COMP- Assignment 4

Fps with nano Modell:

```
2.8834639758120066
2.722536280002363
2.801203749093381
2.8368183404530463
2.8893395644123094
2.8386076806046594
2.8258857708031835
2.785330592918973
2.821750798395338
2.806476240691759
2.8812039454657987
2.8139870649169416
2.783533876151675
```

Fps with small modell:

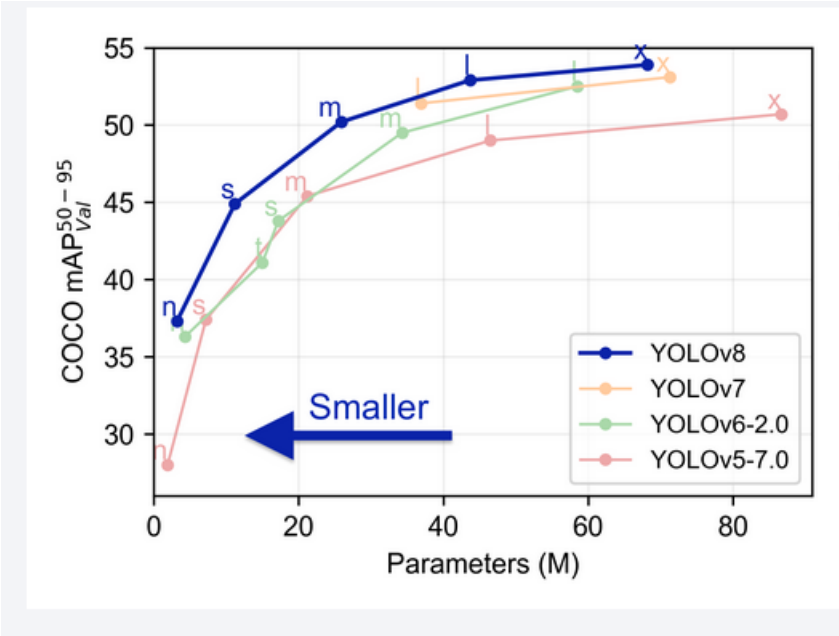
```
1.308897768887377
1.291029778641685
1.1116737693955652
1.6258964750374756
1.3433362030090599
1.3020980762072638
1.3634421786707955
1.3689303403996438
1.3755744673314905
1.5079434529833653
1.4437900701017017
1.3500386088207073
```

Fps with medium model:

For the medium the fps dropped to around 0.8 fps and were therefore way lower than for the nano model due to the higher number of parameters.

```
0.9911570340700305
1.0910437884204787
1.0772057038967415
0.9626743974941984
1.0437078854950796
1.1190764153196167
1.0330256072772537
0.8403754927882015
0.8577833286738801
0.796288206580995
0.8447502296510814
0.8273190375706593
```

Number of parameters:



Model	size (pixels)	mAP^{val} 50-95	mAP^{val} 50	Speed CPU b1 (ms)	Speed V100 b1 (ms)	Speed V100 b32 (ms)	params (M)	FLOPs @640 (B)
YOLOv5n	640	28.0	45.7	**45**	**6.3**	**0.6**	**1.9**	**4.5**
YOLOv5s	640	37.4	56.8	98	6.4	0.9	7.2	16.5
YOLOv5m	640	45.4	64.1	224	8.2	1.7	21.2	49.0

I was playing a video of street traffic on my phone, there the object recognition worked quite well for people and cars using the nano model.

When moving the camera around the room, the results were right around half of the time. Humans worked quite well but aside from that a lot of stuff got wrongly classified as smart phone, even though it was a cup or a water bottle. The results also depend a lot on the light conditions.

For Humans, I would estimate the precision around 60 % while the recall might be around 90 %. Some objects got classified as humans while hardly any humans got misclassified as something else.

The overall recall should be way lower. The overall precision might be at 50 %.