# Praktikum Visual Computing

## Object Detection for Chest X-Ray Imaging

TECHNISCHE UNIVERSITÄT DARMSTADT

By Yihao Li
Tutor: Sanner Autoine

# Chest ImaGenome Dataset

TECHNISCHE UNIVERSITÄT DARMSTADT

- a subset of the MIMIC-CXR dataset
- "golden dataset":manually annotated
  - 256 combinations of relation annotations

| patient_id | | Region | sentence |
|---|---|---|---|
| 15184836 | 00046130-fd952ef0-57f2948d-491a16b4-5db3a18c.dcm | ['cardiac silhouette', 'left hilar structures', 'mediastinum', 'right hilar structures', 'upper mediastinum'] | The cardiomediastinal and hilar contours are within normal limits. |
| 15184836 | 00046130-fd952ef0-57f2948d-491a16b4-5db3a18c.dcm | ['left lung', 'right lung'] | The lung fields are clear. |
| 15184836 | 00046130-fd952ef0-57f2948d-491a16b4-5db3a18c.dcm | ['left apical zone', 'left lung', 'right apical zone', 'right lung'] | There is no pneumothorax, fracture or dislocation. |
| 12930467 | 005043e2-a4e25d1d-aae26631-732a2db0-38412248.dcm | ['left hilar structures', 'left lower lung zone', 'left lung', 'right hilar structures', 'right lower lung zone', 'right lung'] | lower lungs are grossly clear, though there is vascular crowding |

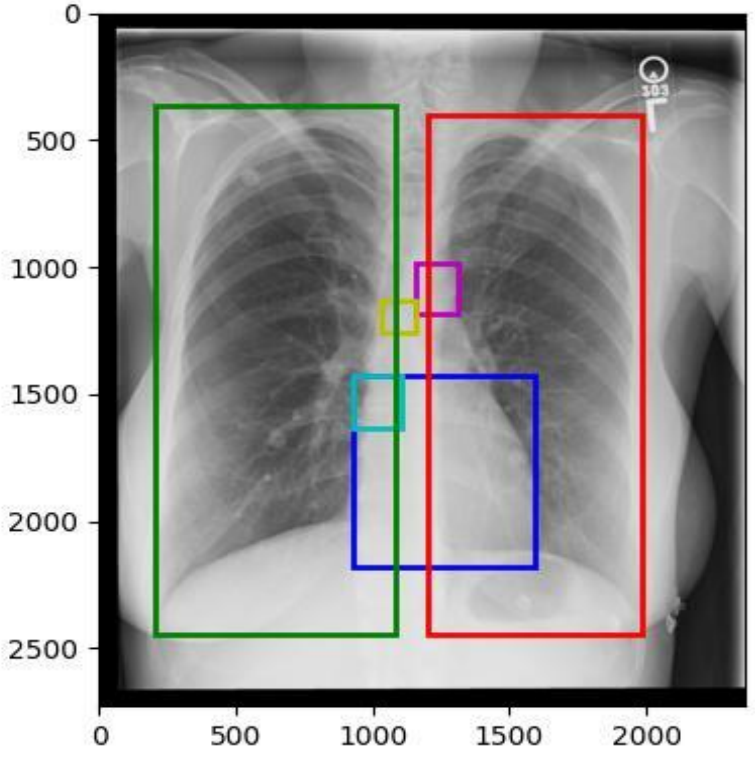
  - over 670, 000 localized comparison relations ( improved, worsened, or no change) between the anatomical locations

# Chest ImaGenome Dataset

TECHNISCHE UNIVERSITÄT DARMSTADT

Average object size: 648 x 586

Average image size: 2675 x 2794

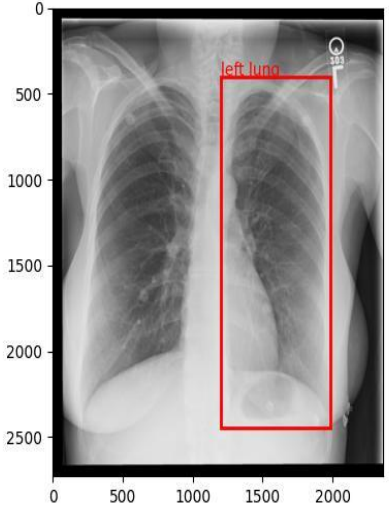


# Chest ImaGenome Dataset

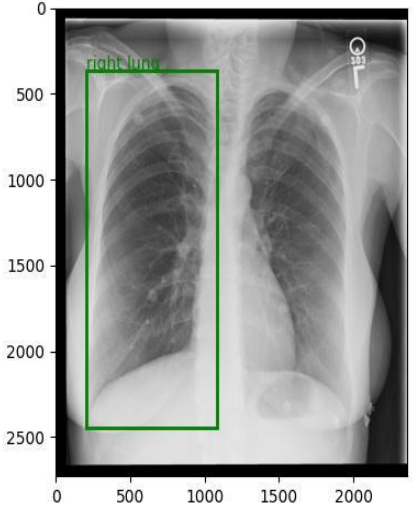
TECHNISCHE UNIVERSITÄT DARMSTADT

## Object Size Biggest 3 :

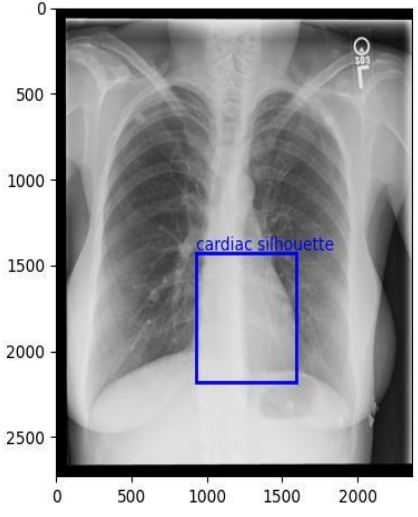
Left lung: 972 x 1809




Right lung: 958 x 1794




Cardiac silhouette: 1070 x 777

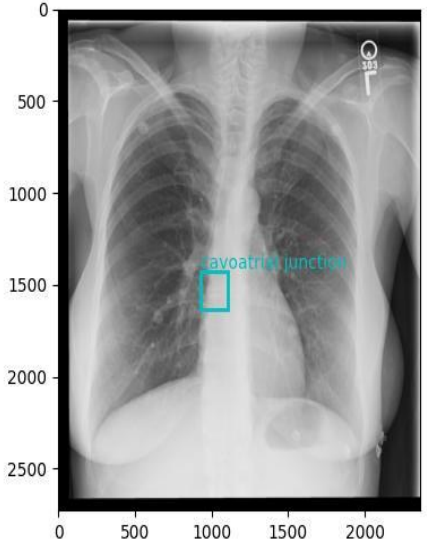


# Chest ImaGenome Dataset

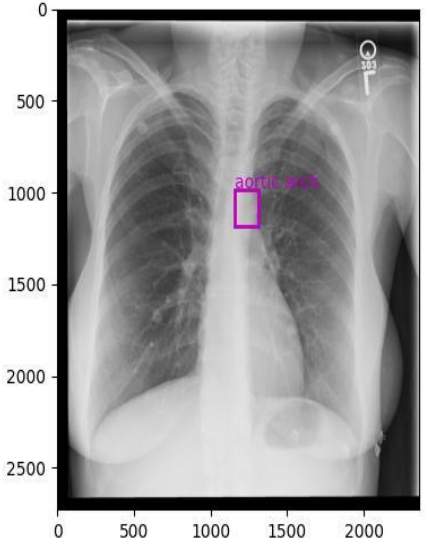
TECHNISCHE UNIVERSITÄT DARMSTADT

## Object Size Smallest 3 :

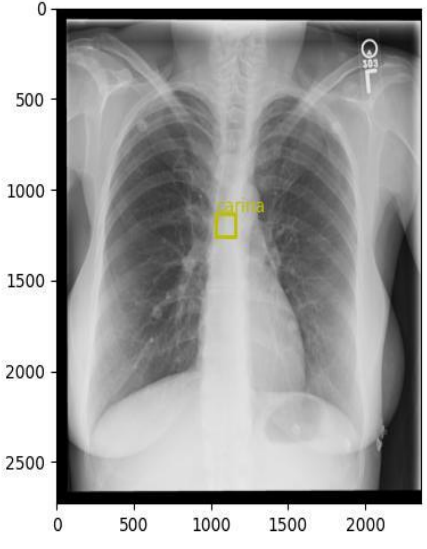
Cavoatrial junction: 314 x 224




Aortic arch: 252 x 226




Carina: 121 x 122



# Detection Method

TECHNISCHE UNIVERSITÄT DARMSTADT

| Stages | Detection Method | Advantages | Disadvantages |
|---|---|---|---|
| One-stage | YOLO | • Transform object detection task into a regression problem<br>• Highly speed up the detection | • Position is not accurate<br>• Detection effect of small and dense instances is not efficient |
| | SSD | • Use feature layers of different scales extracted from the backbone<br>• Speed of detection is fast | • Features for smaller targets may be lost |
| Two-stage | R–CNN | • Extract and learn features from CNNs automatically<br>• Compute the neural network features on each of regions of interest | • Long time to acquire regional targets<br>• Feature extraction is complex |
| | Fast R–CNN | • Run the neural network once on the whole image<br>• Training speed is significantly enhanced | • Restricted by selective search algorithms |
| | Faster R–CNN | • Region Proposal Network(RPN) instead of selective search algorithms<br>• Improve detection speed | • Not satisfy the requirement of real-time detection due to two stage structure |

# Open source Library

TECHNISCHE UNIVERSITÄT DARMSTADT

Mmdetection:

- Maintainced
- Framework base: Pytorch
- Pre-trained models:
  - Fast R-CNN
  - Faster R-CNN
  - SSD
  - RetinaNet
  - YOLOv3
  - other state-of-the-art detection models

# Training

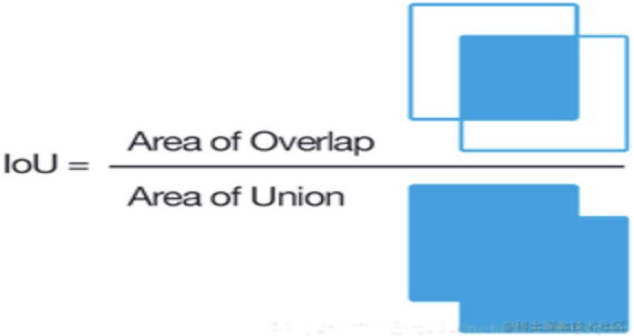
TECHNISCHE UNIVERSITÄT DARMSTADT

COCO format


```
{
    "annotations": [
        {
            "id": 120,
            "image_id": 4,
            "category_id": 4,
            "area": 74529.0,
            "bbox": [245.0, 1759.0, 273.0, 273.0],
            "iscrowd": 0
        }
    ],
    "categories": [
        {
            "id": 1,
            "name": "left lung"
        }
    ],
    "images": [
        {
            "id": 4,
            "width": 3056,
            "height": 2544,
            "file_name": "image1.jpg"
        }
    ]
}
```


- Split data:8:1:1
- Data augmentation:
  - translation: shift small distance
  - rotation: less than 10 degrees

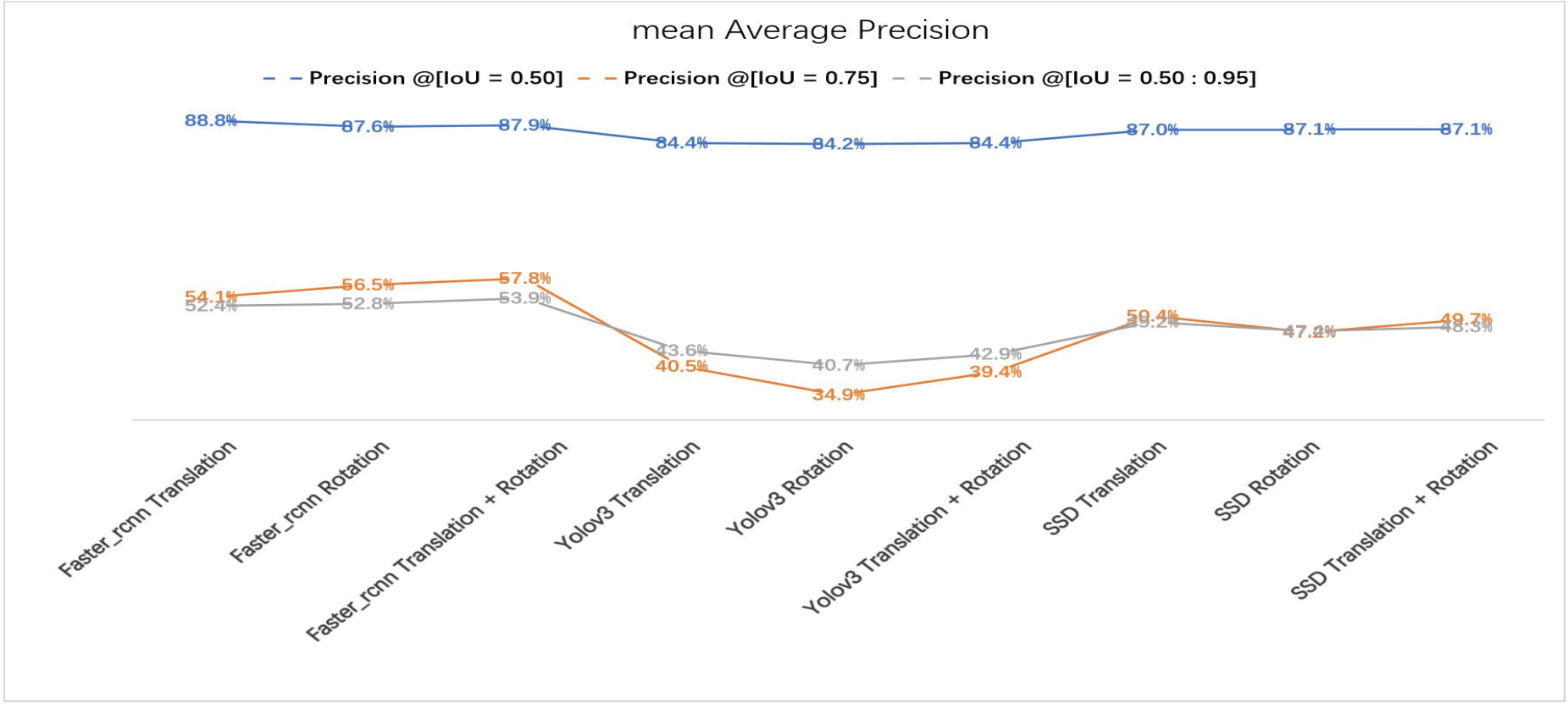
# Evaluations metrics




$$\text{Precision} = \frac{TP}{TP + FP} = \frac{TP}{\text{all detections}}$$

mAP: average of the precision values at different IoU thresholds

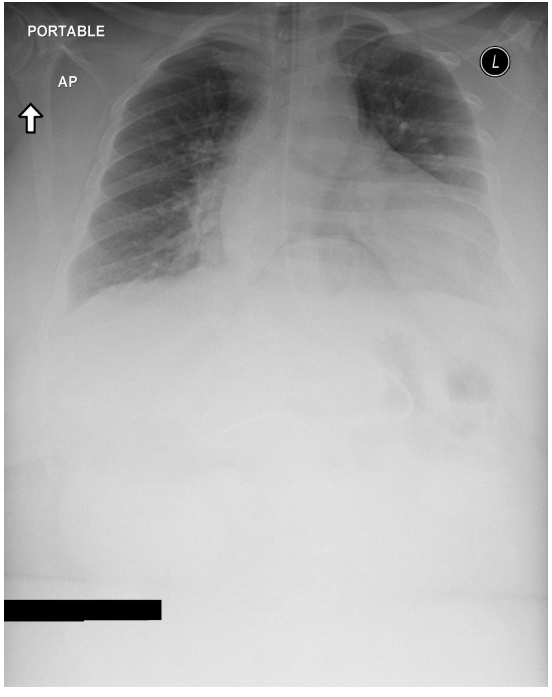
# Result

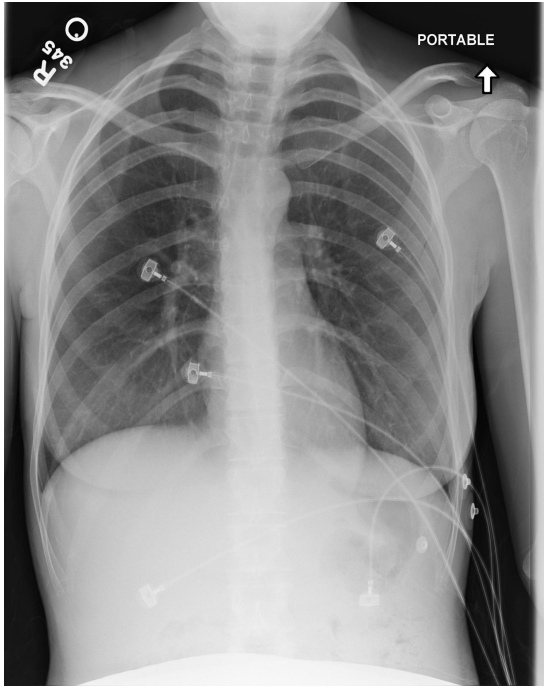

mean Average Precision
Precision @[IoU = 0.50]
Precision @[IoU = 0.75]
Precision @[IoU = 0.50 : 0.95]
88.8%
87.6%
87.9%
84.4%
84.2%
84.4%
87.0%
87.1%
87.1%
54.1%
56.5%
57.8%
40.5%
34.9%
39.4%
59.4%
49.7%
52.4%
52.8%
53.9%
43.6%
40.7%
42.9%
49.2%
47.2%
48.3%
Faster_rcnn Translation
Faster_rcnn Rotation
Faster_rcnn Translation + Rotation
Yolov3 Translation
Yolov3 Rotation
Yolov3 Translation + Rotation
SSD Translation
SSD Rotation
SSD Translation + Rotation

# Result

TECHNISCHE UNIVERSITÄT DARMSTADT

| | | Precision @[IoU = 0.50] | Precision @[IoU = 0.75] | Precision @[IoU = 0.50 : 0.95] |
|---|---|---|---|---|
| Faster_rcnn | Translation | 88,8% | 54,1% | 52,4% |
| | Rotation | 87,6% | 56,5% | 52,8% |
| | Translation + Rotation | 87,9% | 57,8% | 53,9% |
| Yolov3 | Translation | 84,4% | 40,5% | 43,6% |
| | Rotation | 84,2% | 34,9% | 40,7% |
| | Translation + Rotation | 84,4% | 39,4% | 42,9% |
| SSD | Translation | 87% | 50,4% | 49,2% |
| | Rotation | 87,1% | 47,2% | 47,4% |
| | Translation + Rotation | 87,1% | 49,7% | 48,3% |


PORTABLE
AP
L


PORTABLE


R
SEMI-SUPINE
PORTABLE

# Left & Right organs

TECHNISCHE UNIVERSITÄT DARMSTADT

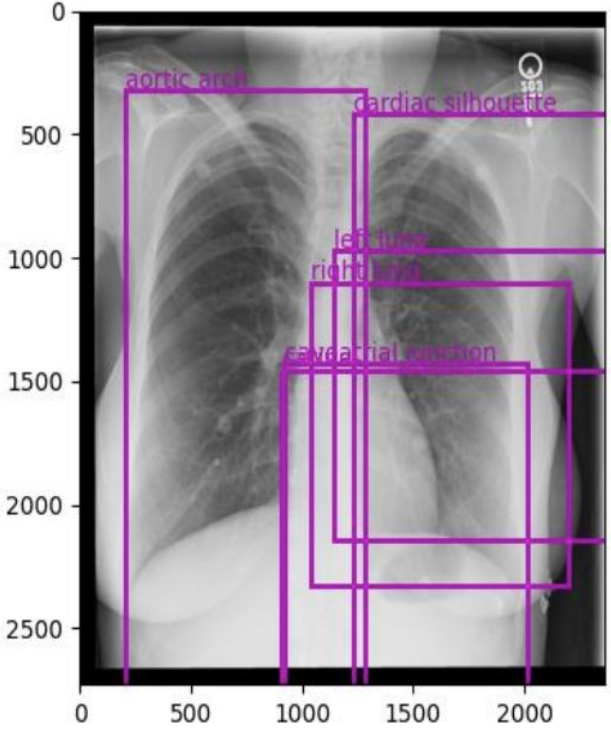
| | Faster_rcnn | | Yolo | | SSD | |
|---|---|---|---|---|---|---|
| | Precision @[IoU = 0.50] | Precision @[IoU = 0.50 : 0.95] | Precision @[IoU = 0.50] | Precision @[IoU = 0.50 : 0.95] | Precision @[IoU = 0.50] | Precision @[IoU = 0.50 : 0.95] |
| left hemidiaphragm | 76% | 43,1% | 63,1% | 25,7% | 72,8% | 35,3% |
| right hemidiaphragm | 89,5% | 57,2% | 86,4% | 39,3% | 85,6% | 46,9% |
| left lung | 100% | 78,1% | 100% | 62,8% | 100% | 61,3% |
| right lung | 99% | 80,0% | 95,7% | 62% | 99,8% | 65,6% |
| left hilar structures | 94,3% | 54,9% | 92,2% | 51% | 97,3% | 53,9% |
| right hilar structures | 92,5% | 56% | 92,5% | 49,9% | 96% | 52,8% |
| left apical zone | 96,7% | 64,9% | 97,6% | 45,1% | 97,6% | 61,5% |
| right apical zone | 95,6% | 58,8% | 89% | 43,6% | 92,5% | 56,3% |
| left clavicle | 85% | 53,5% | 76,9% | 29,4% | 85,7% | 46% |
| right clavicle | 86,9% | 51,1% | 75,1% | 26,4% | 85,1% | 46,9% |
| left costophrenic angle | 62,1% | 28,8% | 66,4% | 32,3% | 65,5% | 29,3% |
| right costophrenic angle | 78,6% | 37,4% | 82,3% | 40,3% | 81,2% | 39,5% |

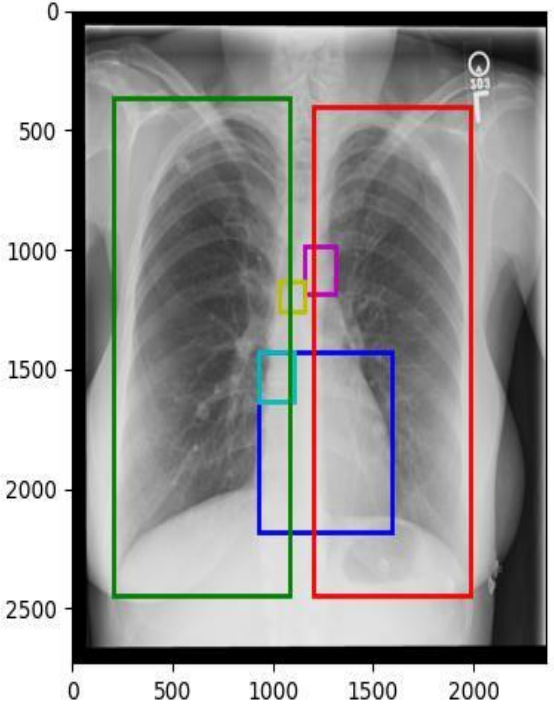
# Big & Small organs

TECHNISCHE UNIVERSITÄT DARMSTADT

| | Faster_rcnn | | Yolo | | SSD | |
|---|---|---|---|---|---|---|
| | Precision @[IoU = 0.50] | Precision @[IoU = 0.50 : 0.95] | Precision @[IoU = 0.50] | Precision @[IoU = 0.50 : 0.95] | Precision @[IoU = 0.50] | Precision @[IoU = 0.50 : 0.95] |
| left lung | 100% | 78,1% | 100% | 62,8% | 100% | 61,3% |
| right lung | 99% | 80,0% | 95,7% | 62% | 99,8% | 65,6% |
| cardiac silhouette | 98,9% | 63% | 96,1% | 53,3% | 96,8% | 54,4% |
| left costophrenic angle | 62,1% | 28,8% | 66,4% | 32,3% | 65,5% | 29,3% |
| carina | 56,1% | 16,8% | 44,1% | 11,9% | 38,8% | 11,2% |
| cavoatrial junction | 53,7% | 16,1% | 57,2% | 17,2% | 59,8% | 18,4% |

# Predictions vs. Ground Truth




Left lung

Right lung

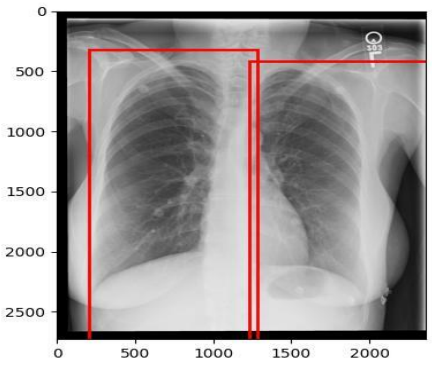
Ardiac silhouette

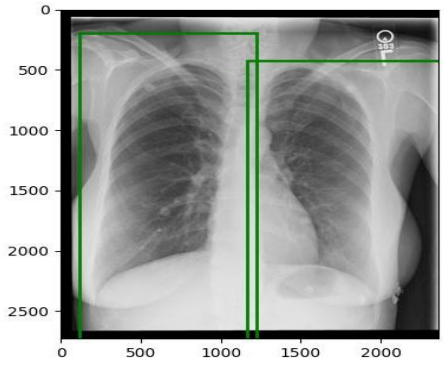
Cavoatrial junction

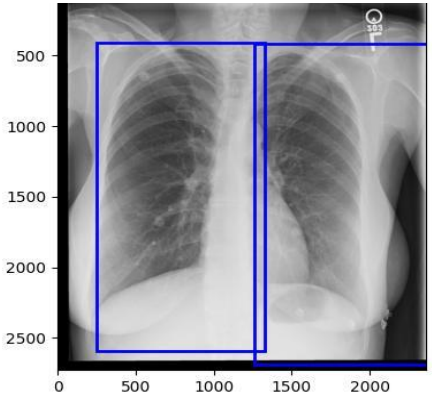
Aortic arch

Carina

# Left & Right Lung



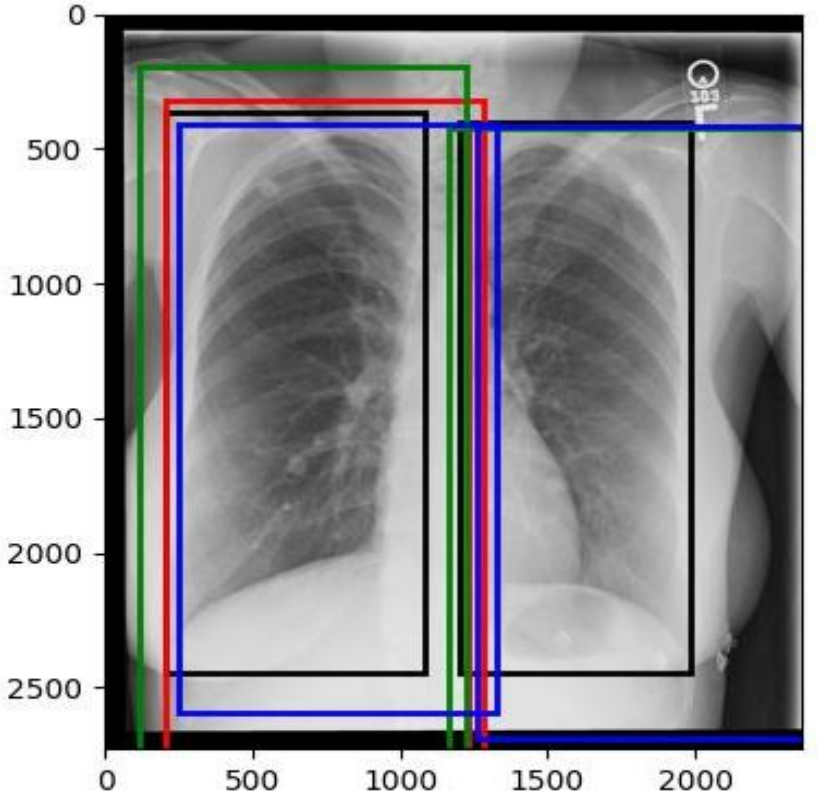




Ground Truth

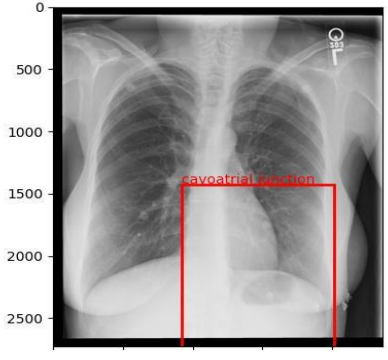
Faster RCNN

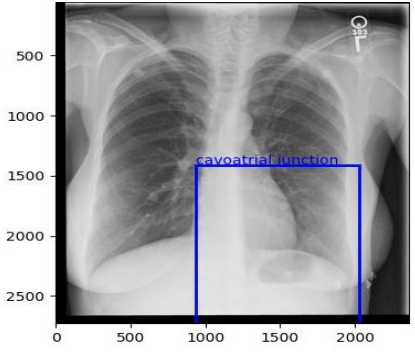
YOLO

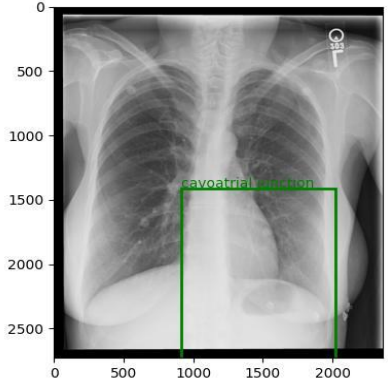
SSD



# Cavoatrial junction



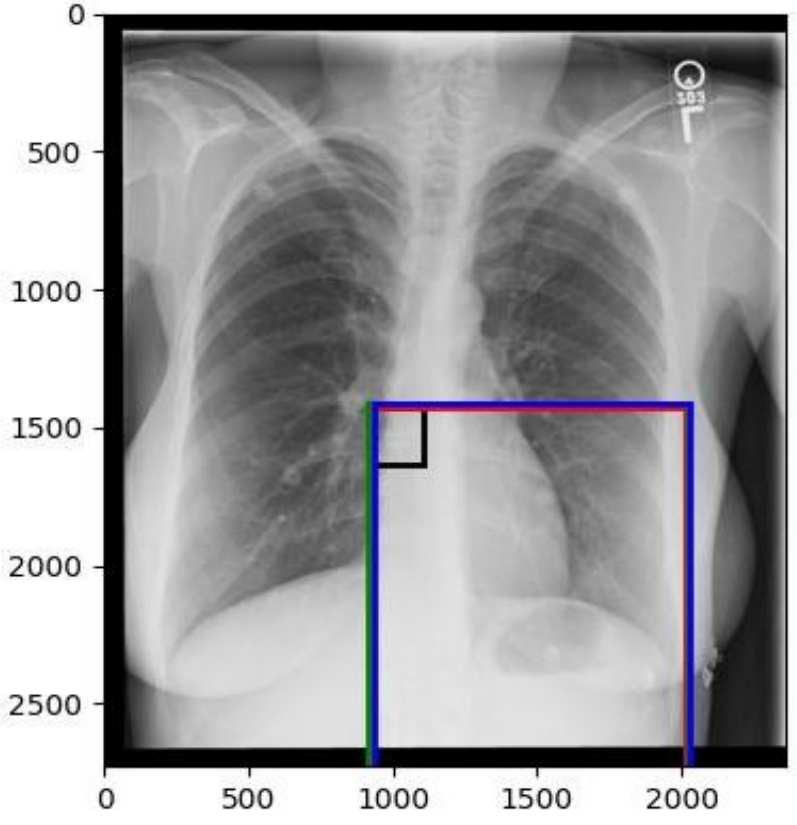



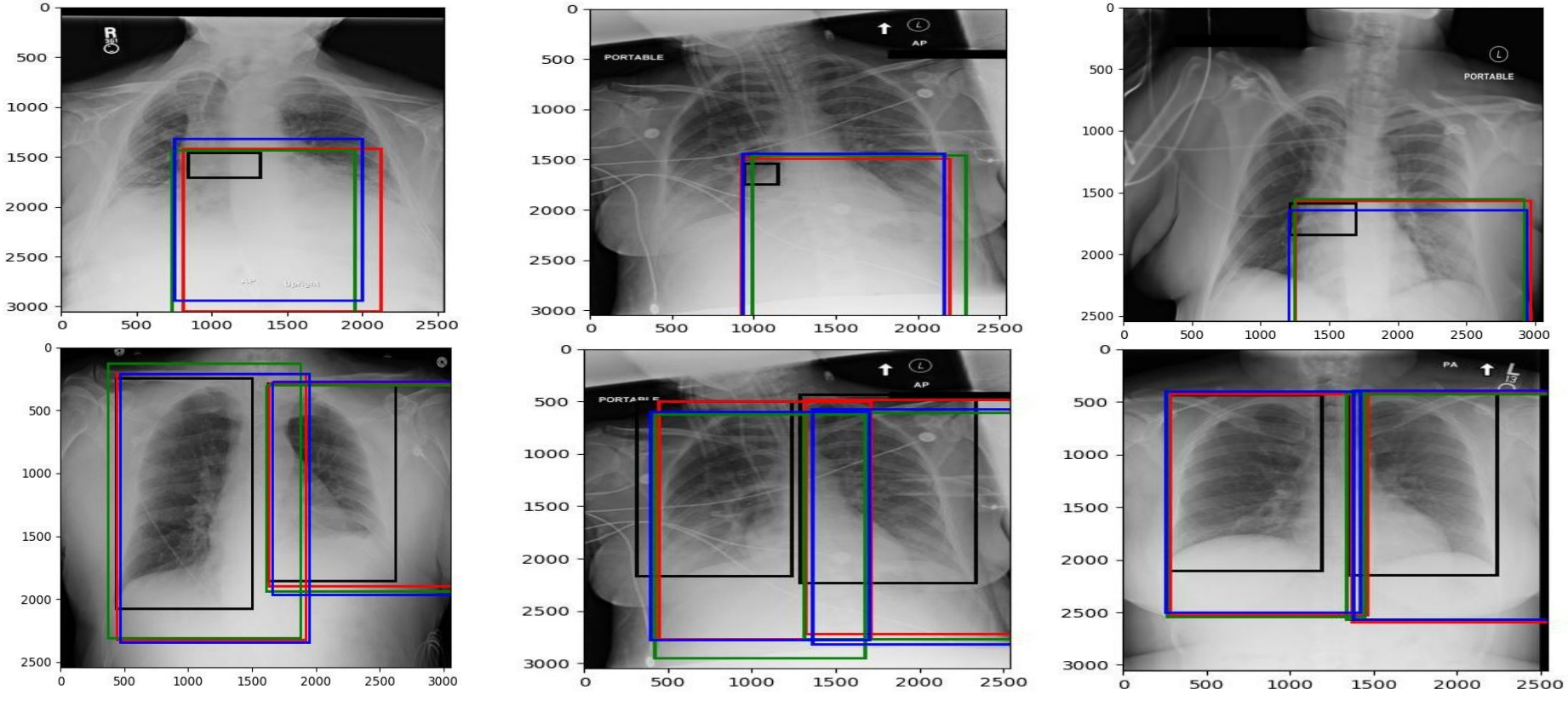

Ground Truth

Faster RCNN

YOLO

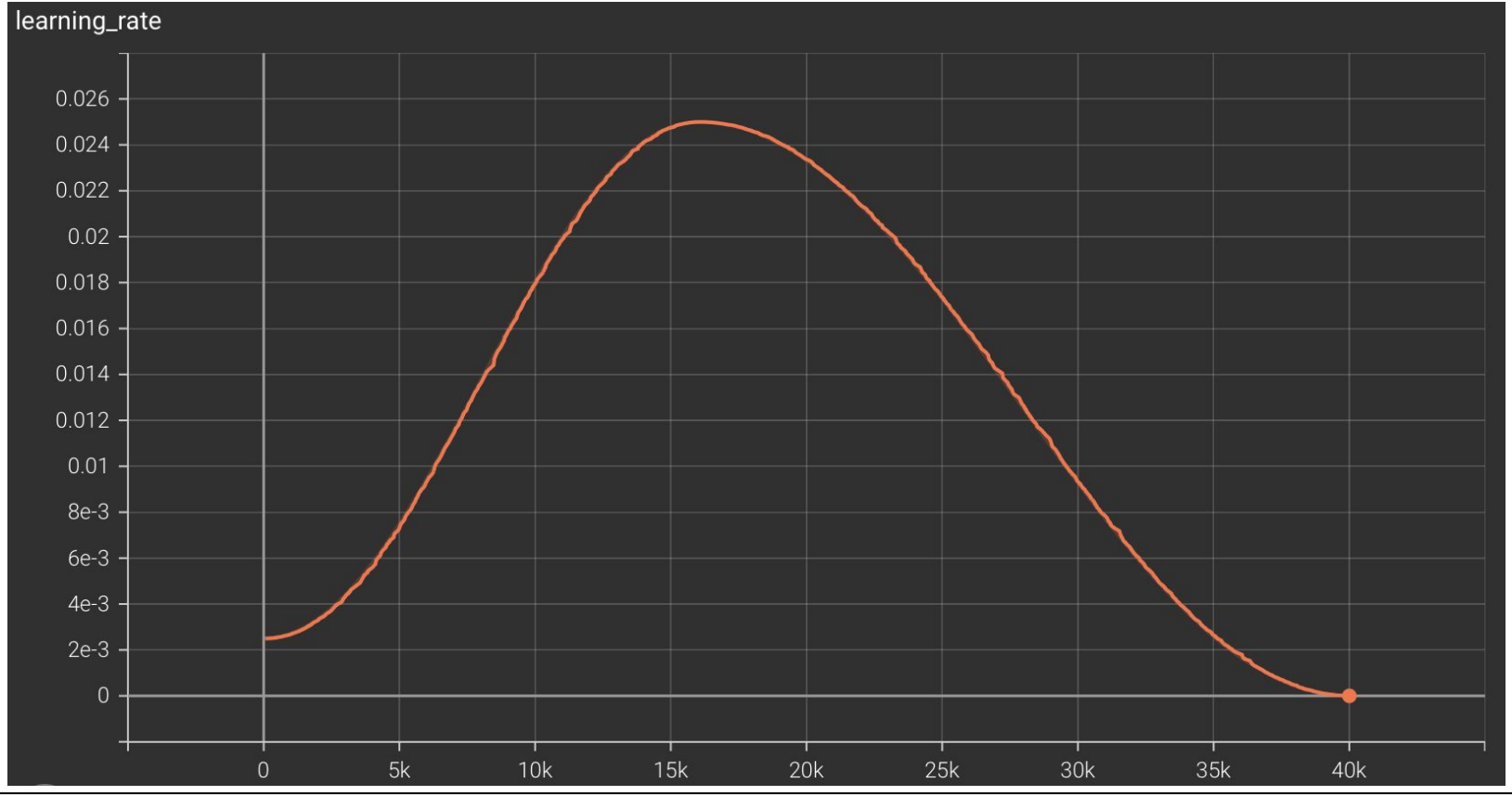
SSD

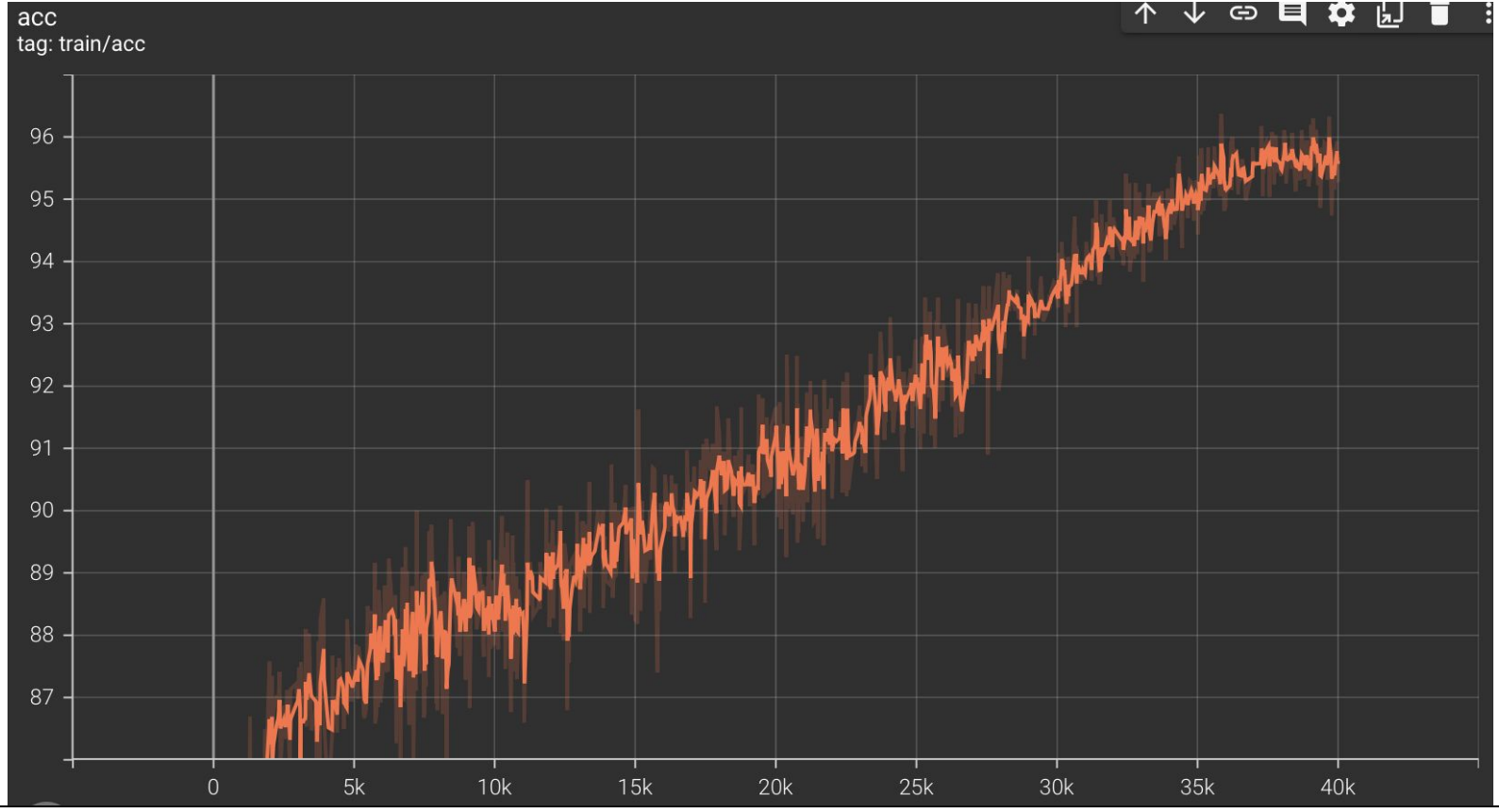


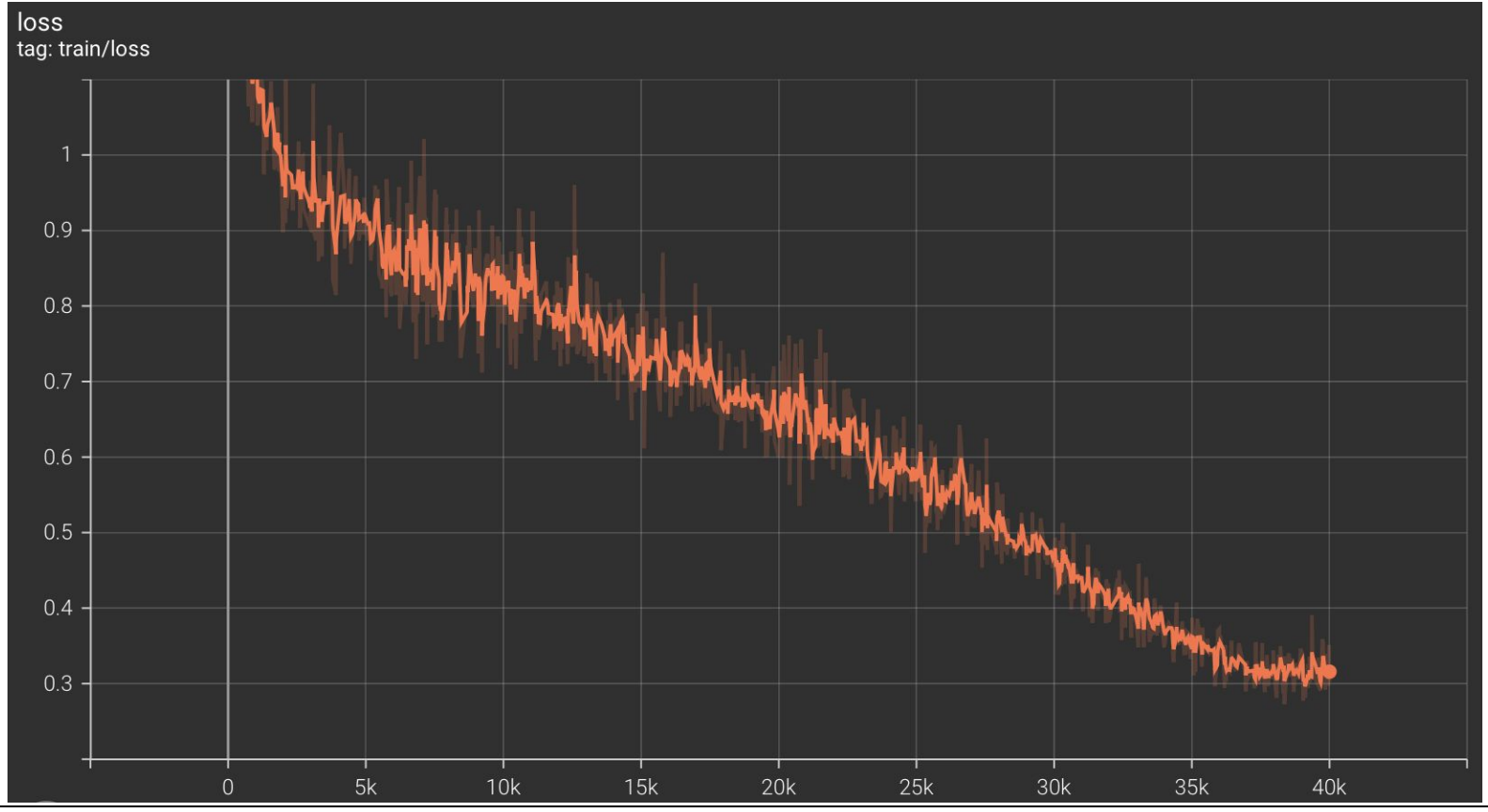

PORTABLE
AP
PORTABLE
PORTABLE
AP
PA

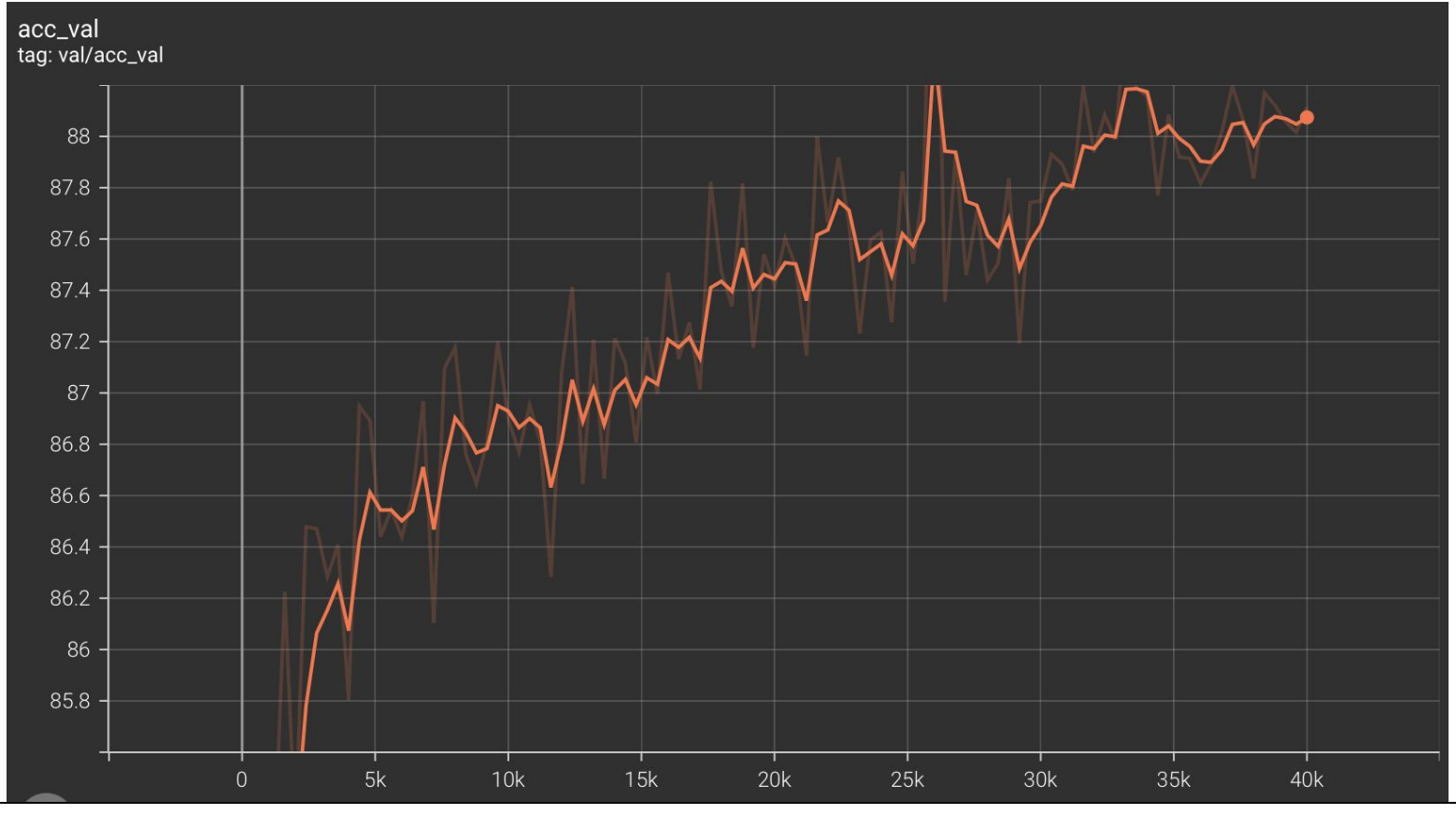
TECHNISCHE UNIVERSITÄT DARMSTADT

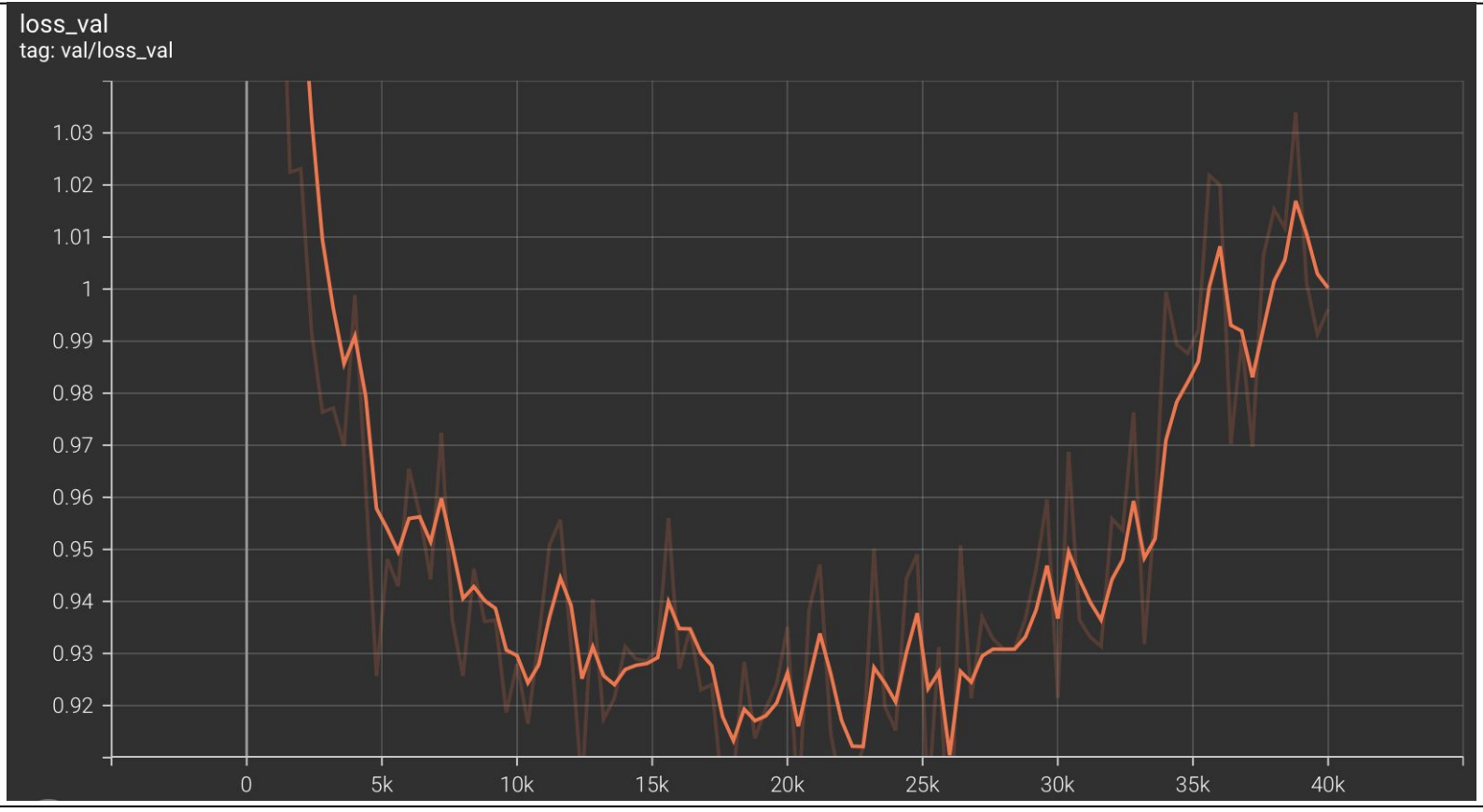
# Thanks!


learning_rate
0.026
0.024
0.022
0.02
0.018
0.016
0.014
0.012
0.01
8e-3
6e-3
4e-3
2e-3
0
0
5k
10k
15k
20k
25k
30k
35k
40k


acc
tag: train/acc
96
95
94
93
92
91
90
89
88
87
0
5k
10k
15k
20k
25k
30k
35k
40k


loss
tag: train/loss
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
5k
10k
15k
20k
25k
30k
35k
40k


acc_val
tag: val/acc_val
88
87.8
87.6
87.4
87.2
87
86.8
86.6
86.4
86.2
86
85.8
0
5k
10k
15k
20k
25k
30k
35k
40k


loss_val
tag: val/loss_val
1.03
1.02
1.01
1
0.99
0.98
0.97
0.96
0.95
0.94
0.93
0.92
0
5k
10k
15k
20k
25k
30k
35k
40k

```
 Average Precision  (AP) @[ IoU=0.50:0.50 | area=   all | maxDets=100 ] = 0.875
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=1000 ] = -1.000
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=1000 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.50 | area= small | maxDets=1000 ] = 0.000
 Average Precision  (AP) @[ IoU=0.50:0.50 | area=medium | maxDets=1000 ] = 0.078
 Average Precision  (AP) @[ IoU=0.50:0.50 | area= large | maxDets=1000 ] = 0.878
 Average Recall     (AR) @[ IoU=0.50:0.50 | area=   all | maxDets=100 ] = 0.900
 Average Recall     (AR) @[ IoU=0.50:0.50 | area=   all | maxDets=300 ] = 0.900
 Average Recall     (AR) @[ IoU=0.50:0.50 | area=   all | maxDets=1000 ] = 0.900
 Average Recall     (AR) @[ IoU=0.50:0.50 | area= small | maxDets=1000 ] = 0.000
 Average Recall     (AR) @[ IoU=0.50:0.50 | area=medium | maxDets=1000 ] = 0.090
 Average Recall     (AR) @[ IoU=0.50:0.50 | area= large | maxDets=1000 ] = 0.904
```

| category | AP | category | AP | category | AP |
|---|---|---|---|---|---|
| right lung | 1.000 | left lung | 1.000 | left costophrenic angle | 0.656 |
| trachea | 0.927 | right costophrenic angle | 0.776 | cavoatrial junction | 0.524 |
| left mid lung zone | 0.900 | left clavicle | 0.856 | right hemidiaphragm | 0.881 |
| aortic arch | 0.895 | left hemidiaphragm | 0.799 | right apical zone | 0.940 |
| carina | 0.470 | left hilar structures | 0.966 | right hilar structures | 0.933 |
| right clavicle | 0.909 | svc | 0.870 | upper mediastinum | 0.973 |
| right atrium | 0.834 | right mid lung zone | 0.907 | left apical zone | 0.968 |
| left upper lung zone | 0.999 | right upper lung zone | 0.969 | left lower lung zone | 0.897 |
| right lower lung zone | 0.917 | cardiac silhouette | 0.977 | None | None |

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.550
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=1000 ] = 0.875
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=1000 ] = 0.586
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=1000 ] = 0.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=1000 ] = 0.023
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=1000 ] = 0.551
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.612
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=300 ] = 0.612
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=1000 ] = 0.612
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=1000 ] = 0.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=1000 ] = 0.031
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=1000 ] = 0.614
```

| category | AP | category | AP | category | AP |
|---|---|---|---|---|---|
| right lung | 0.822 | left lung | 0.800 | left costophrenic angle | 0.324 |
| trachea | 0.506 | right costophrenic angle | 0.391 | cavoatrial junction | 0.172 |
| left mid lung zone | 0.584 | left clavicle | 0.549 | right hemidiaphragm | 0.585 |
| aortic arch | 0.515 | left hemidiaphragm | 0.435 | right apical zone | 0.631 |
| carina | 0.156 | left hilar structures | 0.563 | right hilar structures | 0.584 |
| right clavicle | 0.549 | svc | 0.472 | upper mediastinum | 0.601 |
| right atrium | 0.381 | right mid lung zone | 0.569 | left apical zone | 0.658 |
| left upper lung zone | 0.787 | right upper lung zone | 0.748 | left lower lung zone | 0.615 |
| right lower lung zone | 0.653 | cardiac silhouette | 0.642 | None | None |