Multi-sensor rail track detection in automatic train operations

Master's thesis in Data Science

Student: Attila Kovacs

1st Advisor: Lukas Rohatsch (FH Technikum)

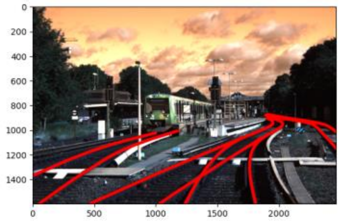
2nd Advisor: Daniele Capriotti (M2C Expert Control GmbH)

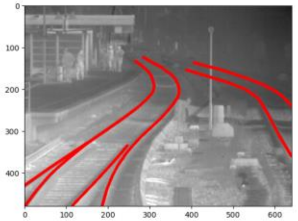
Alignment: 17.01.2024

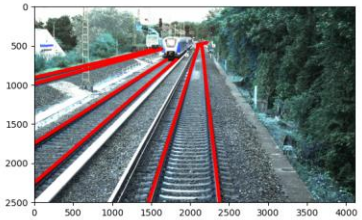
FH University of Applied Sciences TECHNIKUM WIEN

Problem setting

What is Automatic train operations (ATO)
Technology is used to automate
tasks that were previously performed
by rail personnel (e.g., conductor)



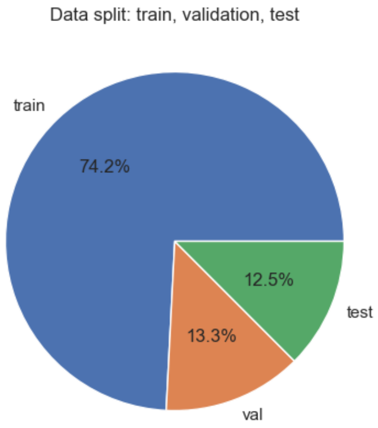




FH University of Applied Sciences TECHNIKUM WIEN

Dataset

- Deutschbahn:
  - Multisensor dataset
    - Infrared
    - RGB 5MP
    - RGB 12MP
    - Left, right, and center, respectively
  - Large similarity between frames of video → split on videos

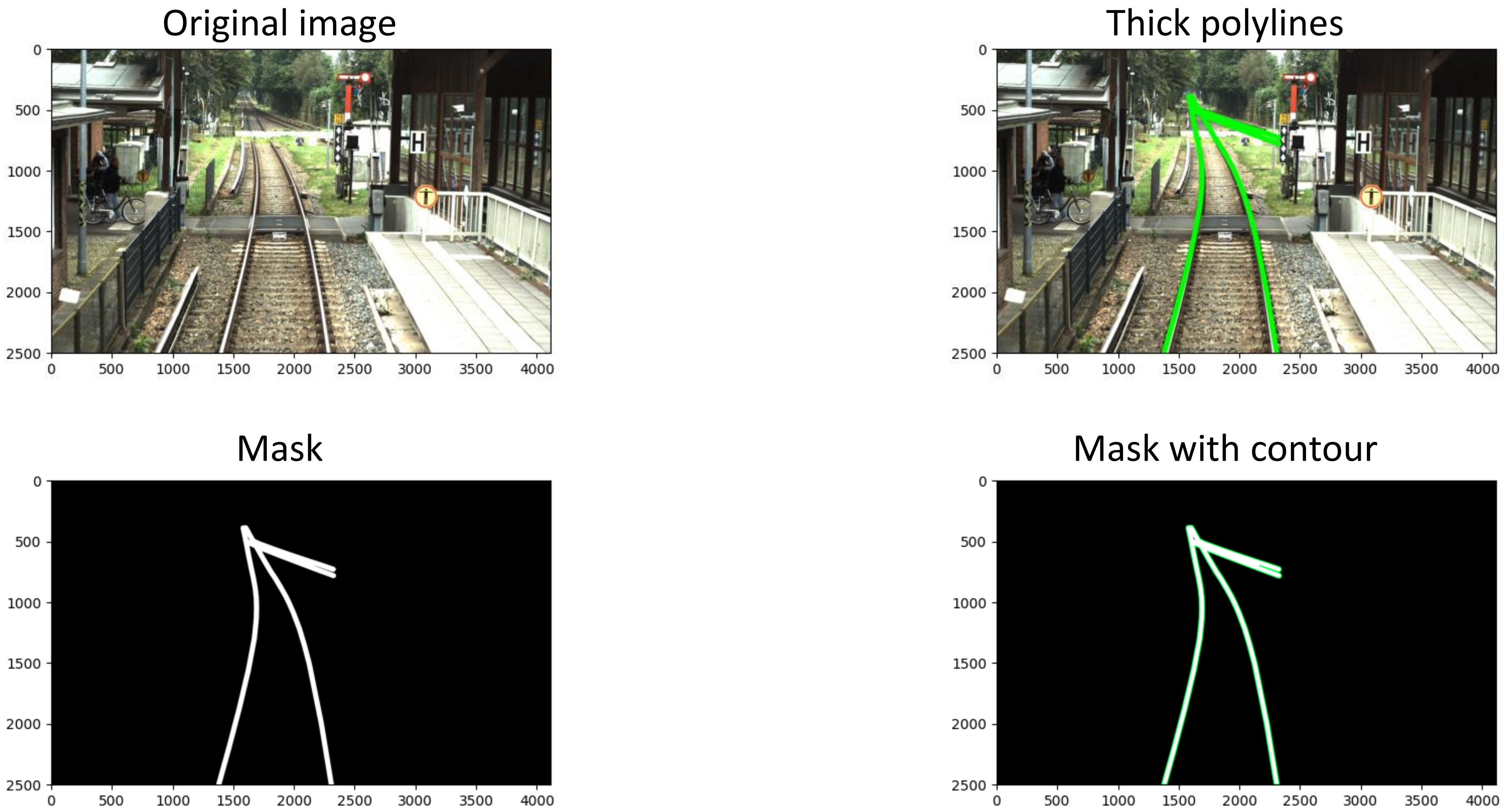


FH University of Applied Sciences TECHNIKUM WIEN

Modelling rail detection as segmentation

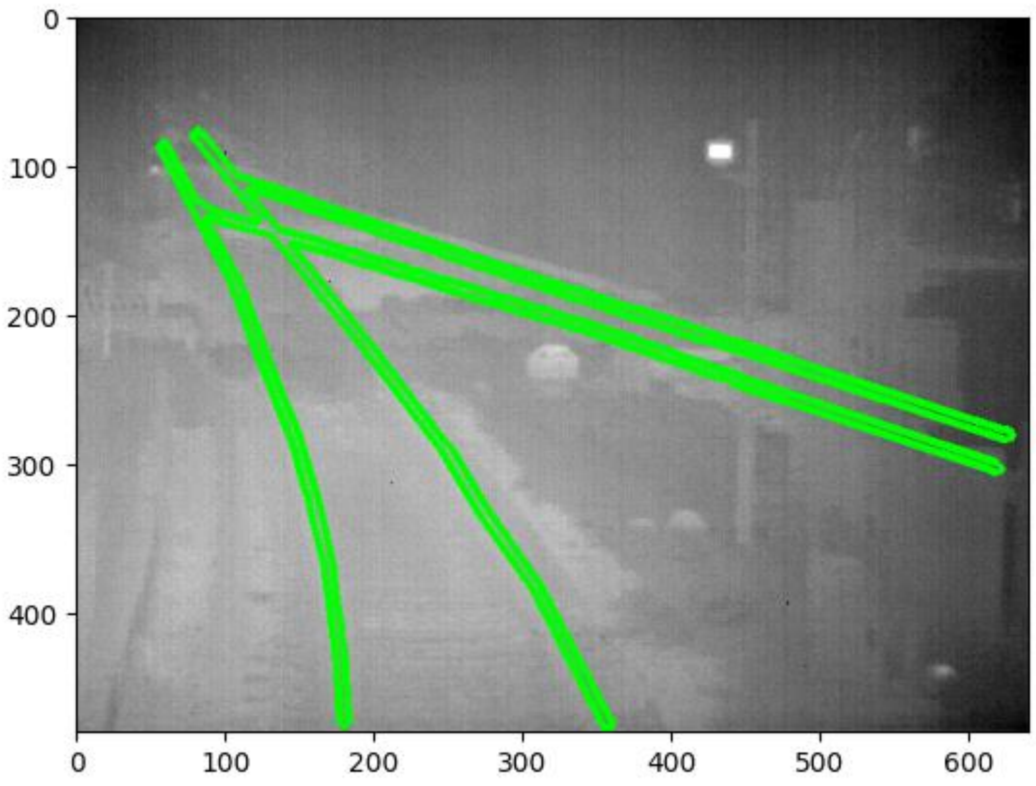
- Semantic segmentation → deep learning algorithm that associates a label or category with every pixel in an image
- Distinguish background pixels and track pixels
- In the data tracks are annotated by lines rather than pixel masks

Modelling rail detection as segmentation

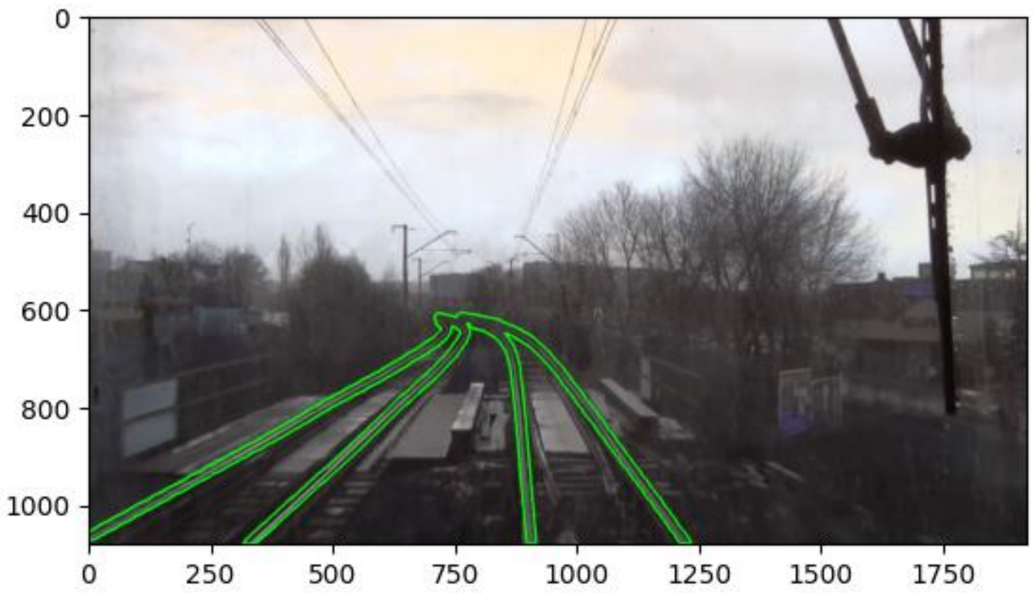


Modelling rail detection as segmentation

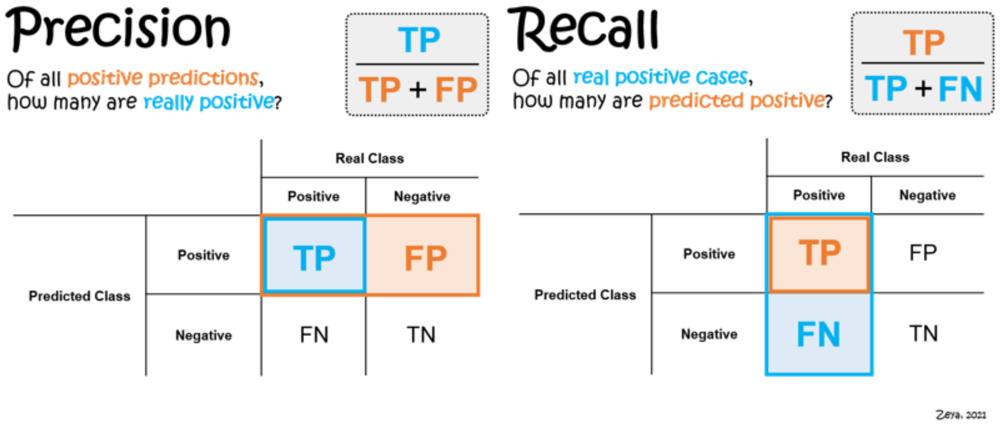
Example IR



Example RailSem



Performance metrics



$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

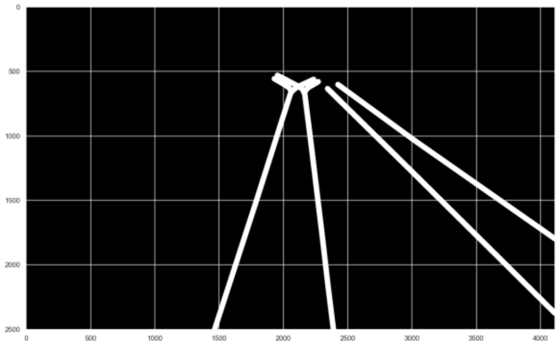
$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

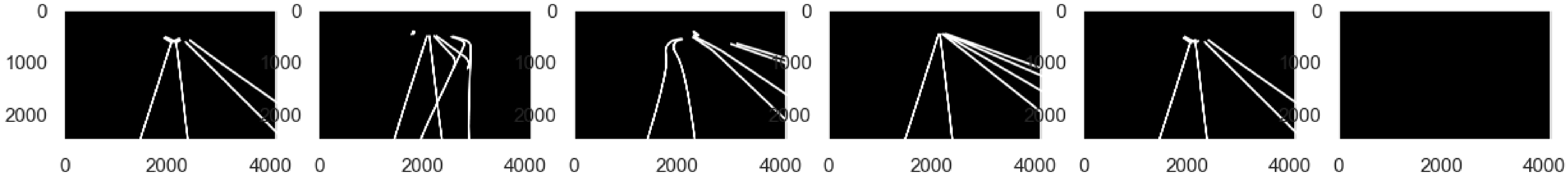
$$Dice = \frac{2TP}{2TP + FP + FN}$$

$$IoU = \frac{TP}{TP + FP + FN}$$

FH University of Applied Sciences TECHNIKUM WIEN

Performance metrics





pixel_acc: 0.9983
dice: 0.9775
precision: 0.9812
recall: 0.9738
iou: 0.9559

pixel_acc: 0.9499
dice: 0.3762
precision: 0.3557
recall: 0.3992
iou: 0.2317

pixel_acc: 0.9205
dice: 0.0567
precision: 0.0514
recall: 0.0633
iou: 0.0292

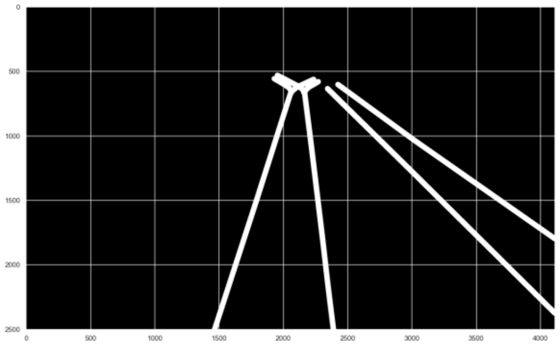
pixel_acc: 0.9435
dice: 0.3703
precision: 0.3201
recall: 0.4392
iou: 0.2272

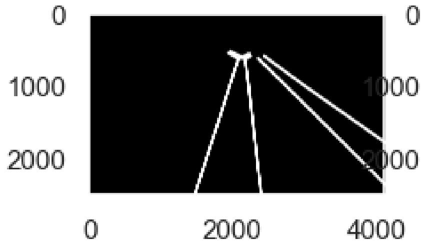
pixel_acc: 0.9979
dice: 0.9714
precision: 0.9787
recall: 0.9643
iou: 0.9444

pixel_acc: 0.9622
dice: 0.0000
precision: 1.0000
recall: 0.0000
iou: 0.0000

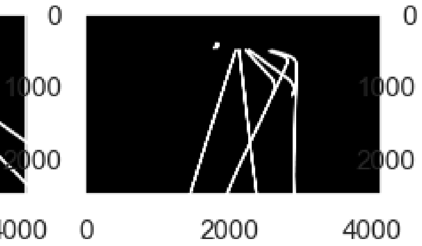
FH University of Applied Sciences TECHNIKUM WIEN

Performance metrics

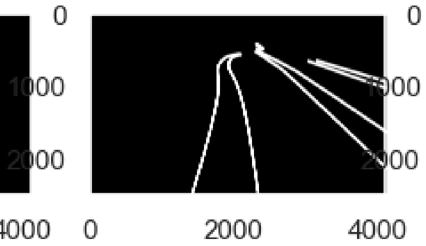




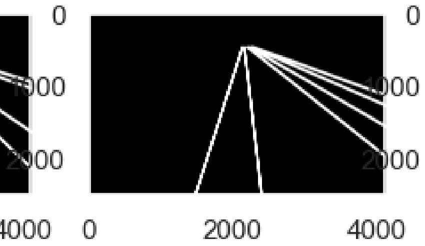
pixel_acc: 0.9983
dice: 0.9775
precision: 0.9812
recall: 0.9738
iou: 0.9559



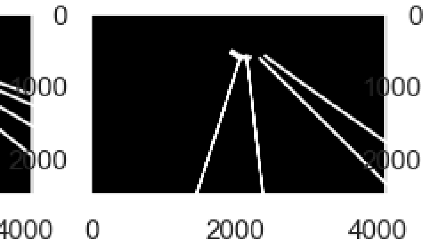
pixel_acc: 0.9499
dice: 0.3762
precision: 0.3557
recall: 0.3992
iou: 0.2317



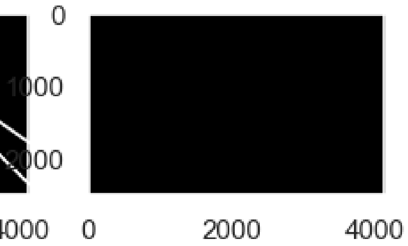
pixel_acc: 0.9205
dice: 0.0567
precision: 0.0514
recall: 0.0633
iou: 0.0292



pixel_acc: 0.9435
dice: 0.3703
precision: 0.3201
recall: 0.4392
iou: 0.2272



pixel_acc: 0.9979
dice: 0.9714
precision: 0.9787
recall: 0.9643
iou: 0.9444

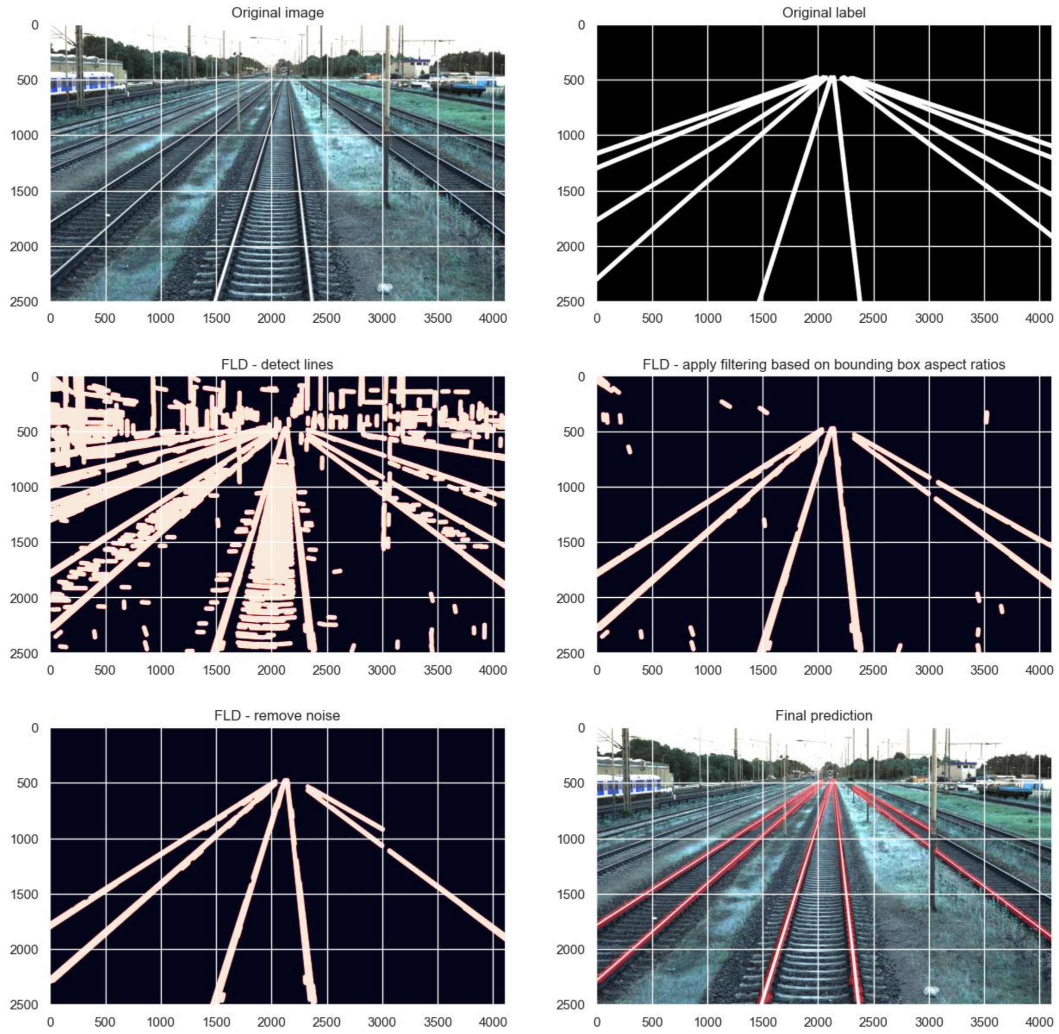


pixel_acc: 0.9622
dice: 0.0000
precision: 1.0000
recall: 0.0000
iou: 0.0000

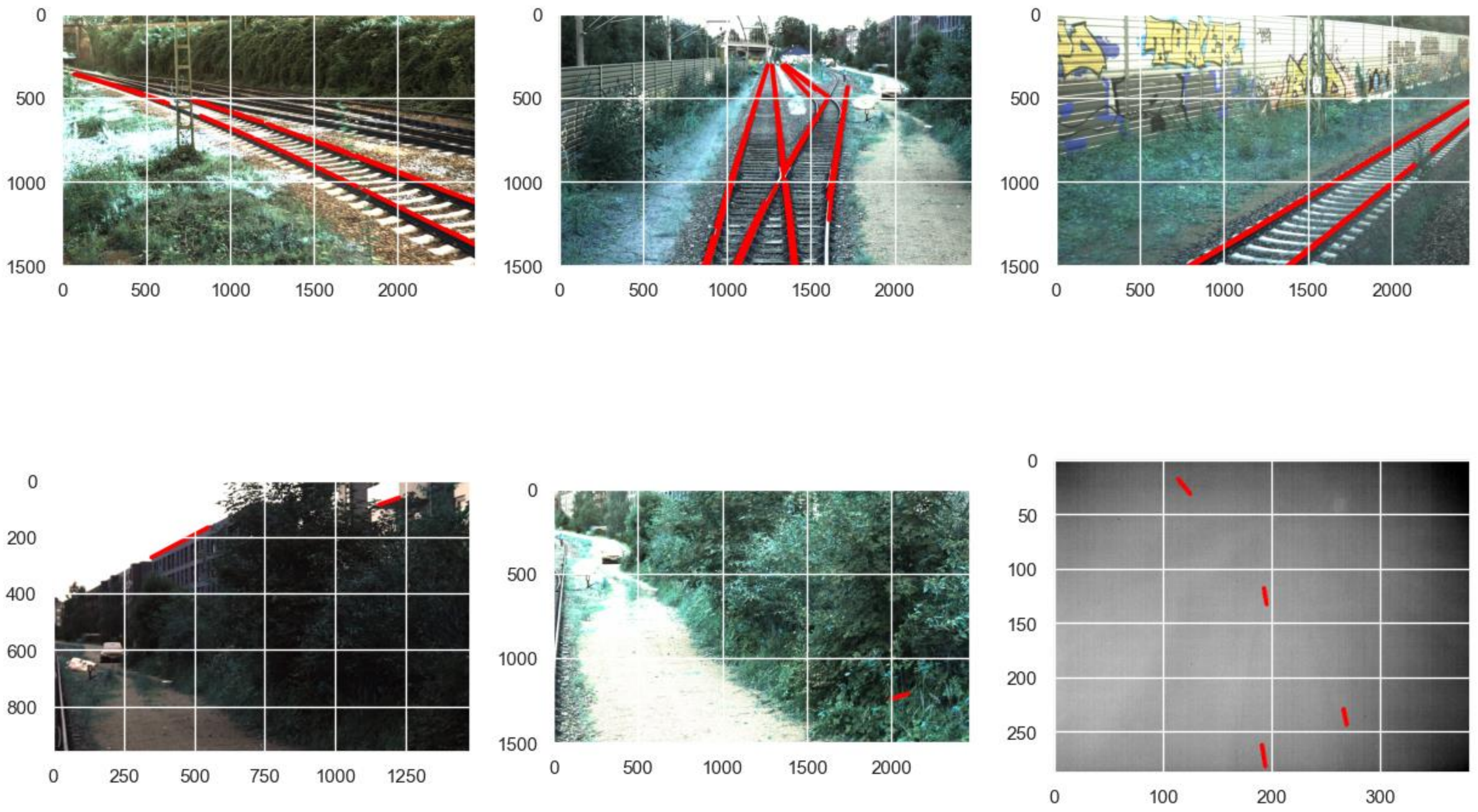
```
precision = (tp + self.eps) / (tp + fp + self.eps)
```

FH University of Applied Sciences
TECHNIKUM
WIEN

Baselining with fast line detection



Baselining with fast line detection
Good examples and bad examples



FH University of Applied Sciences
TECHNIKUM
WIEN

Deep learning approach

- YOLOv8 as framework
- Can be used for segmentation and object detection
- State-of-the-art computer vision model built by Ultralytics
- Easy to use
- Open source

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

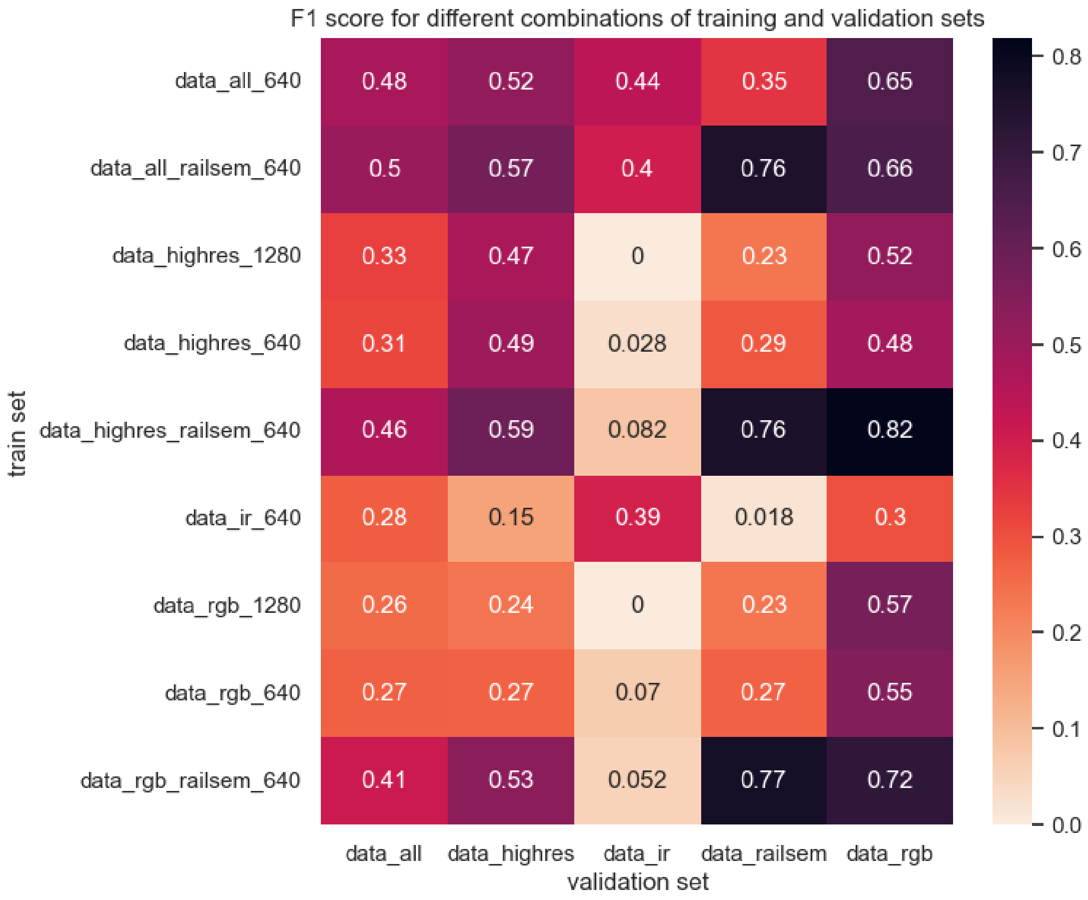
YOLOv8
First experiments - RGB

Without RailSem data

With RailSem data

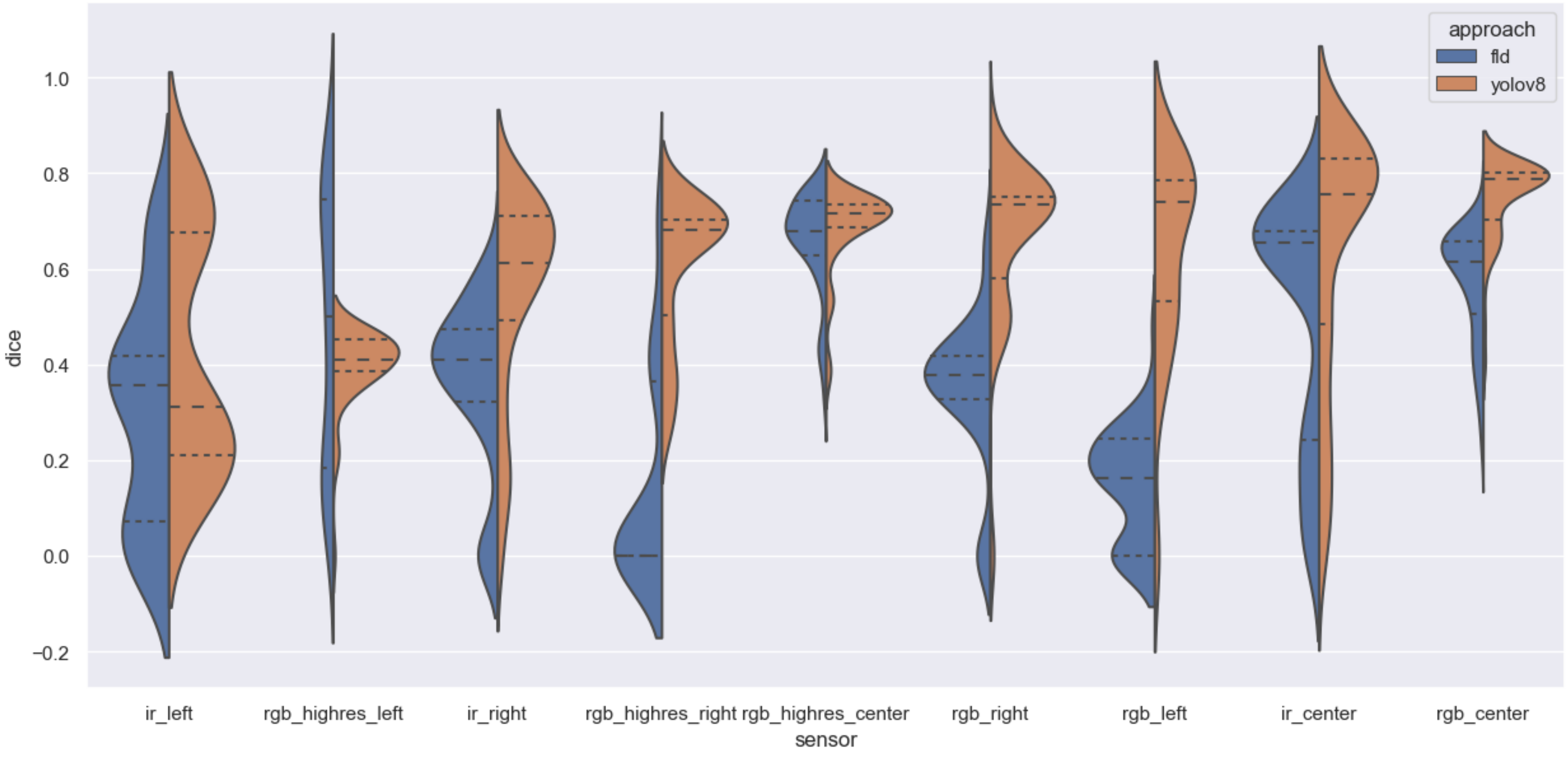
FH University of Applied Sciences TECHNIKUM WIEN

9 models have been trained and validated



FH University of Applied Sciences TECHNIKUM WIEN

YOLOv8 vs FDL
Performance on **validation** data



University of
Applied Sciences

DEMO

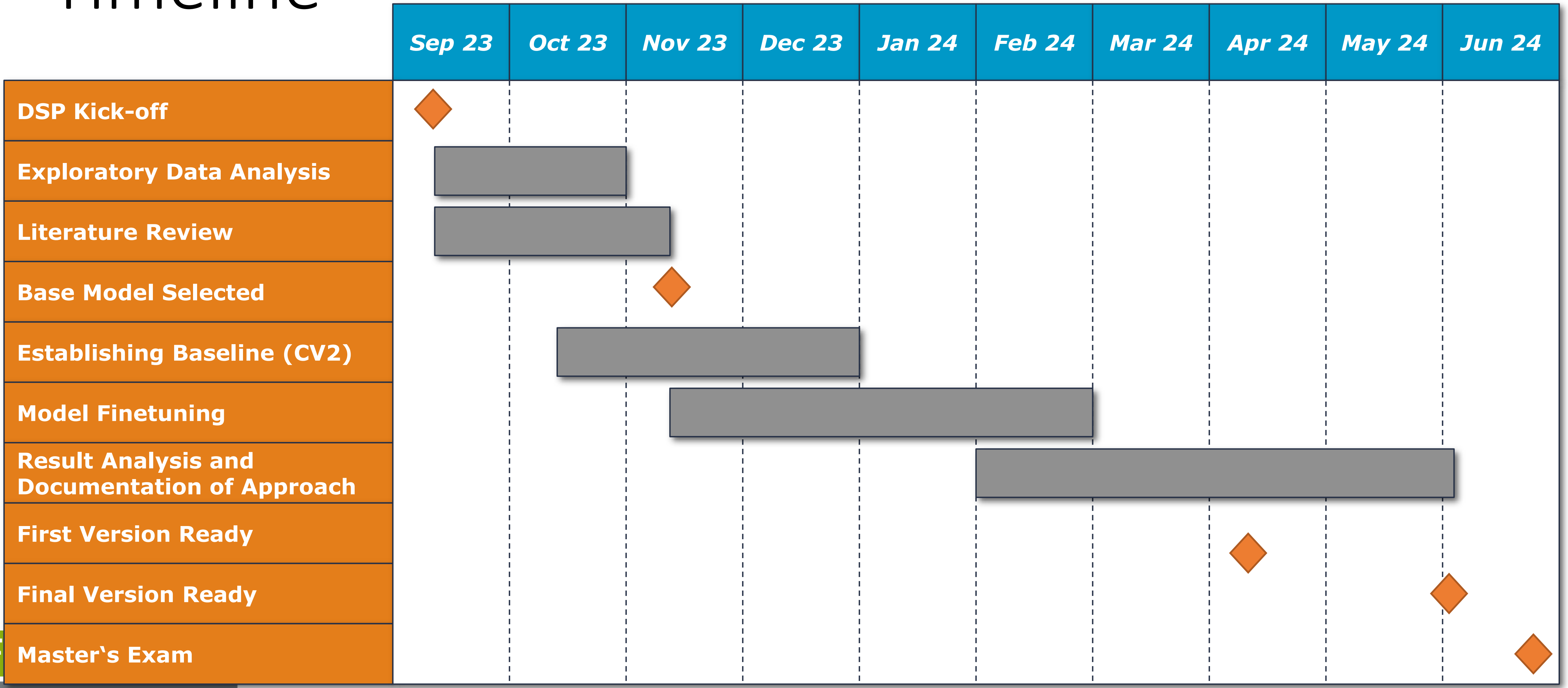
FH University of Applied Sciences TECHNIKUM WIEN

Conclusion

- Image segmentation seems to be a proper approach for modelling and solving the rail track detection problem
- Dice or f1 are well suited to assess the performance of different segmentation approaches
- Fast line detection provides a solid baseline, however YOLO outperforms FDL in most scenarios
- It seems that the performance across different sensors is similar – the orientation of the sensors has a larger impact on the results
- YOLO is well suited to be applied in real life due to fast inference

FH University of Applied Sciences TECHNIKUM WIEN

Timeline



Sep 23
Oct 23
Nov 23
Dec 23
Jan 24
Feb 24
Mar 24
Apr 24
May 24
Jun 24
DSP Kick-off
Exploratory Data Analysis
Literature Review
Base Model Selected
Establishing Baseline (CV2)
Model Finetuning
Result Analysis and Documentation of Approach
First Version Ready
Final Version Ready
Master`s Exam

Thank you!