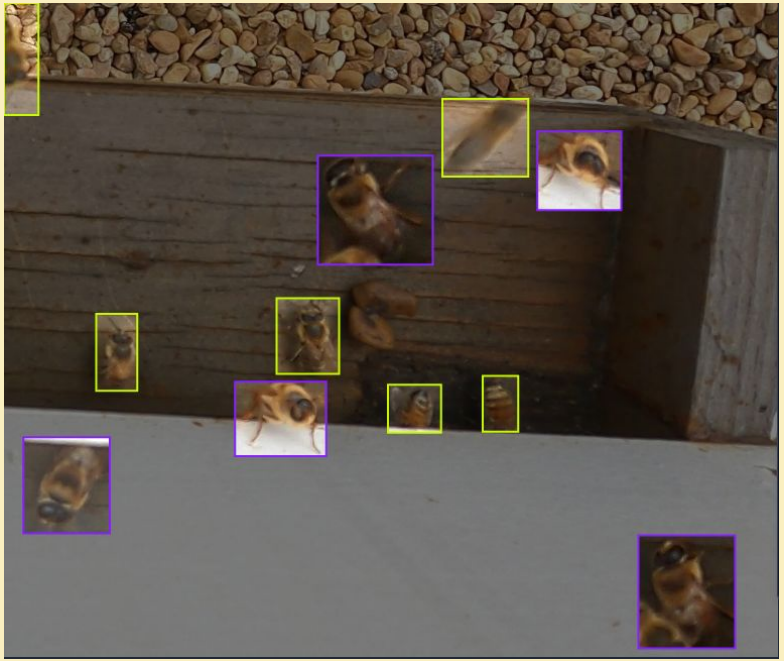
- Removed frames containing only worker bees
- Frames split into **3 overlapping tiles: 3840 × 1080 → 3 × 1280 × 1080**, with **128 px overlap**
- Initial class ratio: **Drone : Worker ≈ 1 : 14**

Manual review: Dataset inspected and labeling errors corrected

Augmentation: Copy–paste augmentation applied to drone instances

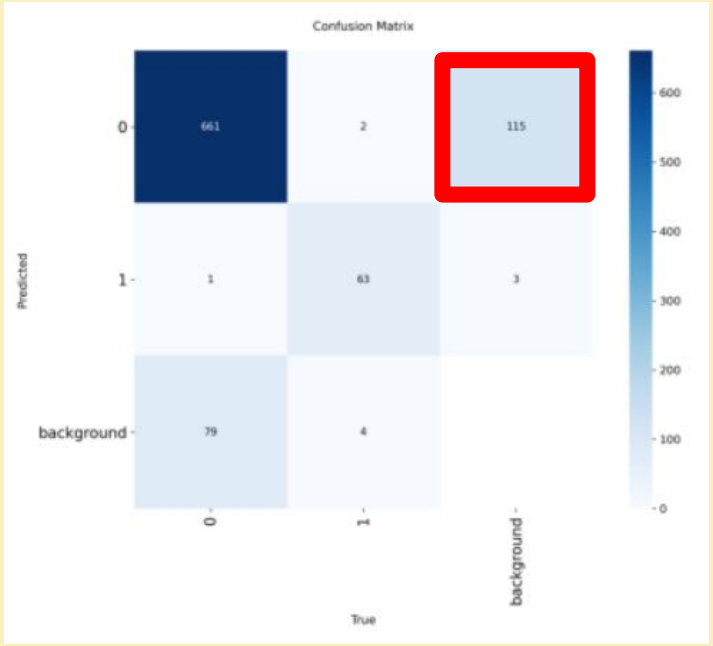
Dataset split: Train / Val / Test = **820 / 100 / 100**

Final class balance: 🐝 **Drone : Worker = 1 : 2**

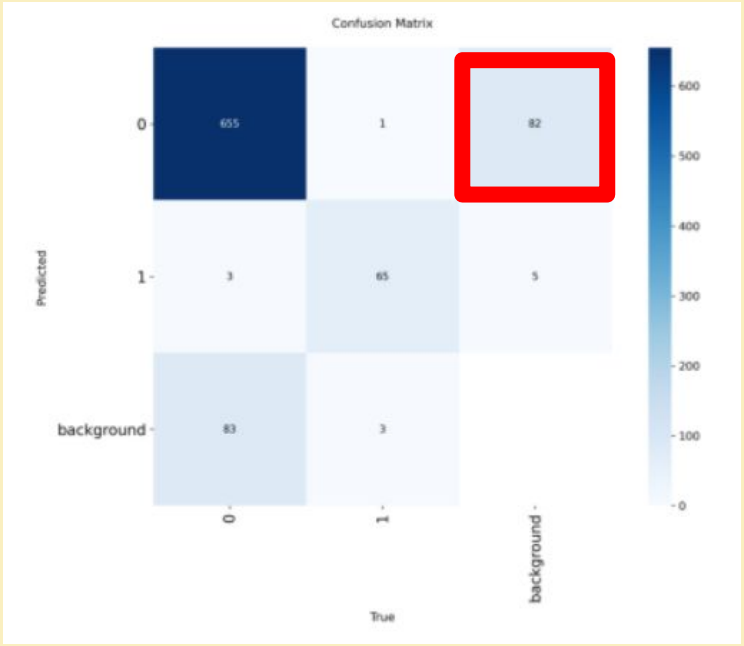


Training

Automatic image splitting



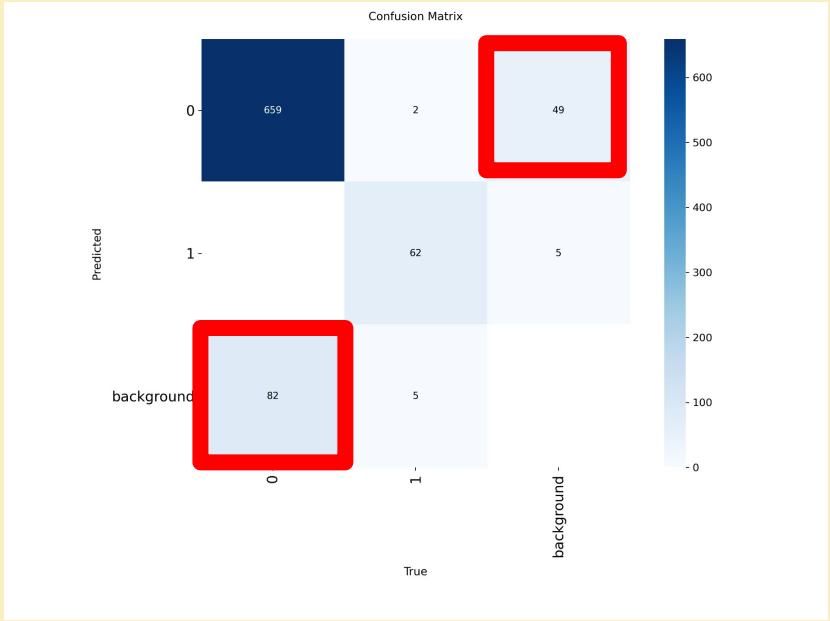
Imbalanced dataset



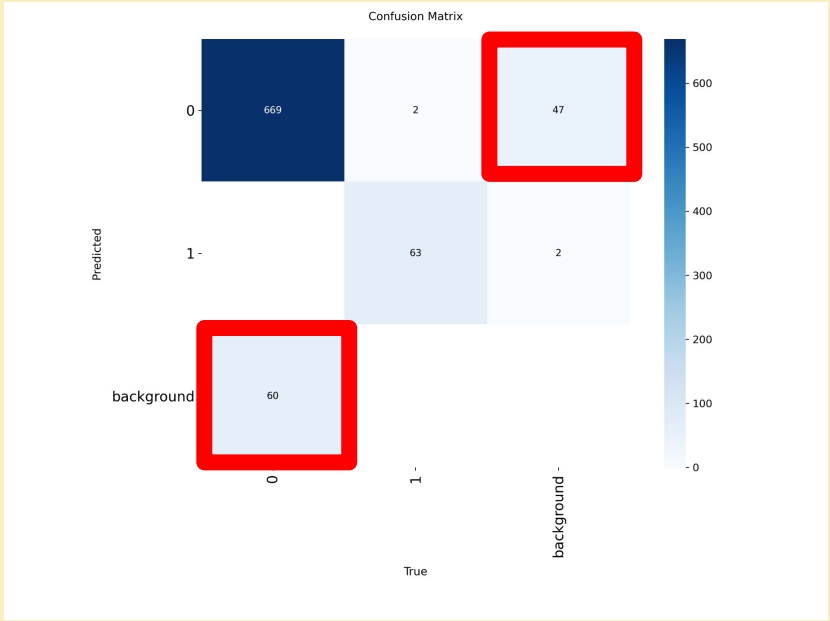
Ratio = 2:1

Training

After manually reviewed



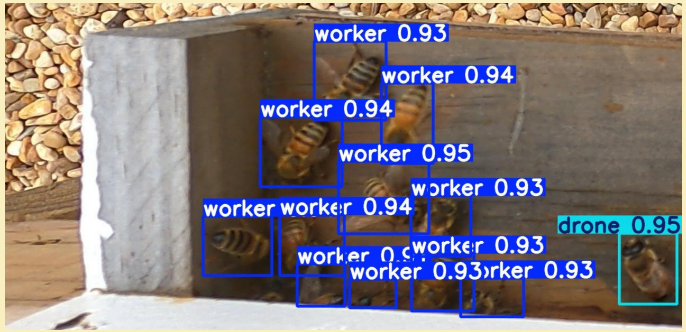
Imbalanced dataset

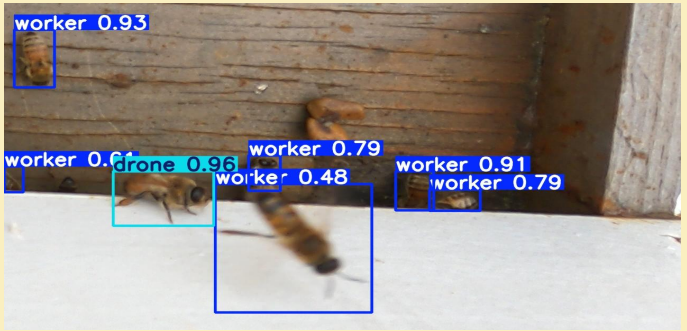


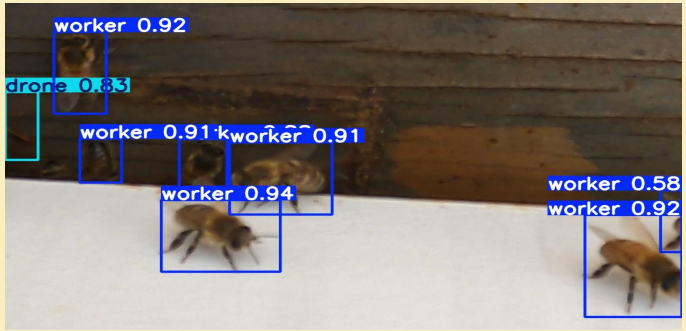
Ratio = 2:1

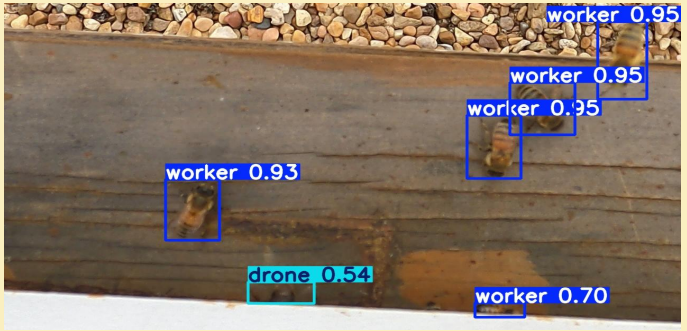
Metric	Imbalanced dataset	Balanced (2:1) dataset
Precision	0.9417	**0.9668**
Recall	0.9078	**0.9332**
mAP@50	0.9526	**0.9662**
mAP@50–95	0.6331	**0.6641**
Train cls loss	0.5425	**0.2367**
Train box loss	1.3727	**0.5520**

Results on test set









What we learned

- Dataset and its quality is important
- Augmentation is useful
- Training takes time

Predictions on Gratheon dataset

Team

- Kreete Kuusk: Trained YOLOv8 for model selection. Manually looked over and edited the final dataset. Communication with project partner and tutor. Prepared slides for presentation.
- Danni Zhang: Trained YOLOv10 for model selection. Preprocessed and augmented data: classes combining, image tiling, extract drone crops, augment them, paste into training images to rebalance the class ratio and visualized it. Wrote README. Prepared slides for presentation. Report and communicate with project partner and tutor.
- Jasper Luik: Searched for datasets on Roboflow. Trained RT-DETR for model selection. Trained initial YOLOv10 models on preprocessed unbalanced and balanced data. Added a README to the repository.
- Rasmus Mirma: Research. Trained YOLOv12 for model selection. Searched for datasets and prepared the second dataset. Trained numerous models on UTHPC with different datasets. Visualized the resulting model on images and videos. Prepared slides for presentation.

Used sources

- *https://bestbees.com/bee-hierarchy/#workerbee*
- *https://gratheon.com/about/products/entrance_observer/*
- *https://www.britannica.com/animal/drone-bee*
- *https://universe.roboflow.com/*
- *https://scholarsjunction.msstate.edu/gri-publications/4/*
- *https://drive.google.com/drive/folders/105PmxDKFUR6NCPLHBkXGdkfcZwWf9ABI*

Github Repository:
https://github.com/KreeteKuusk/Bee-type-detection-ML2025

Thanks!

Any questions?