Safe and Efficient Collaborative Automation Systems

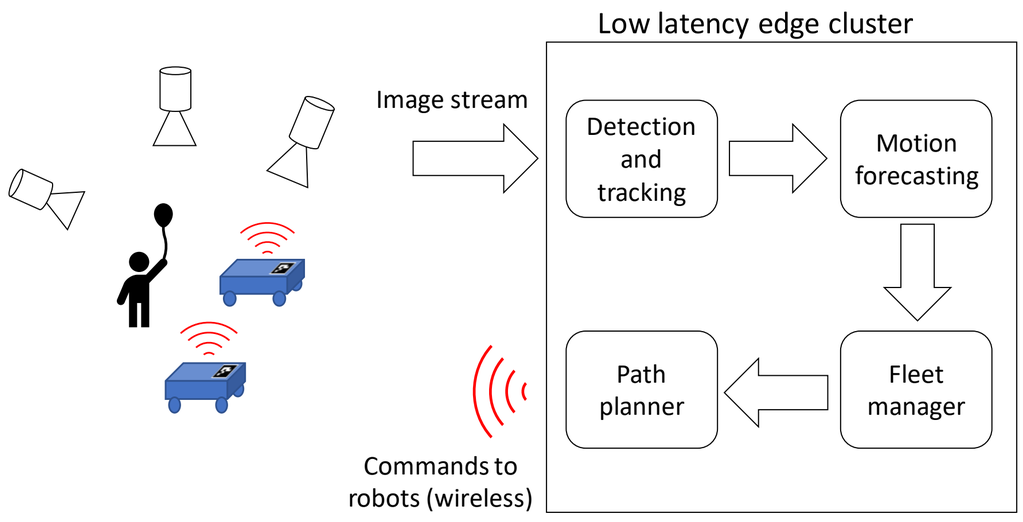
Erik Brorsson, Industrial PhD, Volvo Group Trucks Operations and Chalmers University of Technology
Dept. of Electrical Engineering, Automation Group
Supervisors: Kristofer Bengtsson (Volvo), Knut Åkesson (Chalmers) and Lennart Svensson (Chalmers)

Motivation & Research Goals

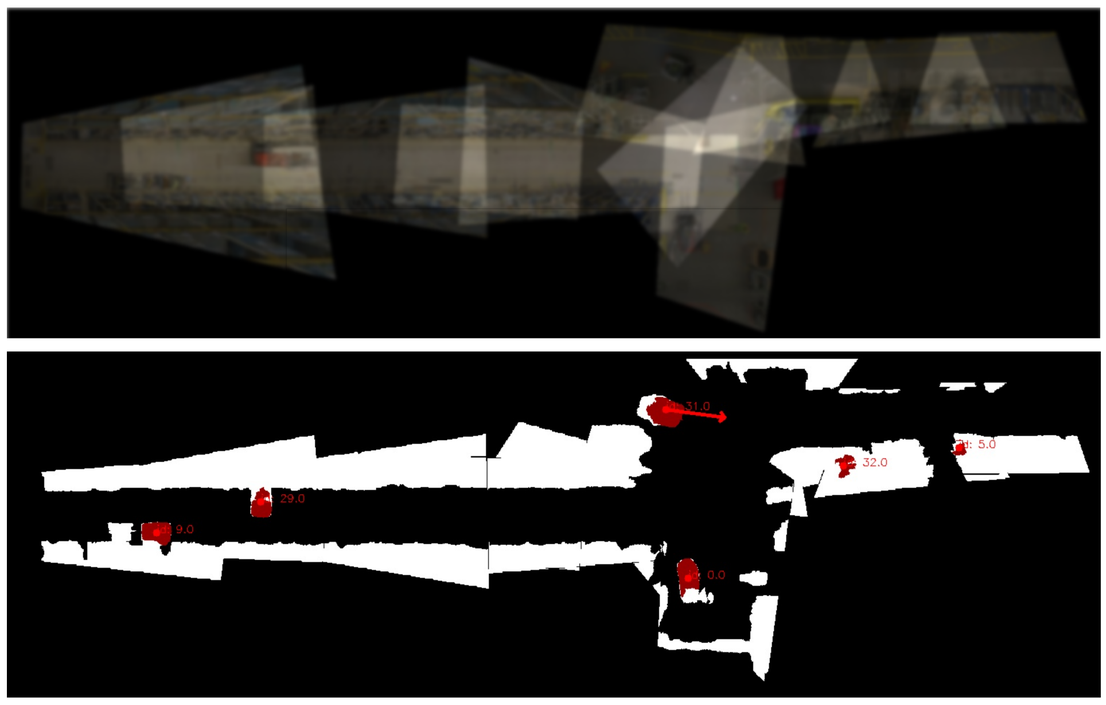
Volvo Group Trucks Operations (GTO) are developing a concept for internal logistics constituting a fleet of autonomous transport robots controlled over wireless communication by a centralized fleet manager. Ceiling-mounted cameras provide a bird's-eye-view of the factory, which enables the robots to navigate safely around pedestrians and forklifts. To ensure safe operation, uncertainties and anomalies in the perception system must be monitored. Additionally, reducing the cost of deployment by leveraging methods of semi-supervised learning during training is of interest. This project aims to realize a safe and practical perception system, where both simulated and unlabeled real-world data may be used during training without compromising the system's reliability.

Methods

In the internal logistics system that is under development at Volvo GTO, the state of the factory is monitored in a real-time database and is the basis for planning the routes of the autonomous transport robots. The system can roughly be divided into (i) camera-based perception, (ii) fleet manager and (iii) path planner.



In this research project, we are mainly concerned with modeling the factory's current state using ceiling mounted cameras with partly overlapping views. The state of the factory is described by an occupancy map and object trajectories in bird's-eye-view.



Research questions include:

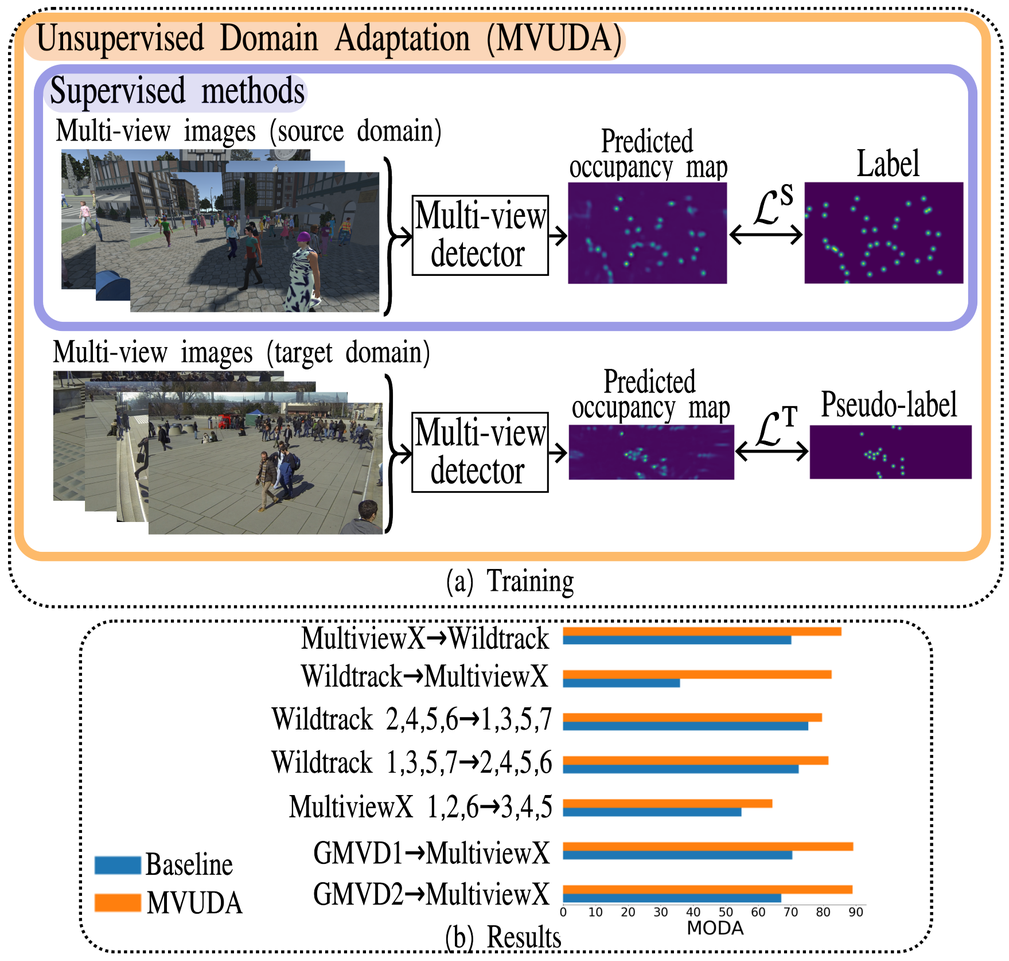
- How can information from different cameras be fused?
- How can the used model be trained efficiently (without much labaled data)?
- How can uncertainties in the perception system be quantified?

Selected Results

In [1], we address multi-view pedestrian detection in a setting where labeled data is collected using a multi-camera setup different from the one used for testing. While recent multi-view pedestrian detectors perform well on the training setup, their performance declines when applied to different camera setups. We propose to adapt the detector to new camera rigs without requiring additional labeled data using Mean Teacher self-training.

Our contributions can be summarized as follows:

1. We unveil the potential of self-training for multi-view pedestrian detection under a strict UDA setting and develop a state-of-the-art method for this problem.
2. We propose a simple yet effective post-processing method that improves pseudo-label reliability and thereby the effectiveness of self-training.
3. We demonstrate the efficacy of our method on multiple established benchmarks and on two new benchmarks, which we introduce to specifically address cross-camera rig adaptation.



References

[1] BRORSSON, Erik, et al. MVUDA: Unsupervised Domain Adaptation for Multi-view Pedestrian Detection. arXiv preprint arXiv:2412.04117, 2024.

WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM