Excel@FIT 2024

# Vision-Guided Target Following: Advancing On-Board Tracking

BRNO UNIVERSITY OF TECHNOLOGY — FACULTY OF INFORMATION TECHNOLOGY


Samuel Ján Mucha
2023/24
Supervisor: doc. Ing. Vítězslav Beran Ph.D.


## Objective

- **Camera-based detection and tracking for drones.**
- **System reliability in dynamic environments.**
- **Real-time performance (<35ms).**
- **Detect and track other drones onboard.**
- **High accuracy and efficiency.**

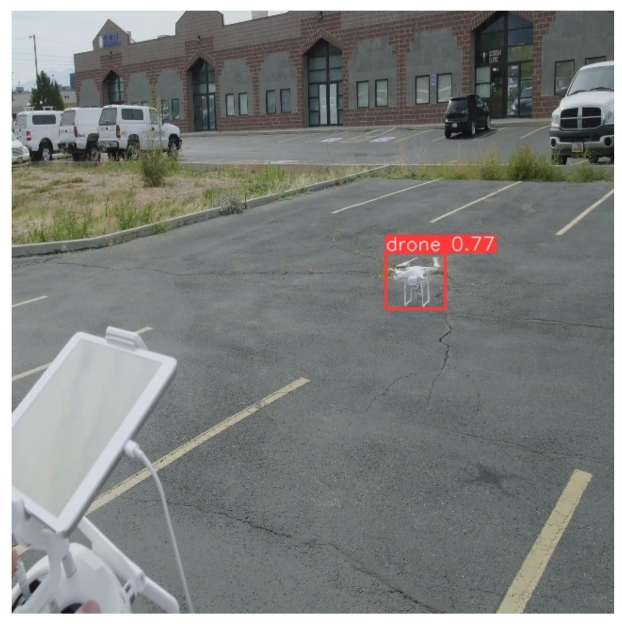
## Existing work

- **GPS limitations**: Inaccurate in urban areas, misses crucial data captured by onboard cameras.
- **Missed opportunities**: GPS reliance overlooks vital information, especially in GPS-interfered regions.
- **Urban challenges**: GPS inadequacies accentuated in urban settings, underscoring the need for onboard camera solutions.

## Dataset

- **Extensive dataset**: Over 10,000 drone images sourced from diverse online and video sources.
- **Varied content**: Captured drones at various distances, from different vendors, ensuring comprehensive coverage.
- **Manual annotation**: More than half of the dataset manually annotated for high-quality labeled data.
- **Largest on Kaggle**: Recognized as one of the largest drone dataset of drones on Kaggle, offering unmatched scale and diversity.
- **Quality assurance**: Rigorous quality control ensures dataset integrity and reliability.
- **Versatile applications**: Suitable for various computer vision tasks beyond drone detection.

Nature Background



Urban Background

Drone Details

Hard To Find Drone

## Detection & Tracking

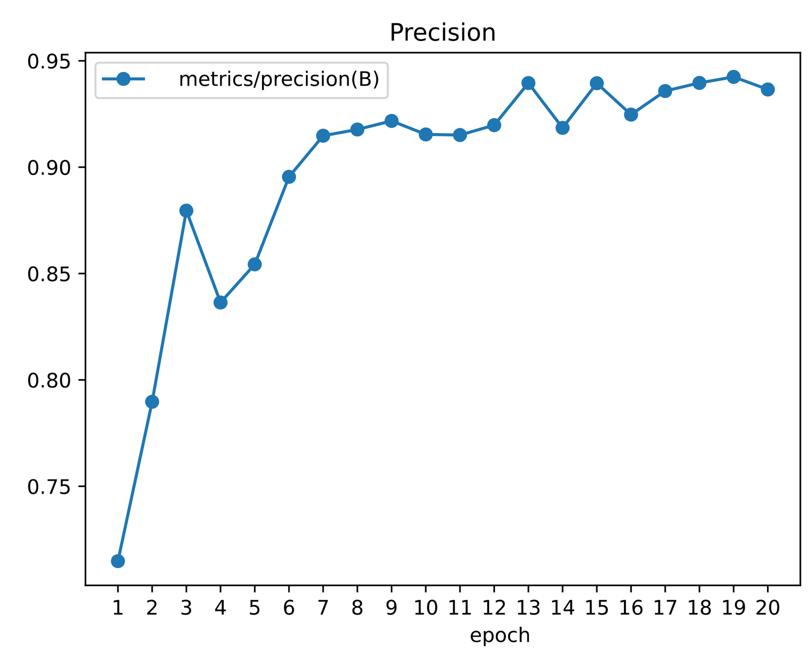
- **Detector**: YOLOv8
  - Training details: Trained for 20 epochs using a dataset split of 70% for training, 15% for testing, and 15% for validation.
  - Performance metrics: Achieved 94% precision and 92% recall.
  - Optimization attempts: Despite efforts to minimize model size through pruning, quantization, and backbone modifications, real-time inference speeds on edge hardware were not attained.

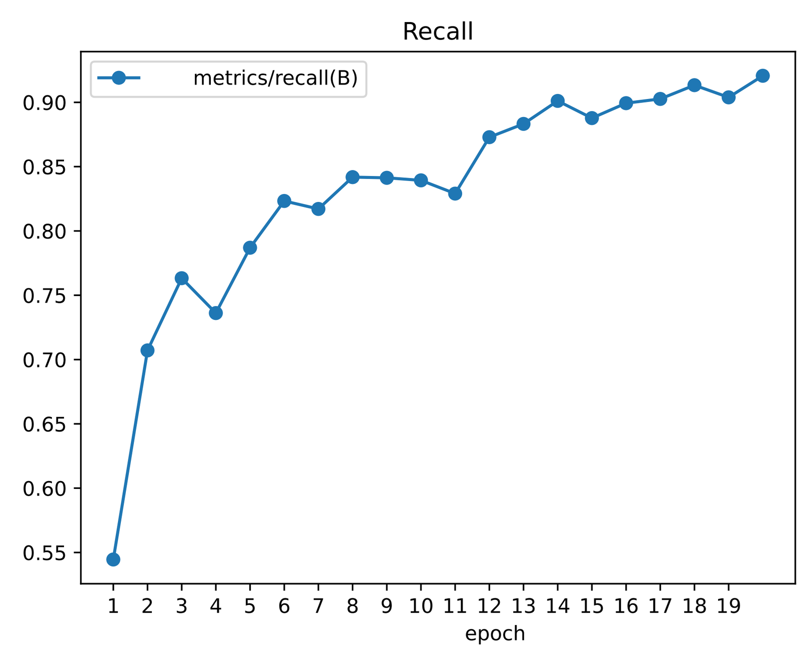
YOLOv8

- **Tracker**: MOSSE, KCF, TLD, CSRT, VIT
  - Purpose: Implemented to compensate for onboard detection limitations and maintain tracking during slow detection speeds.
  - Selection criteria: Different trackers were tested based on their ability to fill detection gaps and their inference speed compared to the detector.
  - Evaluation metrics (tested on a subset of TrackingNet):
    - TLFF (Tracking Length to First Failure): Measures the number of frames until the first failure in tracking.
    - RL (Recovery Length): Indicates the number of frames required for the tracker to recover after a failure.
    - RF (Recovery Failure Percentage): Percentage of failed recovery attempts.
    - UNR (Unnecessary Full Sequence Tracking Percentage): Percentage of tracking sequences that were fully tracked without failure.

| [ms/per frame] | Raspberry Pi 4 | Nvidia Jetson Nano | Intel Core i7-9750HQ | NVIDIA GTX 1660ti |
|---|---|---|---|---|
| PyTorch | 1244 | 316 | 51 | 7 |
| ONNX | 600 | X | 47 | X |

YOLOv8 Detection Speeds On Various Platforms and Inference Engines



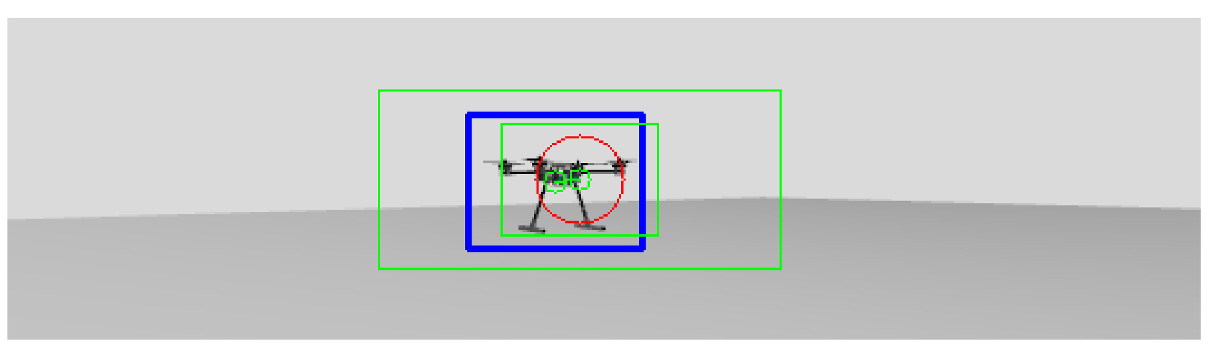




| | TLFF | RL | RF | UNR | Speed [ms/per frame] |
|---|---|---|---|---|---|
| MOSSE | 127.12 (26.67%) | 32.71 (6.85%) | 46% | 14% | 0.183 |
| KCF | 158.01 (33.34%) | 31.66 (6.57%) | 39% | 16% | 1.186 |
| TLD | 41.67 (8.66%) | 27.08 (5.81%) | 14% | 2% | 21.612 |
| CSRT | 188.72 (39.69%) | 21.05 (5.68%) | 18% | 19% | 21.482 |
| VIT | 228.69 (48%) | 17.75 (4%) | 6% | 28% | 9.230 |

Evaluation of Trackers on TrackingNet; mean num. of frames is 478

## Testing in Simulation

- **Testing Environment**: Simulator (Gazebo) with ROS2 and PX4 drone stack integration.
- **Surveillance Implementation**: Utilized a simple box minimization algorithm for proof of concept.
- **Control Algorithm**: Drone bounding box utilized from detection is used for movement in YZ plane; bounding box size employed for X-axis movement.



Following Drone Perspective

PX4 autopilot

ROS