Multi-sensor rail track detection in automatic train operations

Master’s thesis in Data Science

Student: Attila Kovacs

1st Advisor: Lukas Rohatsch (UAS Technikum Wien)

2nd Advisor: Daniele Capriotti (M2C Expert Control GmbH)

06.06.2024

FH University of Applied Sciences TECHNIKUM WIEN

Motivation

Automatic train operations (ATO)

- ATO systems use advanced technologies to perceive and interpret the railway environment to facilitate autonomous operations with minimal human intervention

Automatic rail track detection

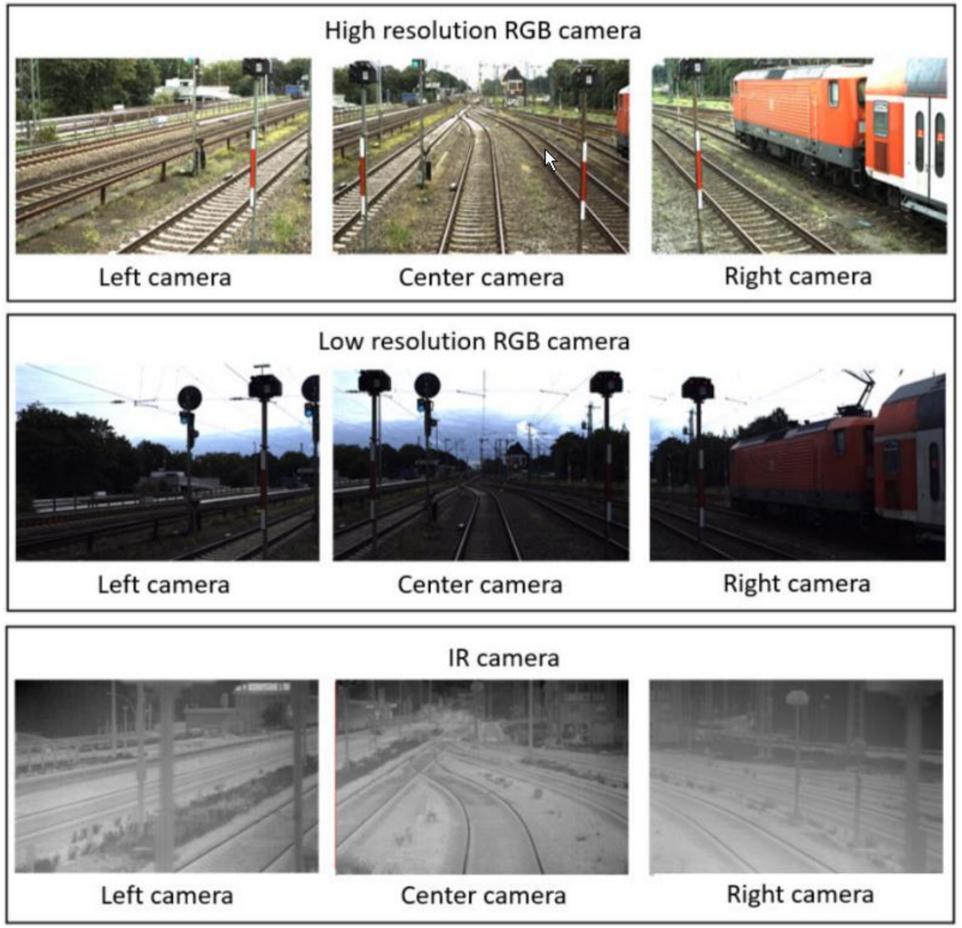
- Computer vision-based rail track detection is a crucial component for autonomous train navigation as it enables trains to understand and navigate complex rail networks

FH University of Applied Sciences TECHNIKUM WIEN

Motivation

Input: images generated by different sensors mounted on locomotive

Goal: identify and segment rail tracks



FH University of Applied Sciences TECHNIKUM WIEN

Dataset

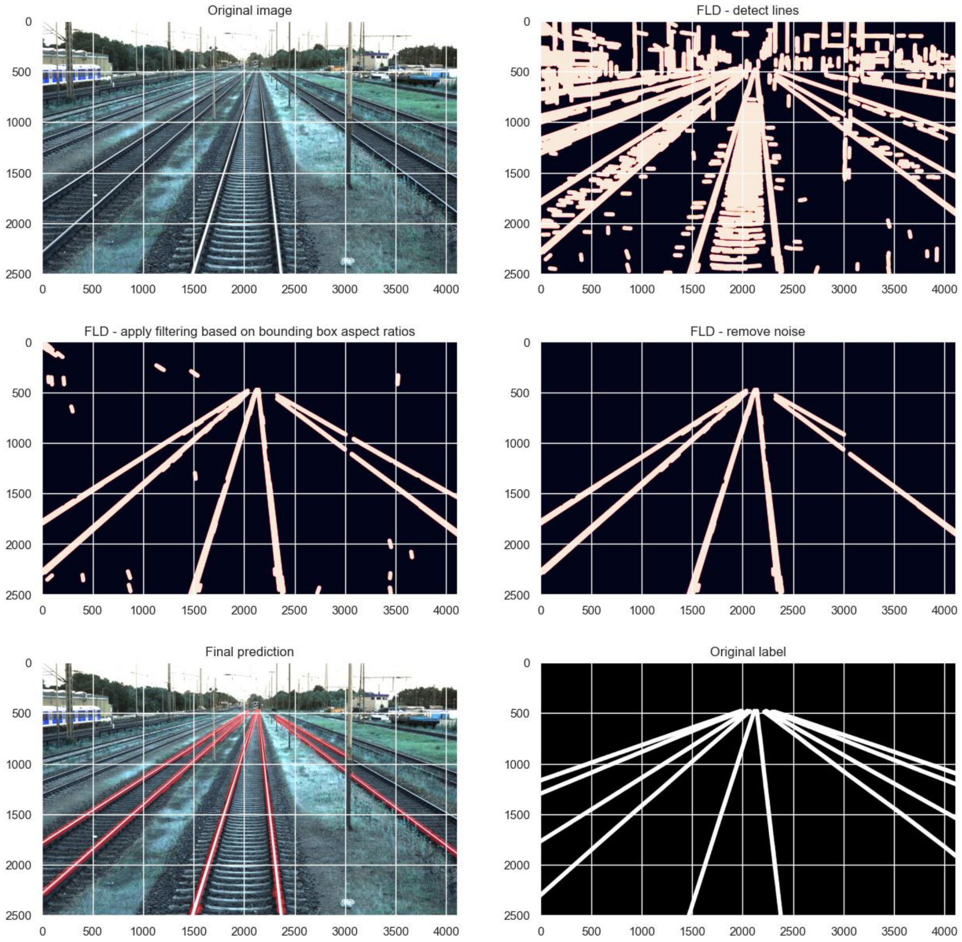
- OSDaR23 dataset (Digitale Schiene / Deutsche Bahn) → **Training and evaluation**
  - 7.421 frames from 45 video sequences
  - 27.386 labels
  - Low/high resolution RGB camera and infrared camera
- RailSem19 dataset (Austrian Institute of Technology) → **Training**
  - 8.500 images
  - 58.483 labels
  - Only RGB images
- Video stream (M2C / DB Cargo) → **Evaluation**
  - 1:14h video
  - Different scenarios such as tunel, double/single track, side walls etc.

FH University of Applied Sciences TECHNIKUM WIEN

Solution approach

- **Non-AI-based segmentation with fast line detection (FLD)**
  - Detect edges in image
  - Extract line segments
  - Grouping of line segments based on orientation and proximity
- **Deep-learning based approach based on YOLOv8**
  - Train model with pre-labeled images
  - Classify each pixel in an image according to its category (rail tracks vs. background)
  - Convert pixels into polylines
- **Evaluation criterion**
  - Dice score – best suited for unbalanced datasets, e.g., when the background is dominant

Non-AI-based segmentation with fast line detection (FLD)



Apply FLD

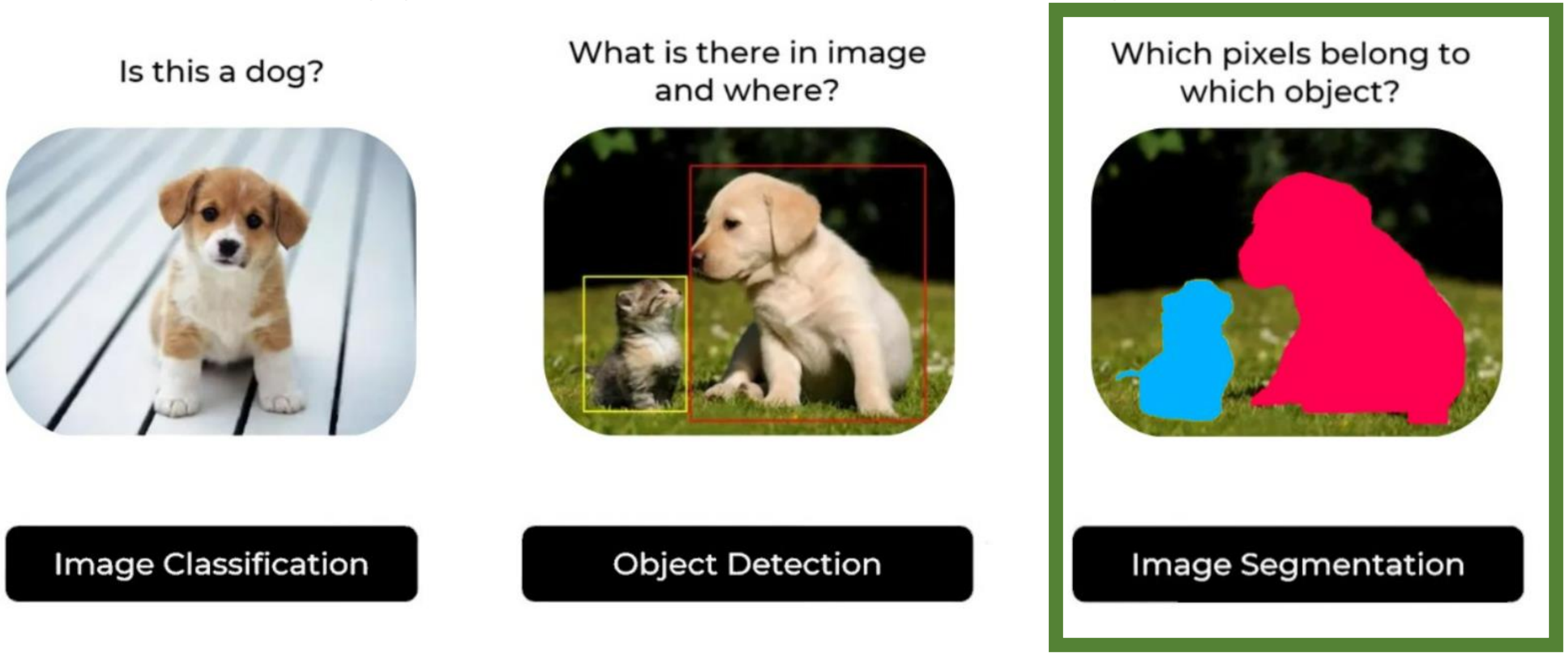
Filtering by removing „unusual“ lines

Removing noise based on clustering

Labels are given in the dataset

Deep-learning based approach based on YOLOv8

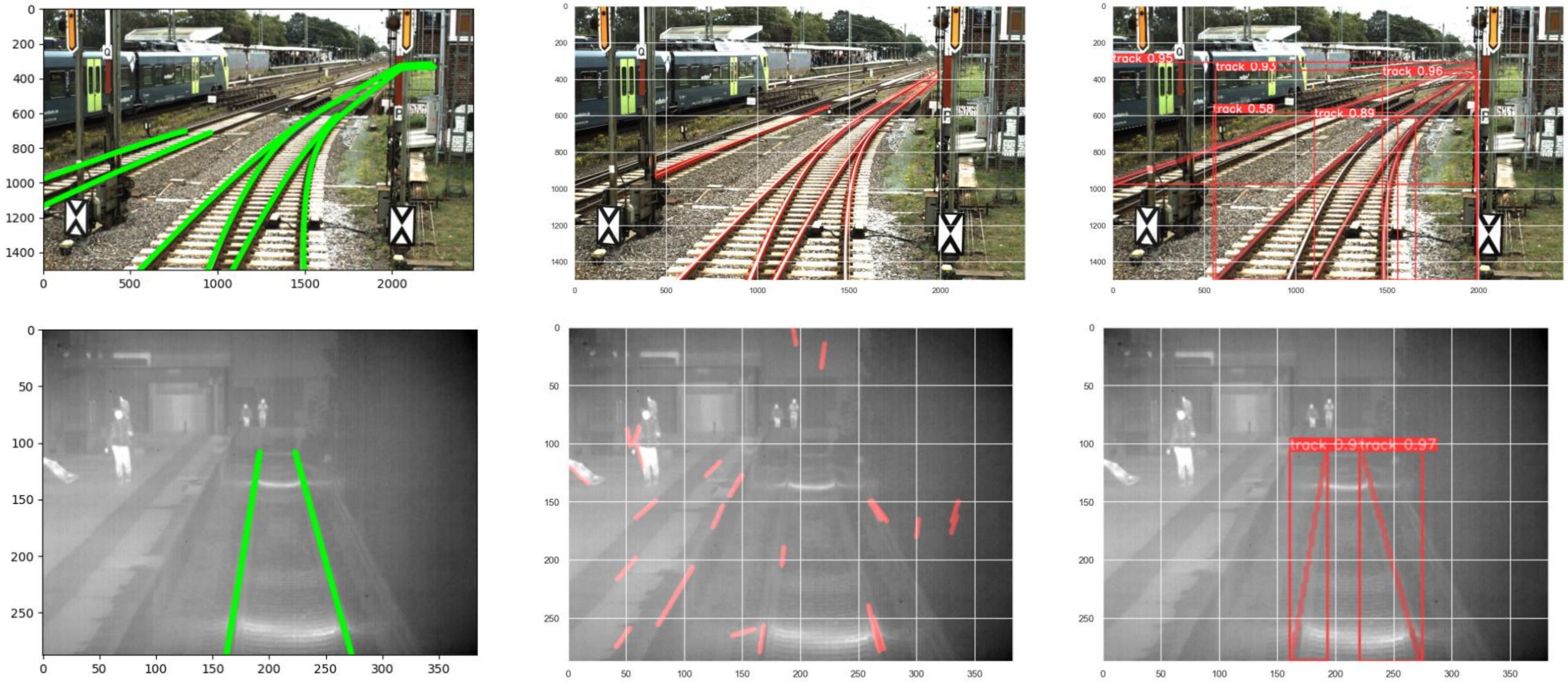
There are differenet applications in AI-based computer vision



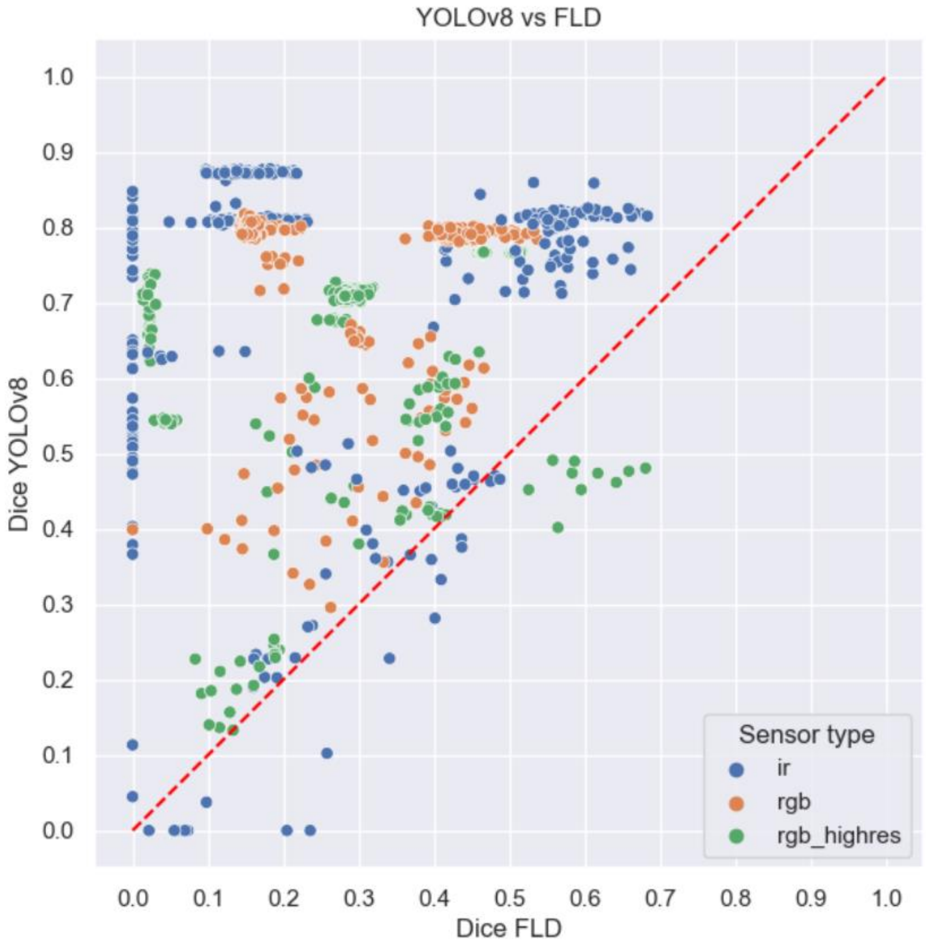
FH University of Applied Sciences TECHNIKUM WIEN

Results – Visual inspection

Ground truth | FLD | YOLO

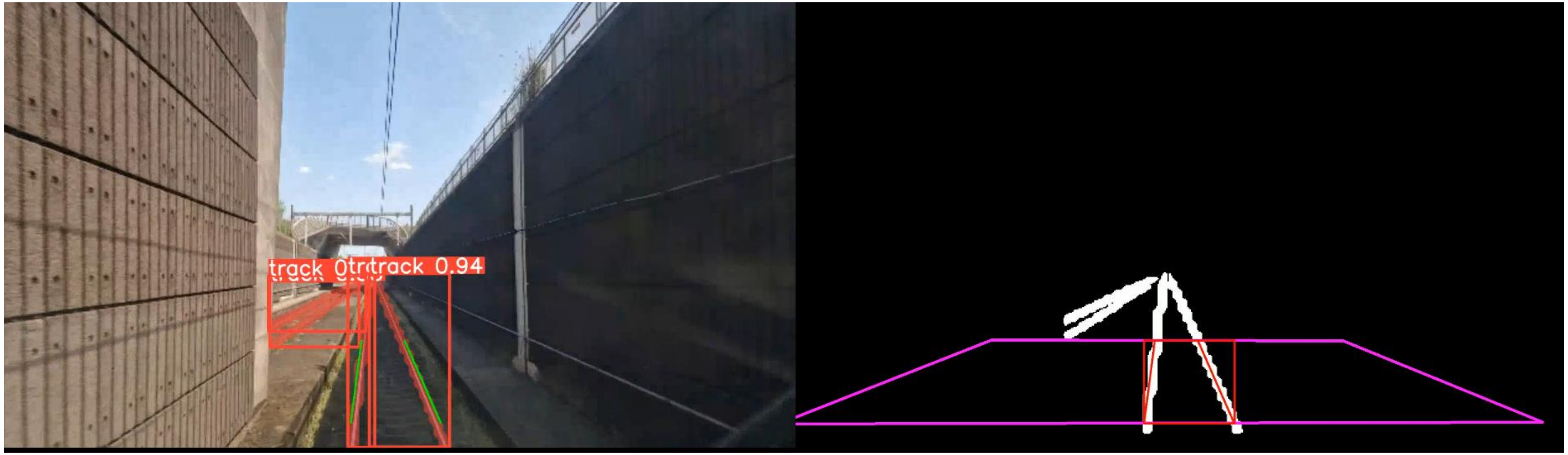


Results – FLD vs. YOLOv8 on OSDaR23



FH University of Applied Sciences TECHNIKUM WIEN

Results – Video stream: incorportating domain knowledge



University of Applied Sciences TECHNIKUM WIEN

Conclusion

- First project to investigate different sensor types in rail track detection
- Devised traditional base-line approach (FLD) and AI-based approach (YOLO)
- YOLO outperforms FLD in almost all test images based on Dice-score
- YOLO seems to provide very good results on infrared images
- **Best performace is achieved if AI-based track detection is enhanced by domain knowledge**

FH University of Applied Sciences TECHNIKUM WIEN

Thanks!

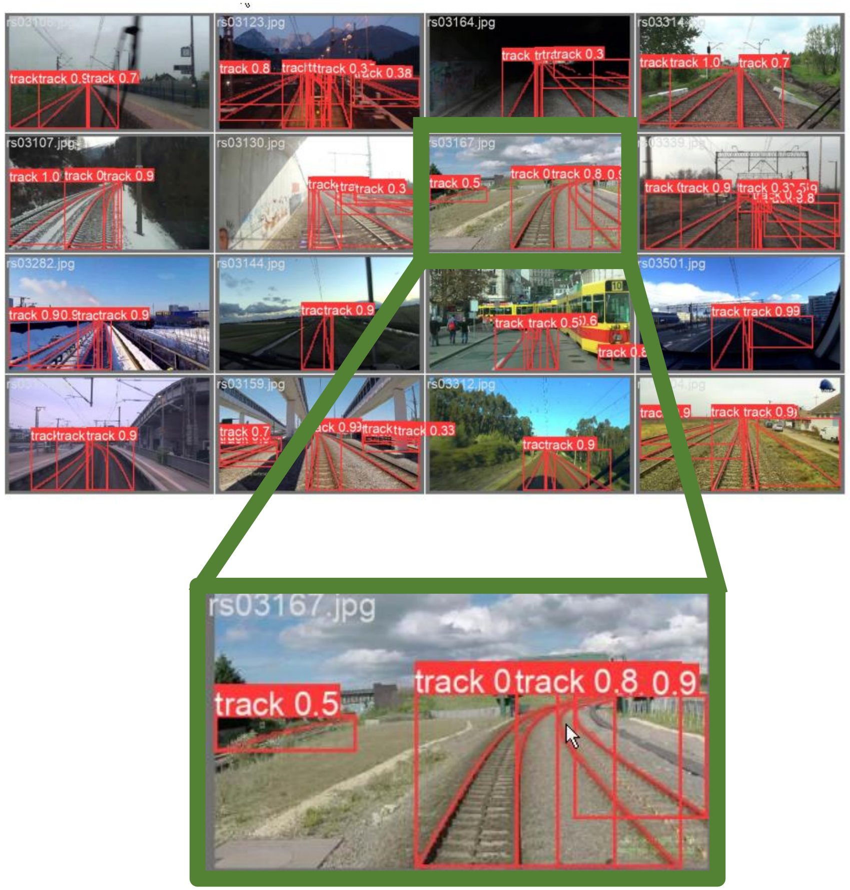
Challenges

- Size of dataset becomes often a limiting factor (disk, RAM, CPU)
- Refactoring existing approaches for lane detection proved difficult
- Organization of experiments
- A lot of custom code
- Resource intensive training despite HPC
- Working with HPC (scheduling, data transfer, versioning)
- Model tuning can take up to one week

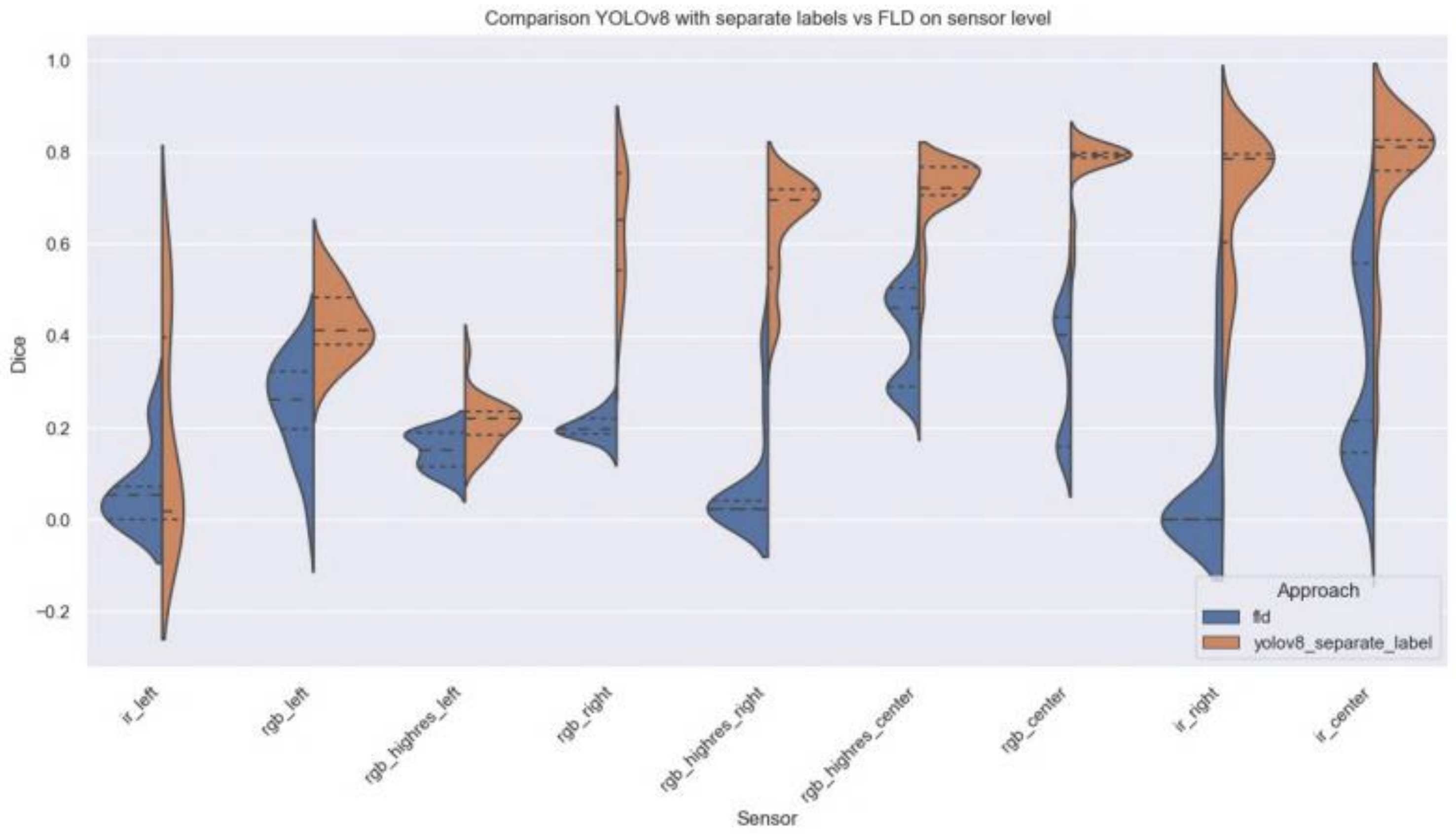
Deep-learning based approach based on YOLOv8

In each step of the training process, the model is provided with labeled images

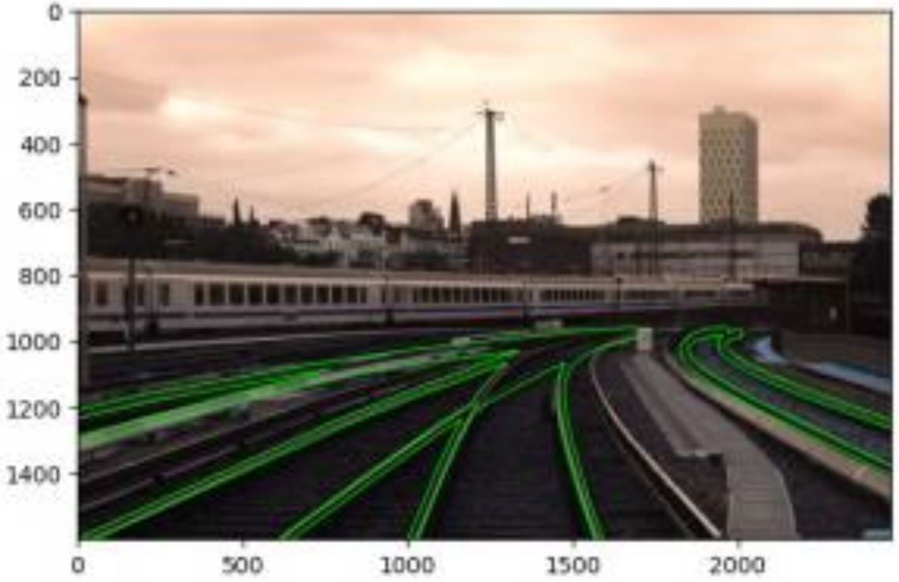
The model is evaluated on images from the validation set

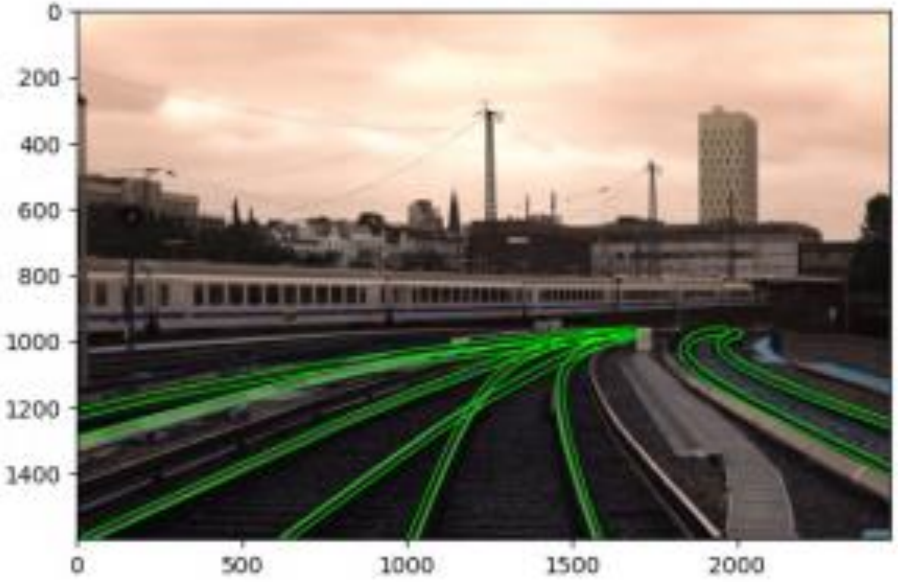


Results per sensor



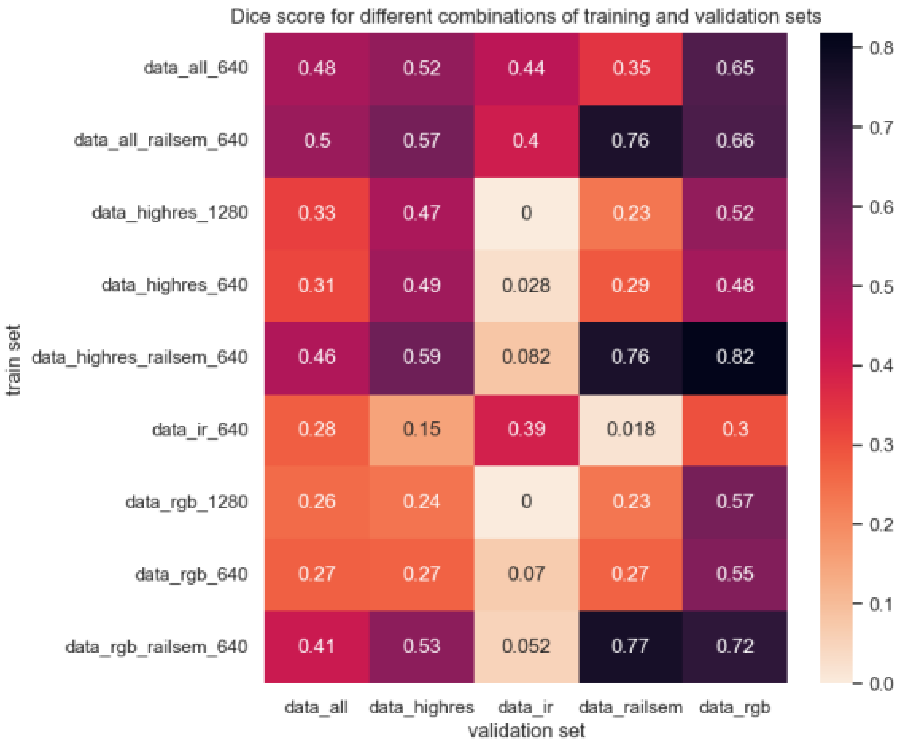
Labelling approaches



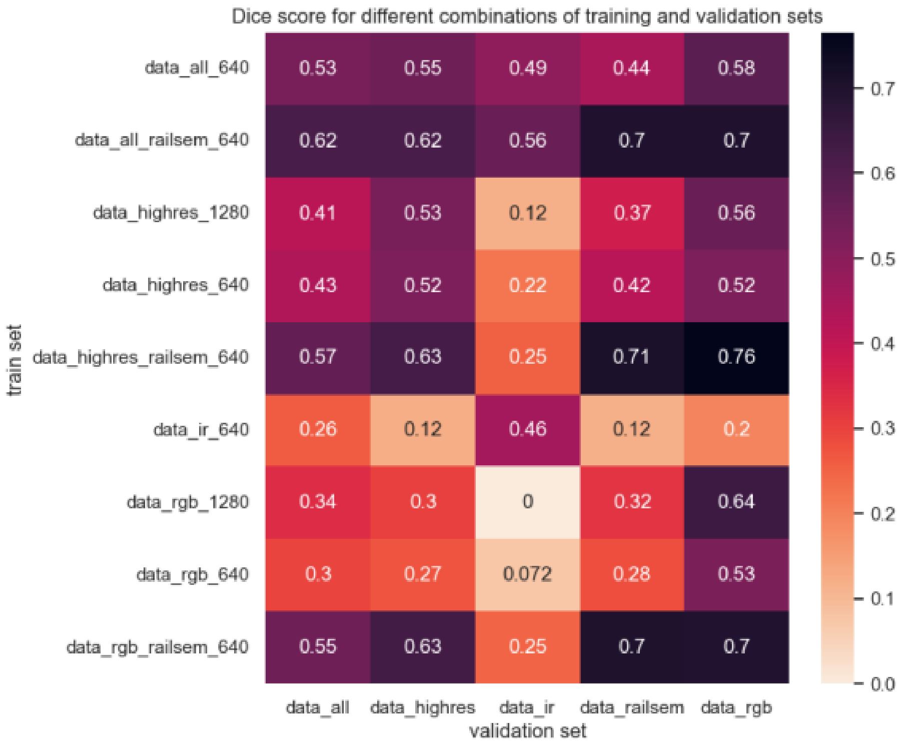


FH University of Applied Sciences TECHNIKUM WIEN

Tested YOLOv8 models



(a) Models with unified labels.



(b) Models with separate labels.

Frames of a video sequence