Camera-Based Object Detection and Tracking for a Decentralized Sensor System

Student

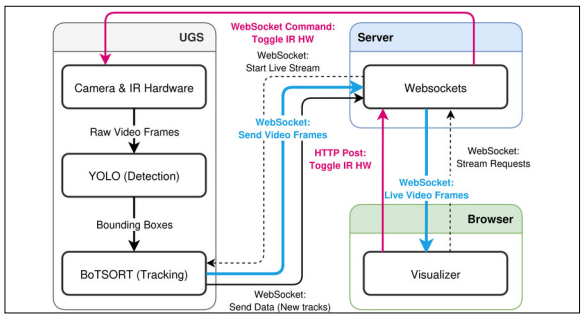
Mike Peng

Initial Situation: RUAG is developing a decentralized sensor system to detect and track persons, vehicles, and drones day and night. The previous system relied on motion detection, which was highly prone to errors and lacked classification. A deep learning solution optimized for edge hardware is needed to meet these requirements. The goal is to reach a precision and recall of 90% while simultaneously ensuring real-time object tracking and visual surveillance via a web browser live stream.

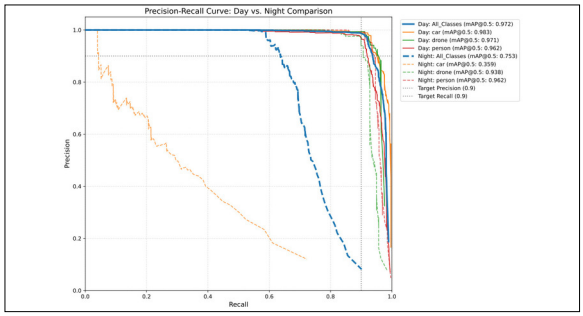
Approach / Technology: The technical realization focuses on training a lightweight object detection model, implementing a multi-object tracker, and developing a web-based live stream architecture. For nighttime operation, the camera module features IR LEDs and an IR filter, which can be remotely controlled via the web visualizer. A YOLO model was trained using public datasets covering persons, vehicles, and drones. To specifically enhance drone detection accuracy, a custom dataset featuring drone footage was created and integrated into the training pipeline. To ensure reliable tracking independent of unstable mobile network frame rates, the BoTSORT algorithm was deployed directly on the Raspberry Pi. Each sensor system processes the camera data locally and only transmits data to the server upon detecting a new track, or when a live stream is requested by the user.

Result: The system successfully achieves the targeted 0.9 precision and recall. However, clear distance limits had to be established. In daylight, drones are reliably detected within 8 meters, whereas the detection range for persons is 80 meters and for cars is 90 meters. These distances are heavily reduced at night, with a maximum detection range of 5 meters for persons and 3m for drones. Cars are poorly detected at night because the IR LEDs are too weak to sufficiently illuminate their contours. Despite these hardware limitations, the successfully integrated edge-tracking and web-streaming architecture, capable of running at 5 frames per second, demonstrates the feasibility and efficiency of this decentralized sensor concept.

System architecture overview
Own presentment



Precision-Recall plot, with distance limits. The poor performance of cars is clearly visible.
Own presentment



Web-based visualization, displaying the video stream with tracked objects and the system's position on the global map.
Own presentment

Advisors
Hannes Badertscher, Thomas Unterer

Subject Area
Artificial Intelligence

Project Partner
RUAG AG, Emmen, LU