Multi-sensor rail track detection in automatic train operations

Master's thesis in Data Science

Student: Attila Kovacs

1st Advisor: Lukas Rohatsch (FH Technikum)

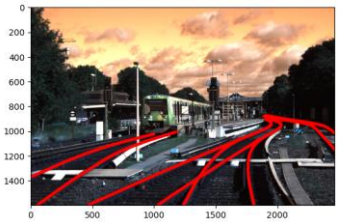
2nd Advisor: Daniele Capriotti (M2C Expert Control GmbH)

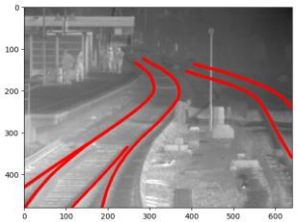
Alignment: 20.10.2023

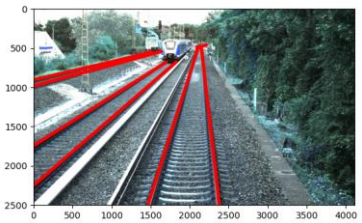
FH University of Applied Sciences
TECHNIKUM WIEN

Problem setting

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







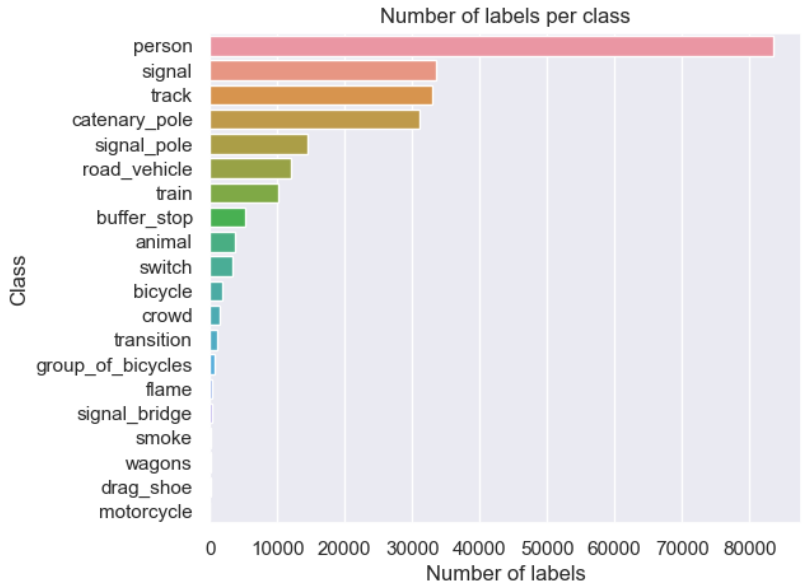
FH University of Applied Sciences
TECHNIKUM WIEN

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)

Data source

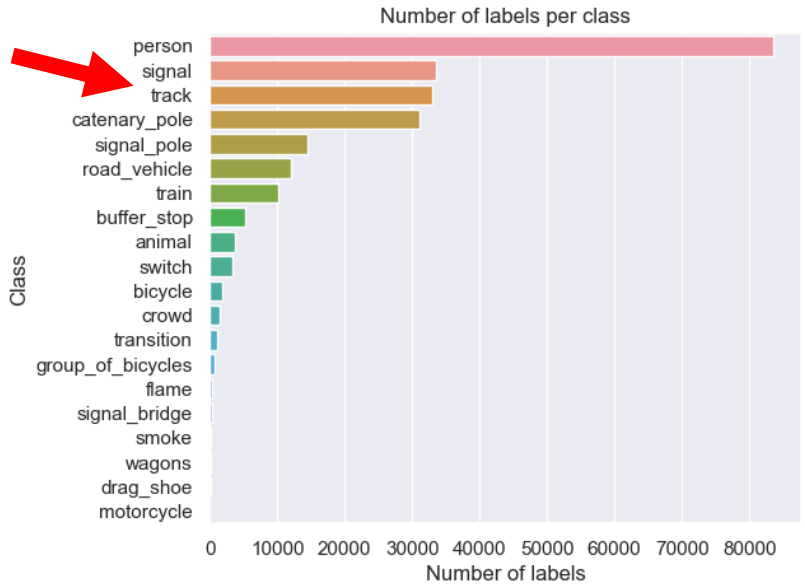
- Deutsch Bahn / Digitale Schiene Deutschland (https://digitale-schiene-deutschland.de/en/news/OSDaR23-multi-sensor-data-set-for-machine-learning).
- Images were generated between 09.09.2021 and 15.09.2021
- Total number of images: 13.952
- Sensors: 12MP RGB, 5MP RGB, IR cameras, lidar, radar, GPS, inertia sensor



FH University of Applied Sciences TECHNIKUM WIEN

Data source

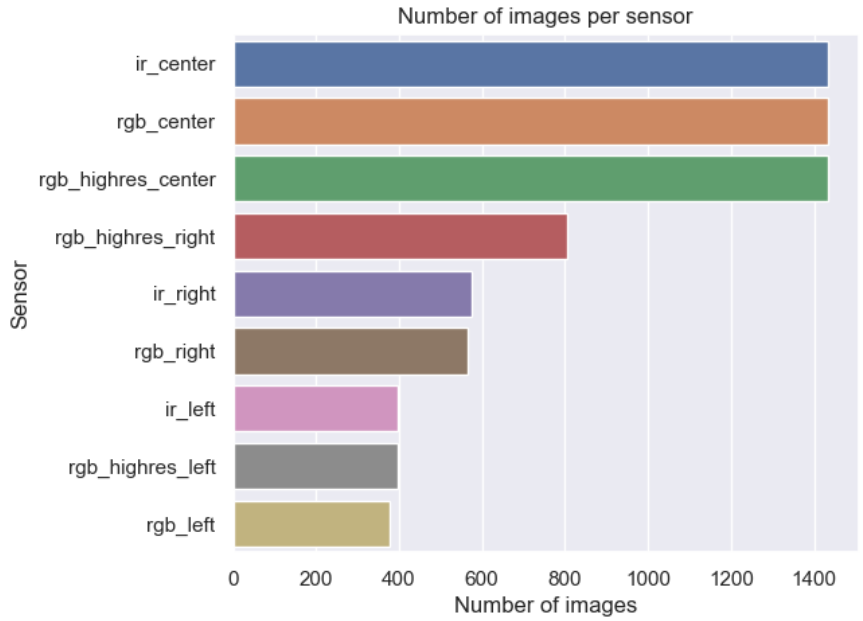
- Deutsch Bahn / Digitale Schiene Deutschland (https://digitale-schiene-deutschland.de/en/news/OSDaR23-multi-sensor-data-set-for-machine-learning).
- Images were generated between 09.09.2021 and 15.09.2021
- Total number of images: 13.952
- Sensors: **12MP RGB, 5MP RGB, IR cameras**, lidar, radar, GPS, inertia sensor

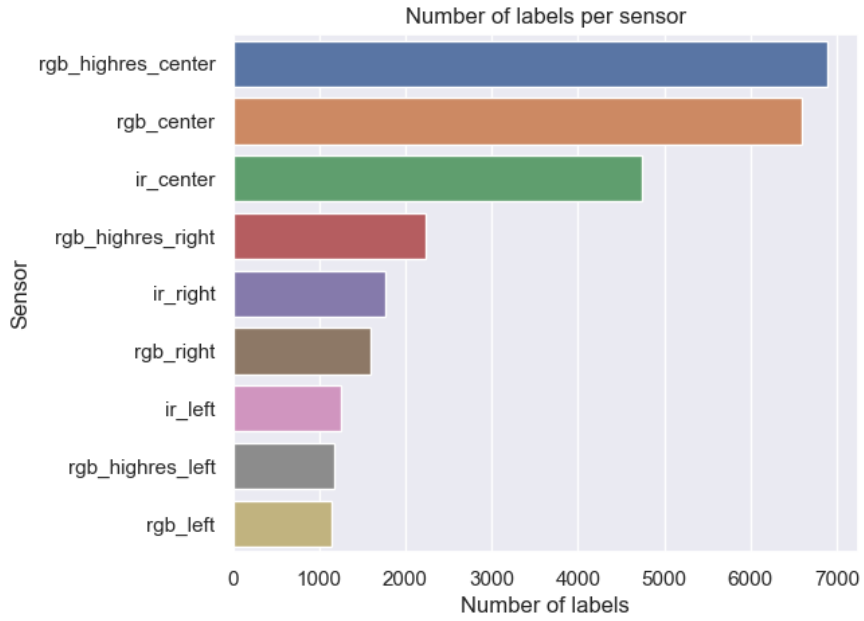


FH University of Applied Sciences TECHNIKUM WIEN

Relevant dataset (7.421 images, 27.386 labels)

- Most images are from forward facing cameras
- The number of labels in the image depends on sensor orientation and type





FH University of Applied Sciences
TECHNIKUM WIEN

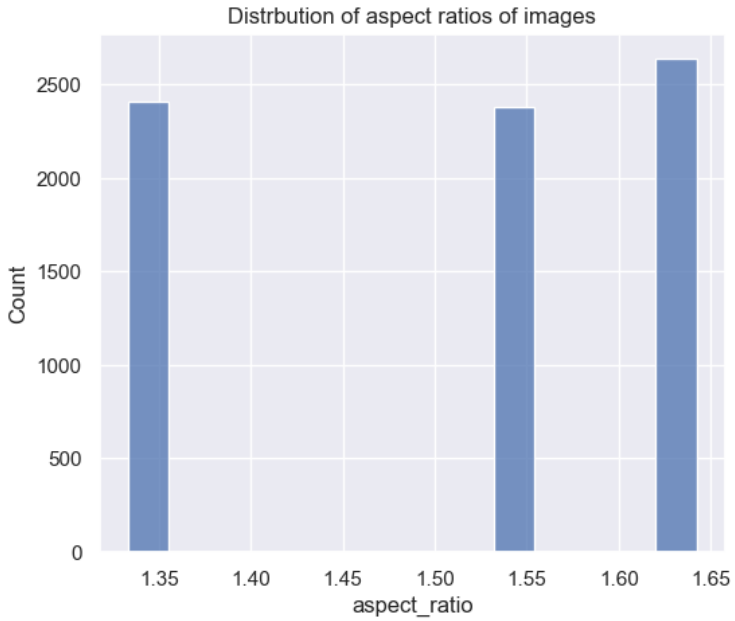
Sensors

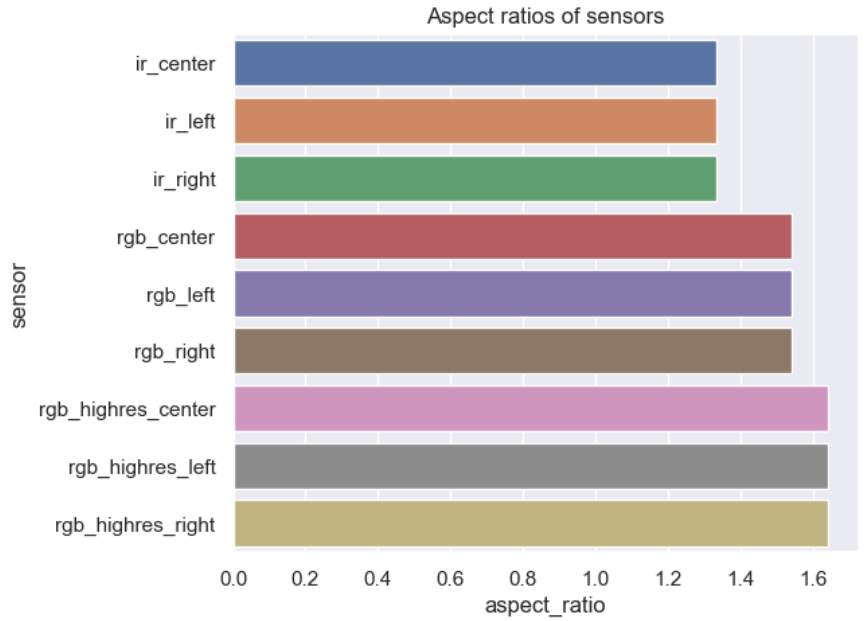
Three 12MP RGB cameras	
Type	Teledyne GenieNano 5GigE C4040
Sensor data	RGB images (8 Bit, PNG)
Resolution	4 112 × 2 504 px
Sampling frequency	10 Hz (synchronized)
Alignment	trident (in driving direction diagonal left, central and diagonal right)
Three 5MP RGB cameras	
Type	Teledyne GenieNano C2420
Sensor data	RGB images (8 Bit, PNG)
Resolution	2 464 × 1 600 px
Sampling frequency	10 Hz (synchronized)
Alignment	trident
Three IR cameras	
Type	Teledyne Calibir DXM640
Sensor data	grayscale images (8 Bit, PNG)
Resolution	640 × 480 px
Sampling frequency	10 Hz (synchronized)
Alignment	trident

FH University of Applied Sciences TECHNIKUM WIEN

Aspect ratios (= width / height)

- All images are generated with one of three camera types
- Double check whether there are other images in the set

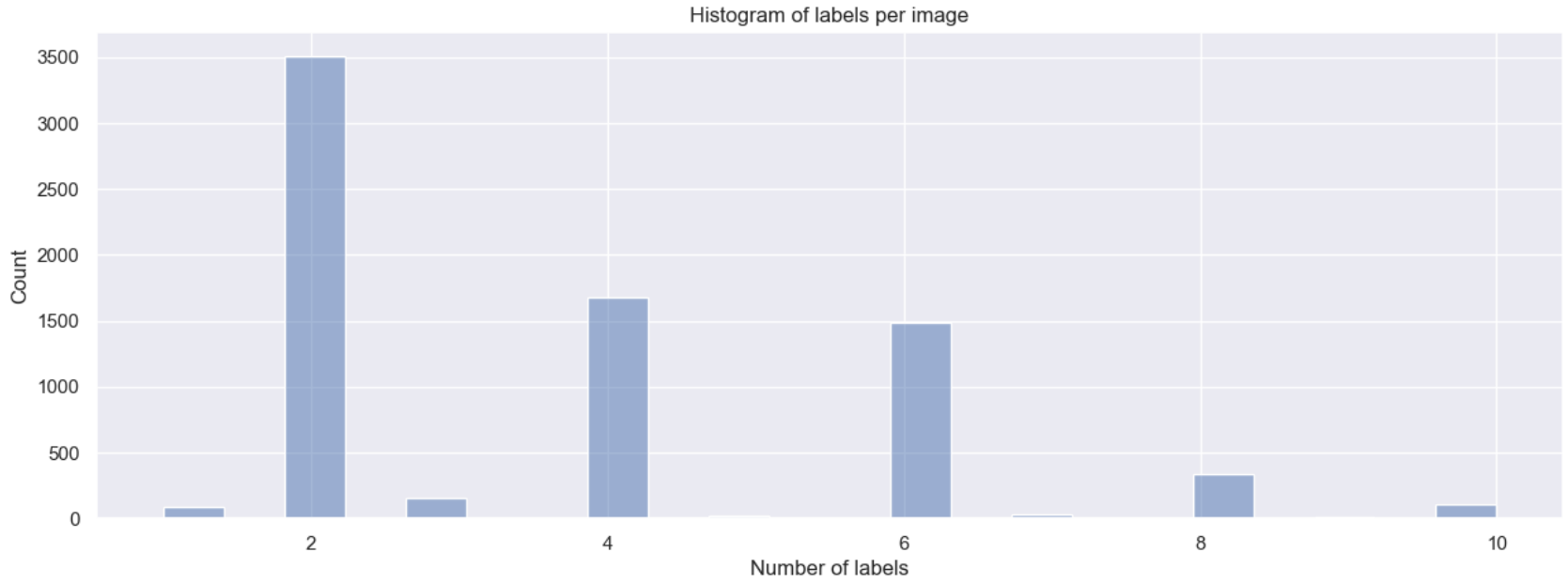




FH University of Applied Sciences TECHNIKUM WIEN

Labels per image

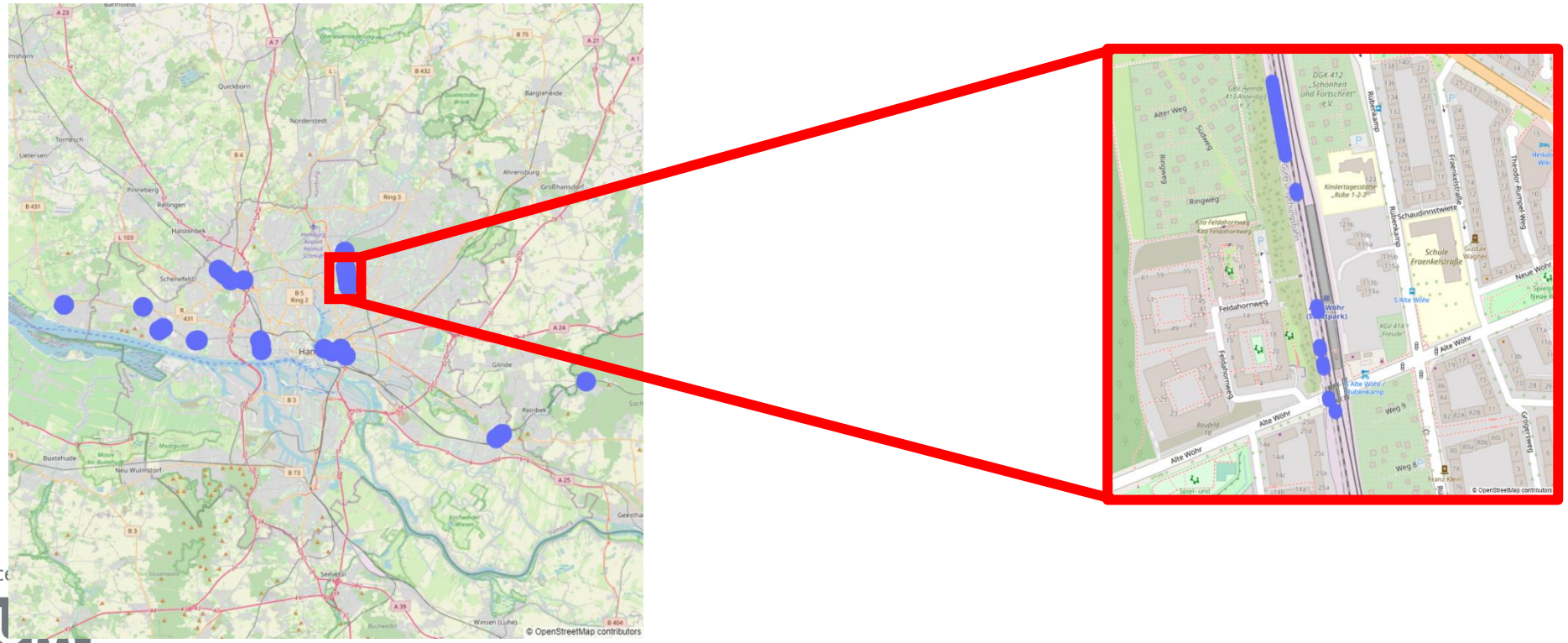
- Most tracks are labelled in pairs
- However, there is also a small number of images with uneven number of tracks



FH University of Applied Sciences TECHNIKUM WIEN

Location of images

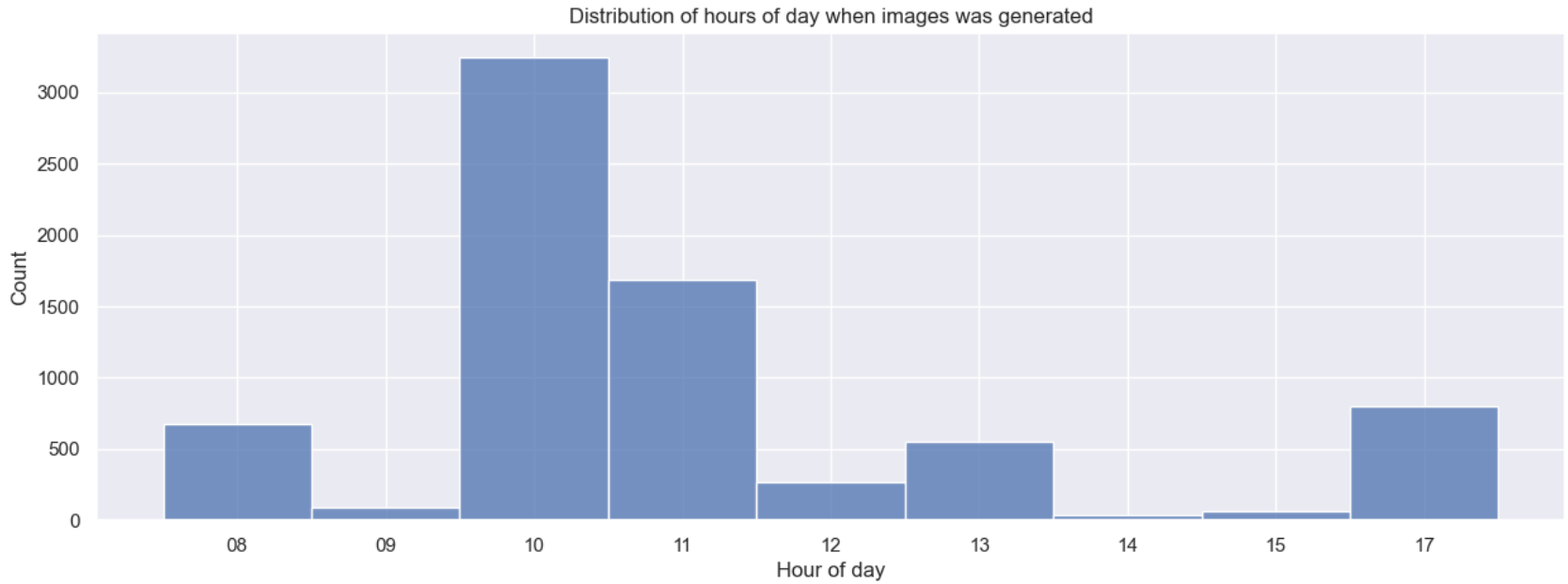
- All images/videos were taken around Hamburg



FH University of Applied Science
TECHNIKUM WIEN

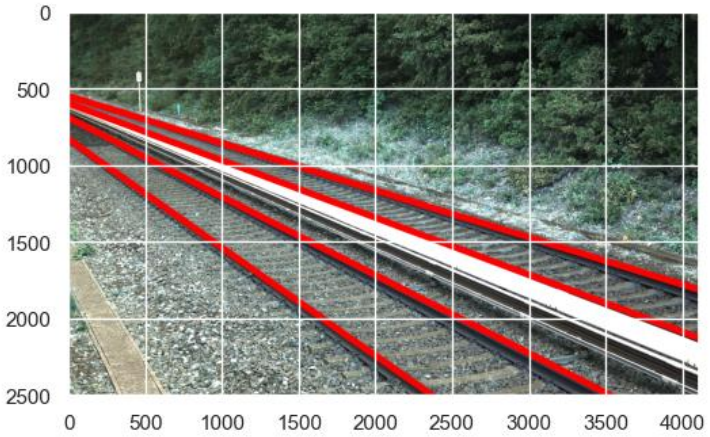
Time of images

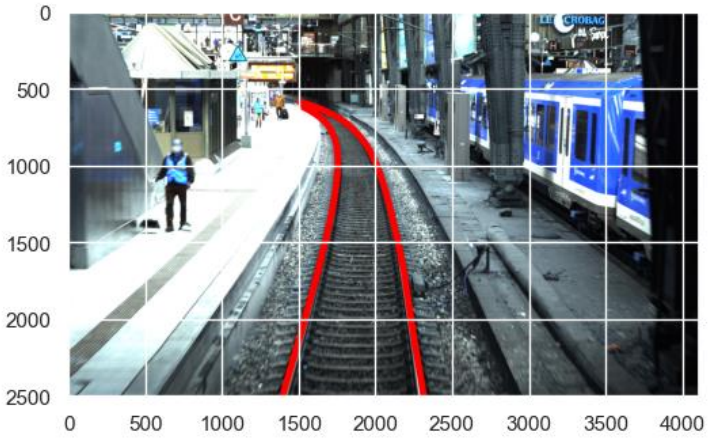
- All images were taken between 8 AM and 5 PM

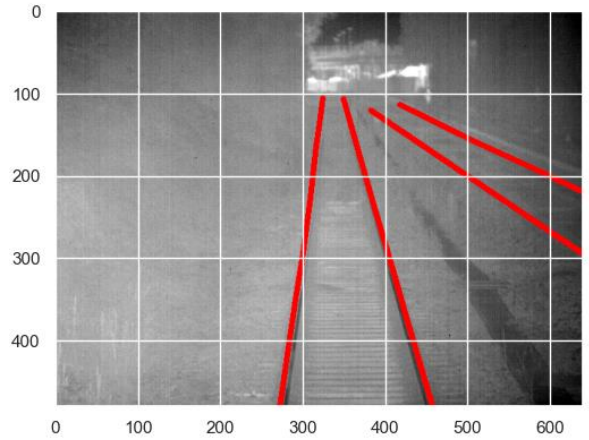


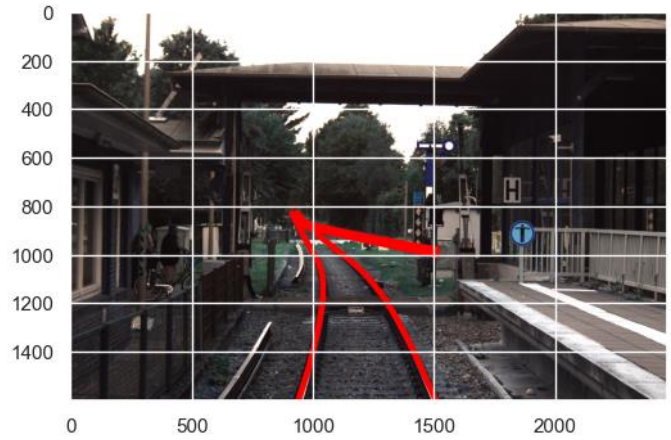
FH University of Applied Sciences TECHNIKUM WIEN

Some examples of labeled images





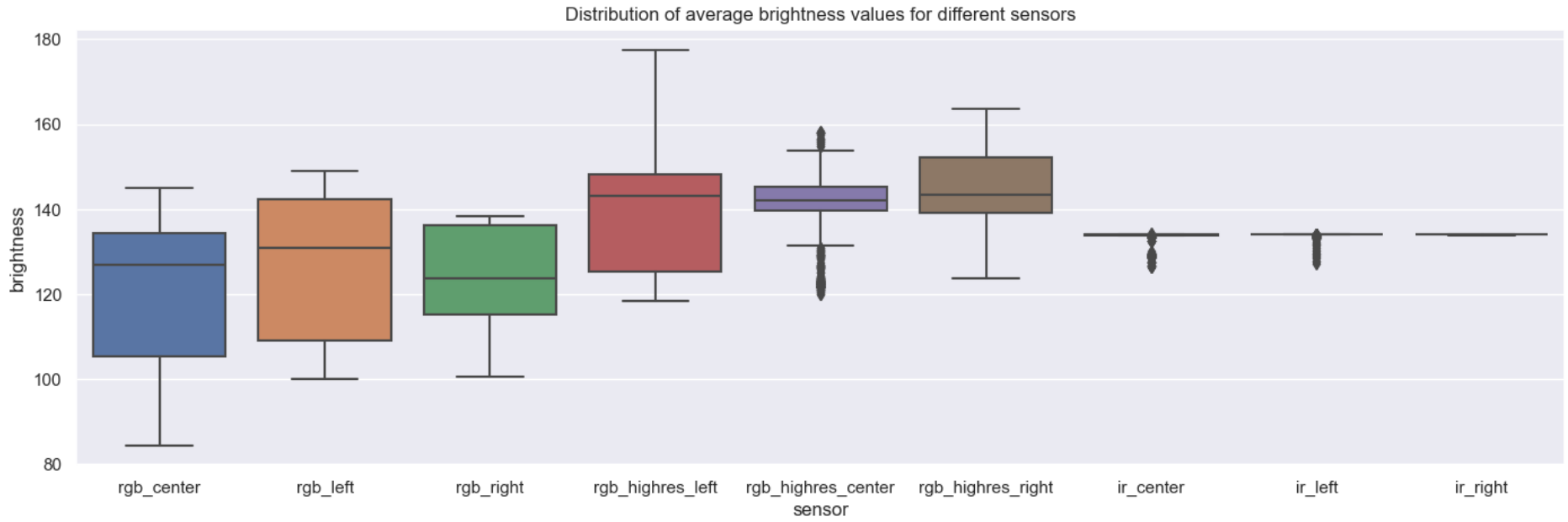




FH University of Applied Sciences
TECHNIKUM WIEN

Brightness of images

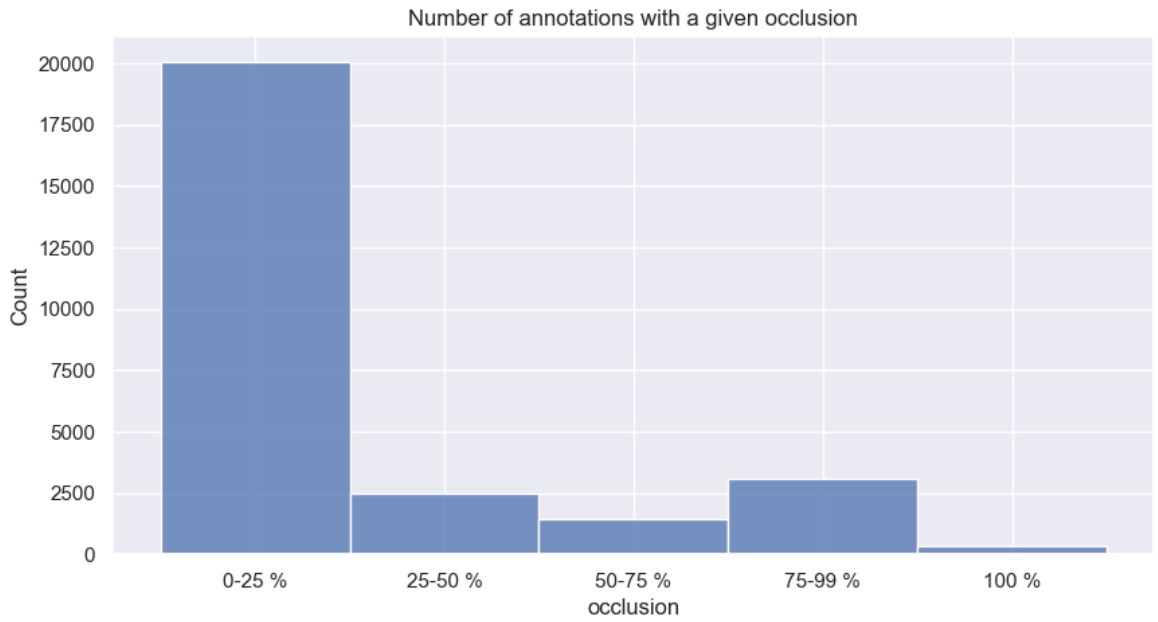
- RGB images are darker thank high resolution images (not pixel related)



FH University of Applied Sciences TECHNIKUM WIEN

Occlusion

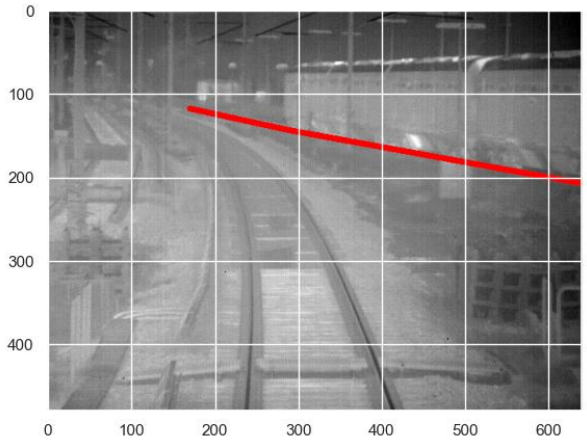
- Most of the labels have a good visibility

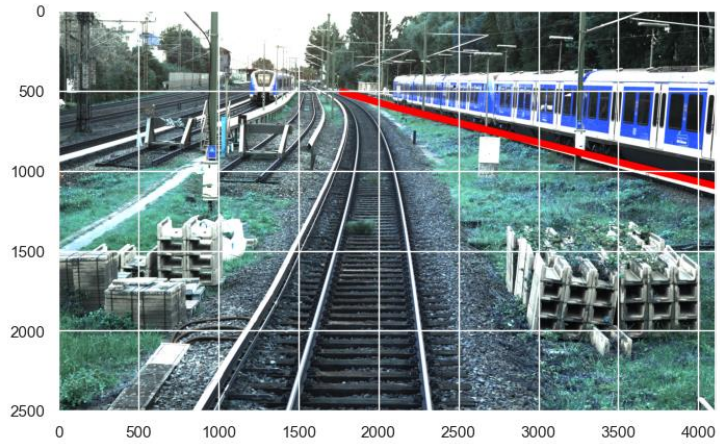


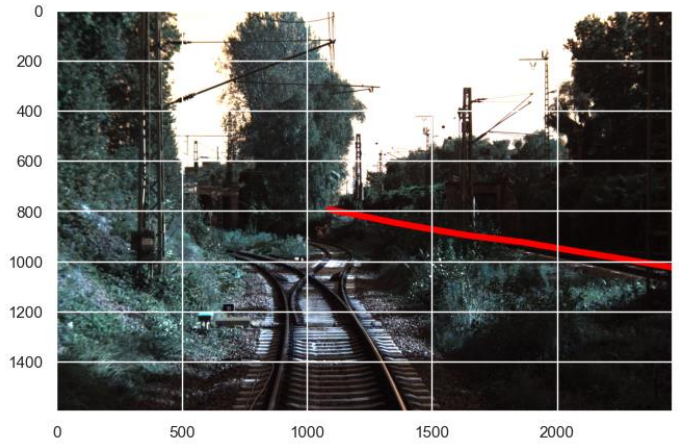
FH University of Applied Sciences TECHNIKUM WIEN

Examples of occlusion = 100%

- Most of the labels have a good visibility



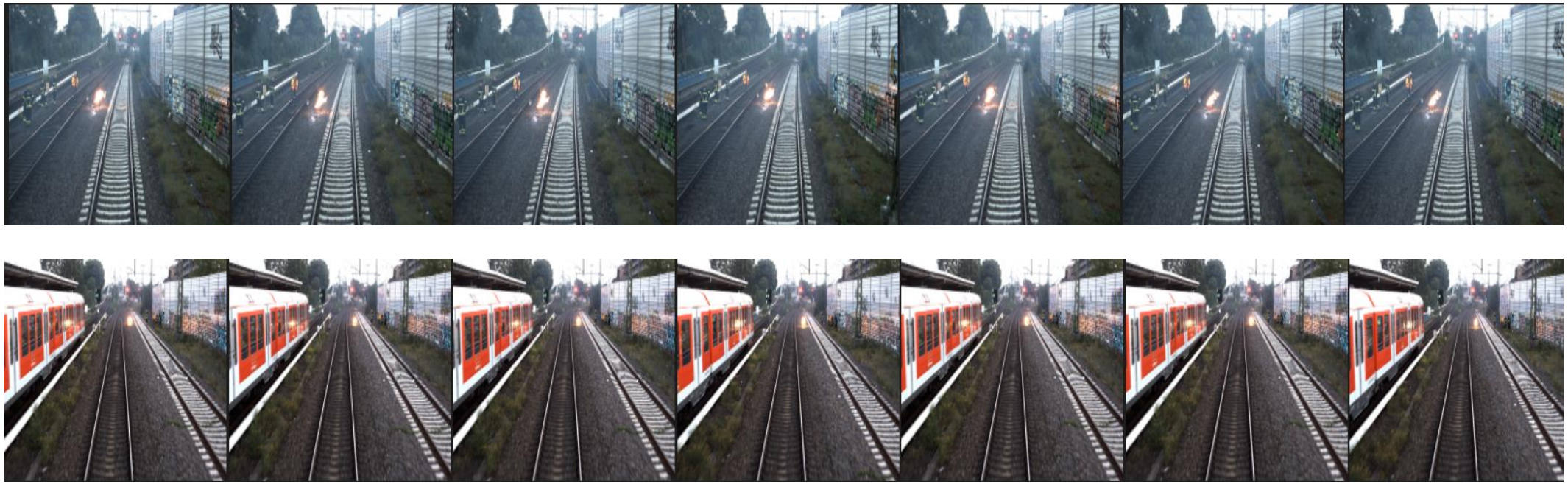




FH University of Applied Sciences TECHNIKUM WIEN

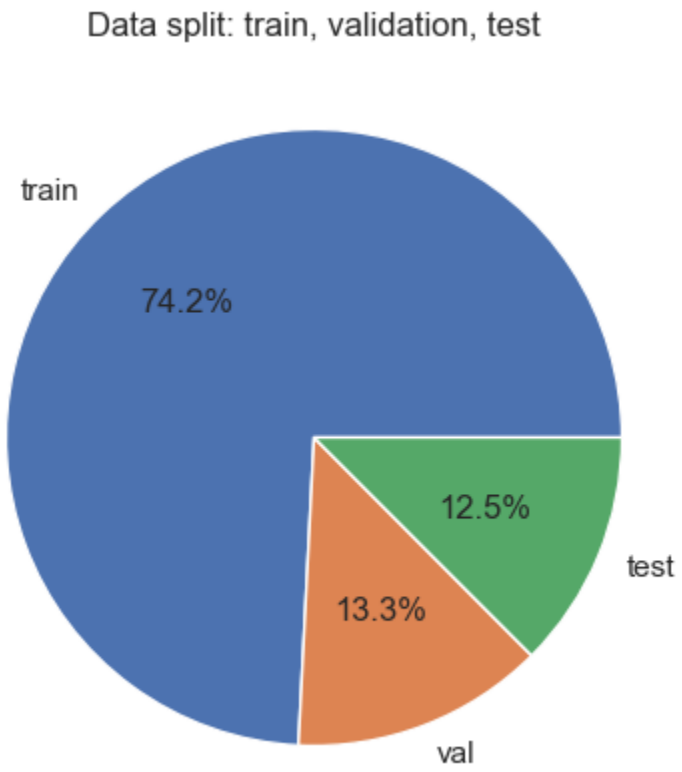
Splitting into trail, validation and test set

- Many images are very similar which needs to be taken into account



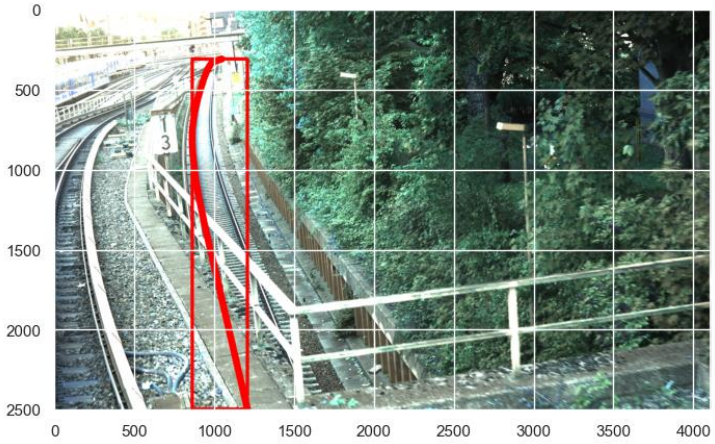
Splitting data set

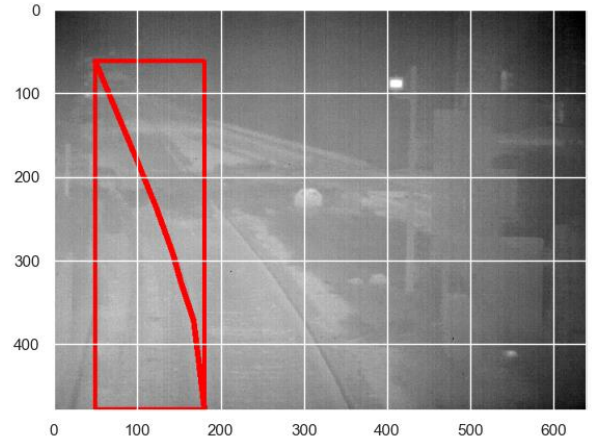
- Given the large similarity between images, we split data by randomly assigning videos to either train, validation or test set
- Advantage: fair testing as we have a „data leakage“ between the sets
- Disadvantage: validation batches might have low variety

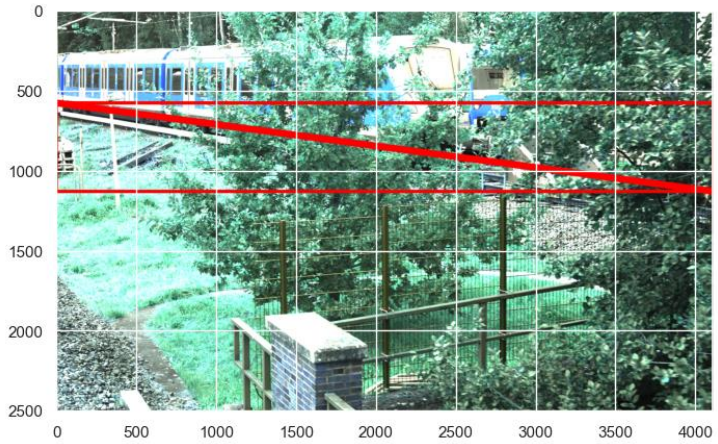


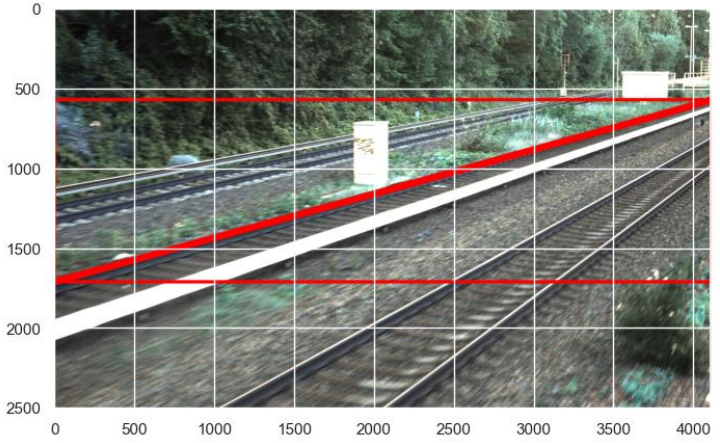
FH University of Applied Sciences TECHNIKUM WIEN

Orientation of labels



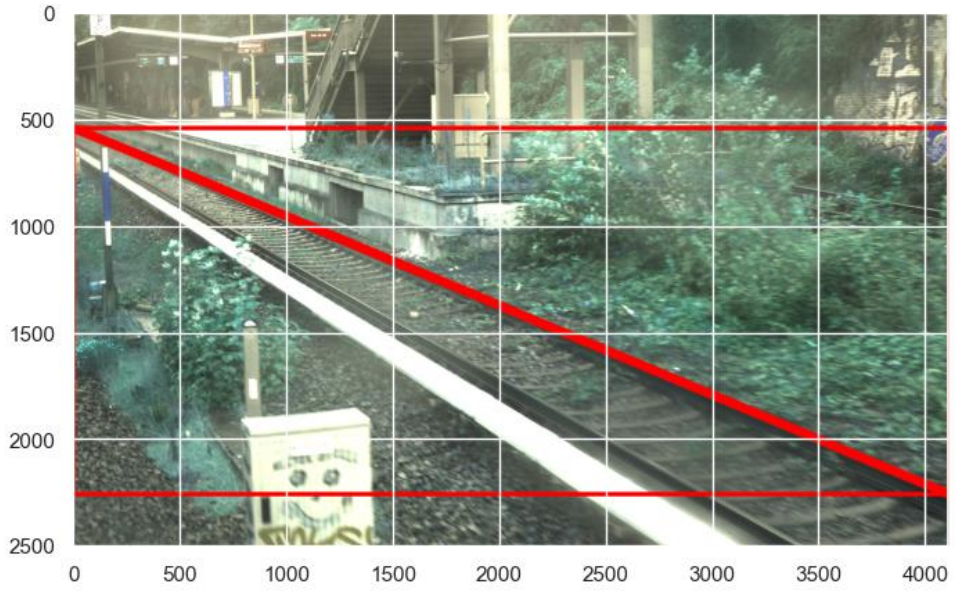


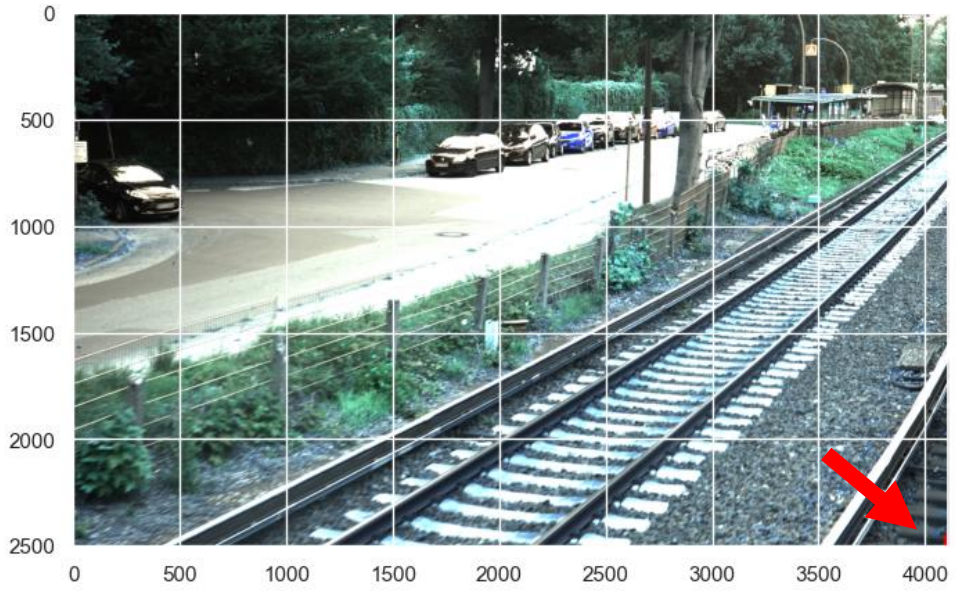




FH University of Applied Science
TECHNIKUM WIEN

Size of labels

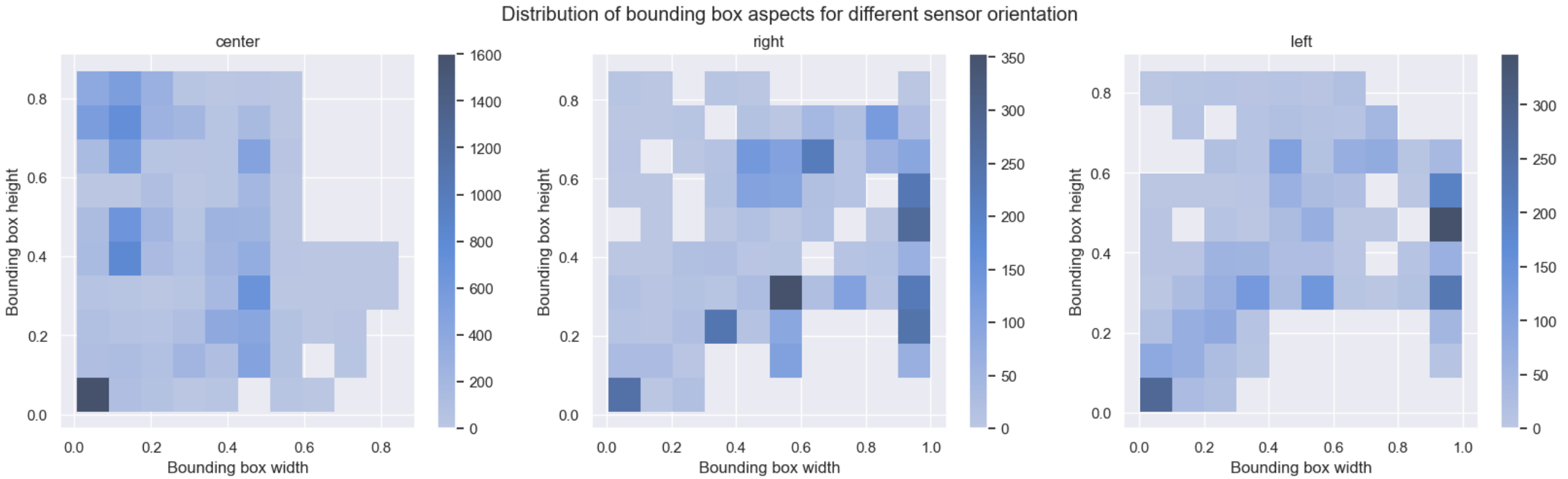




FH University of Applied Sciences
TECHNIKUM WIEN

Histogram of bounding box shapes and sizes

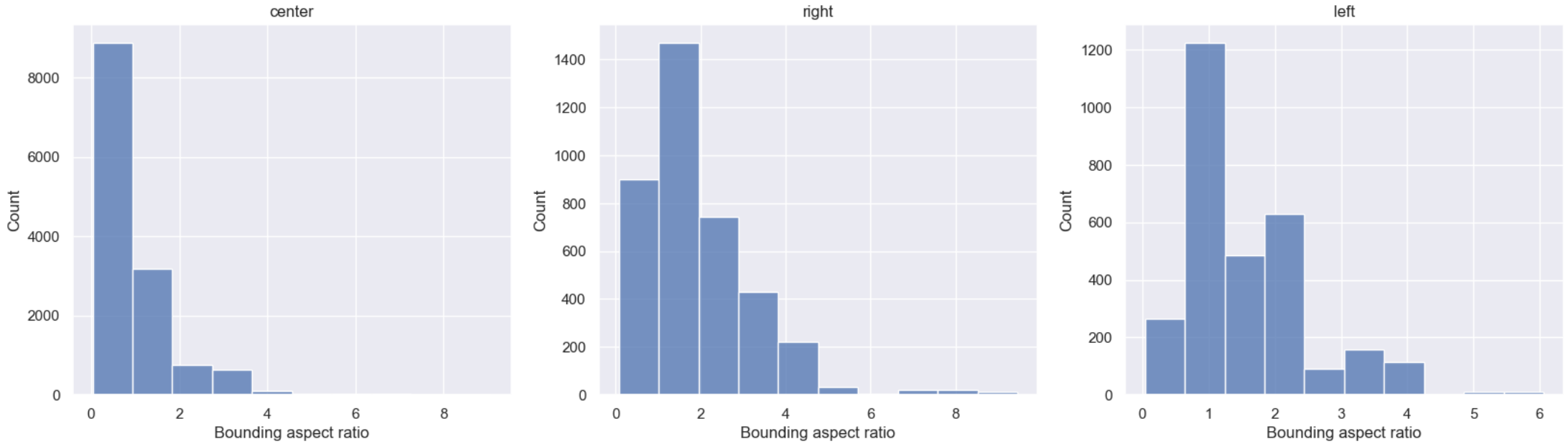
- Bounding boxes format is related to the orientation of sensors



FH Applied Sciences TECHNIKUM WIEN

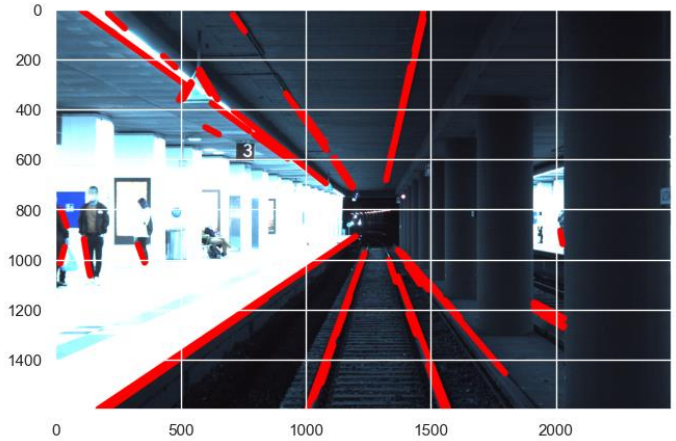
Aspect ratio of bounding boxes

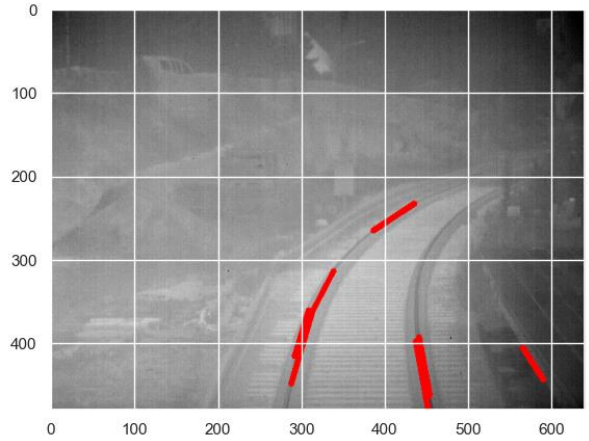
- These features can guide the gradient treshold appraoch

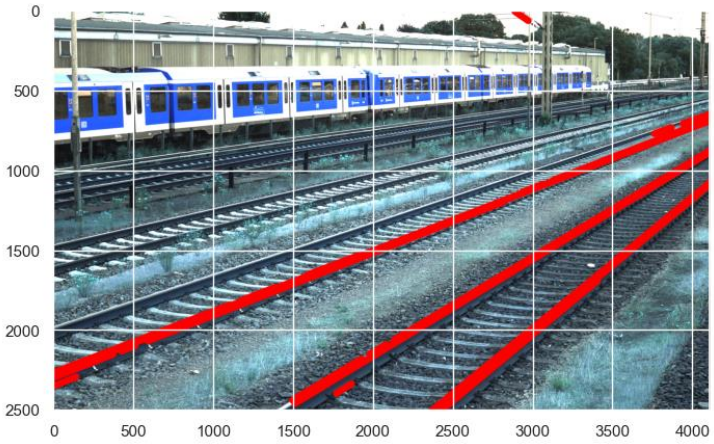


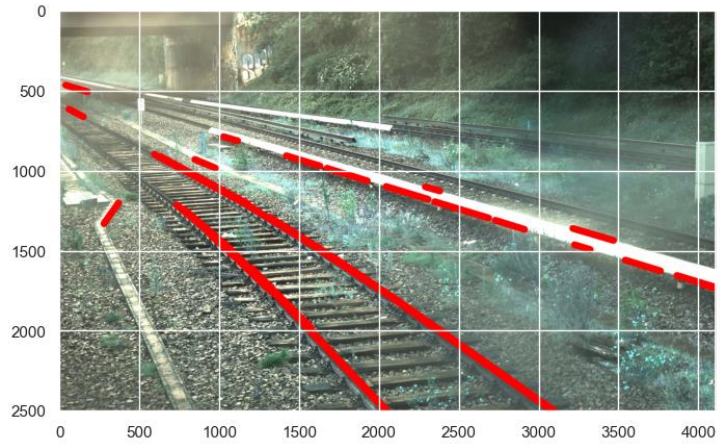
FH University of Applied Sciences TECHNIKUM WIEN

Baselining with CV2-FastLineDetection Filtering based on aspect ratio









FH University of Applied Sciences TECHNIKUM WIEN

Next steps

- Establish baseline
  - Finetune parameters → separately for highres/rgb/ir
  - Analyze results, select optimization function
- Generate label masks for images
- Select image segmentation model

Fragen

- Ist die aktuelle Richtung in Ordnung?
  - Fragestellung: Wo liegen Schienen genau (Hauptschiene, Nebenschienen)
  - Anreichern der Daten mit alternativ Sets
- FastLineDetection als Baseline
  - Finetuning
  - Parametertuning
  - Result selection
- Modellierung als „Segmentation Model“
- Lösungsansatz YOLOv8
- Nächste Abstimmung: 17.11.2023 13:00

FH Technikum Wien