Multi-sensor rail track detection in automatic train operations

Master's thesis in Data Science

Student: Attila Kovacs

1st Advisor: Lukas Rohatsch (FH Technikum)

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

Alignment: 22.12.2023

FH University of Applied Sciences
TECHNIKUM WIEN

YOLO experiments

- Model backbone: yolov8n-seg
- Combined loss function consisting of
  - bounding box loss (error between the predicted and the ground truth boxes' geometry)
  - objectness loss (how confident the model is about the presence of an object in the bounding box)
  - segmentation loss (how close the predicted segmentation map is to the ground truth map)
- Paramters
  - epochs: 300
  - image size: 640, 1280 (where applicable)
  - batch size: auto selection based on image size and available memory
  - Yolo defaults
- Logging: Comet ML

Model statistics

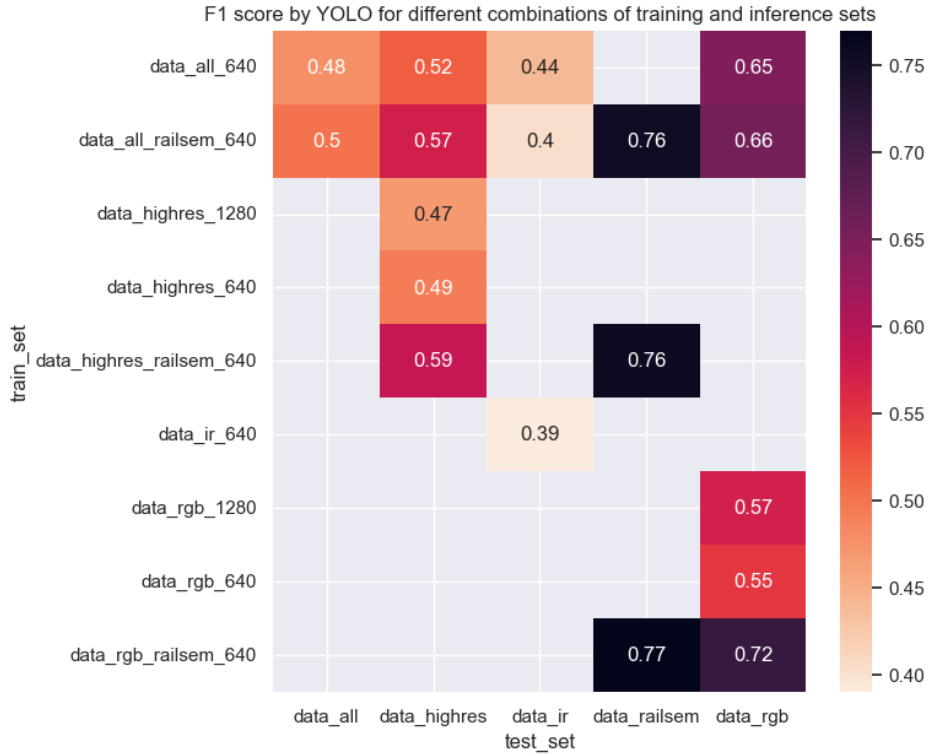
model	train set	epochs	image size	batch	wall clock time
yolov8n-seg.pt	rgb	121/300	640	23	00:37:57
yolov8n-seg.pt	ir	196/300	640	23	00:51:51
yolov8n-seg.pt	rgb	117/300	1280	5	01:39:58
yolov8n-seg.pt	rgb+railsem	138/300	640	23	02:44:37
yolov8n-seg.pt	highres	81/300	1280	5	04:20:46
yolov8n-seg.pt	highres+railsem	110/300	640	23	13:36:26
yolov8n-seg.pt	highres	130/300	640	23	13:53:44
yolov8n-seg.pt	all	190/300	640	23	23:27:11
yolov8n-seg.pt	all+railsem	224/300	640	23	23:42:53

Logging

FH University of Applied Sciences TECHNIKUM WIEN

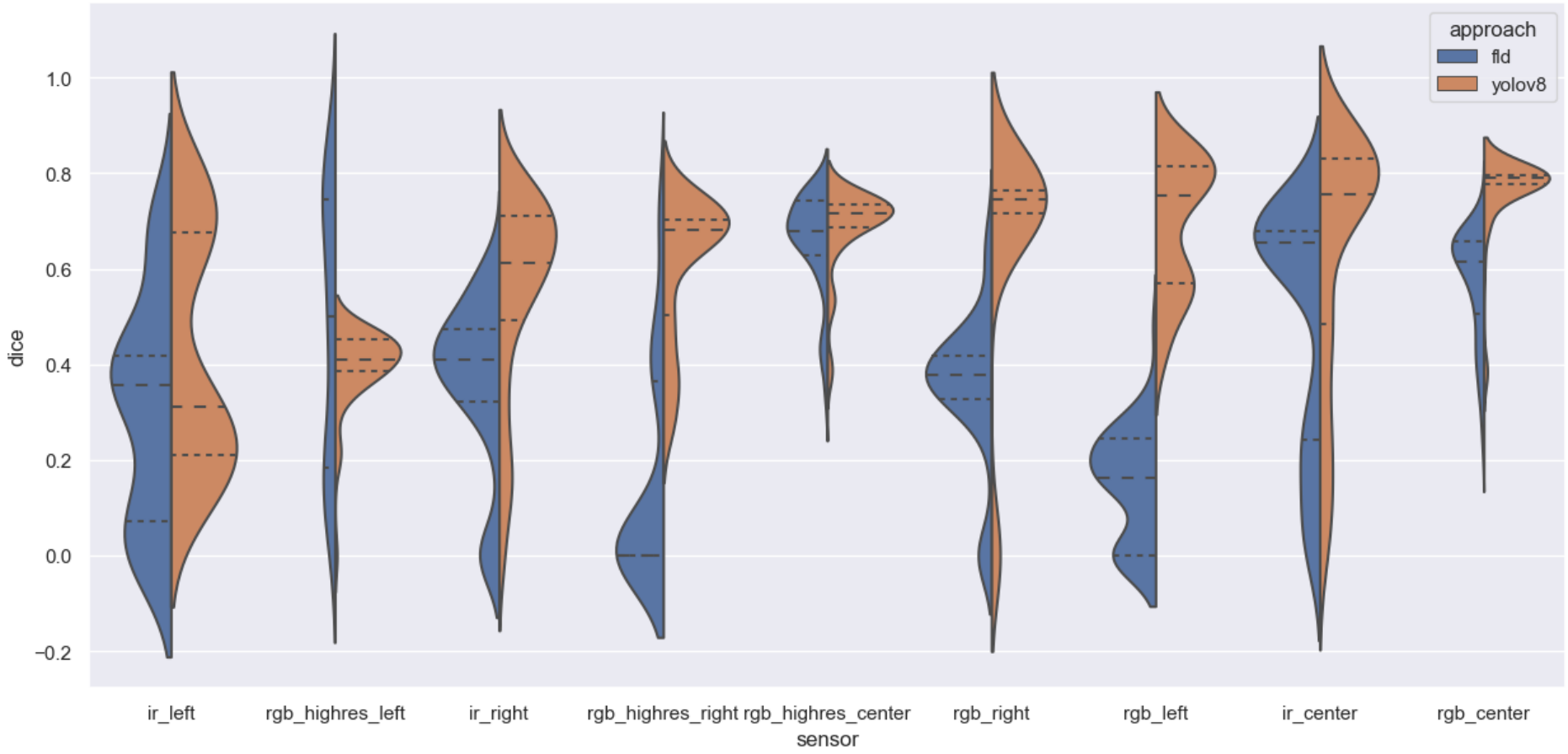
Test model on different sets

- Yolo was used to apply the different models to different validation sets
- Evaluation metric is the segmentation f1 score (= dice)
- For each data set we select the best model



FH University of Applied Sciences TECHNIKUM WIEN

YOLO vs Baseline



Video stream

- Video

Next steps

- Finetuning
- Finalizing the models
- Evaluation on test-set

Research questions

- Which modelling technique and which deep learning model can be applied to solve the rail detection problem?
- How can the efficiency of standard, high-resolution, and infrared cameras be compared against each other; does a higher resolution result in a higher accuracy?
- What is the trade-off between model accuracy and speed of providing predictions when applied to a video stream in real time?
- How do deep learning models perform compared to gradient-based thresholding approaches in terms of, e.g., accuracy (share of correctly identified objects) or F1-score (mean of precision and recall)