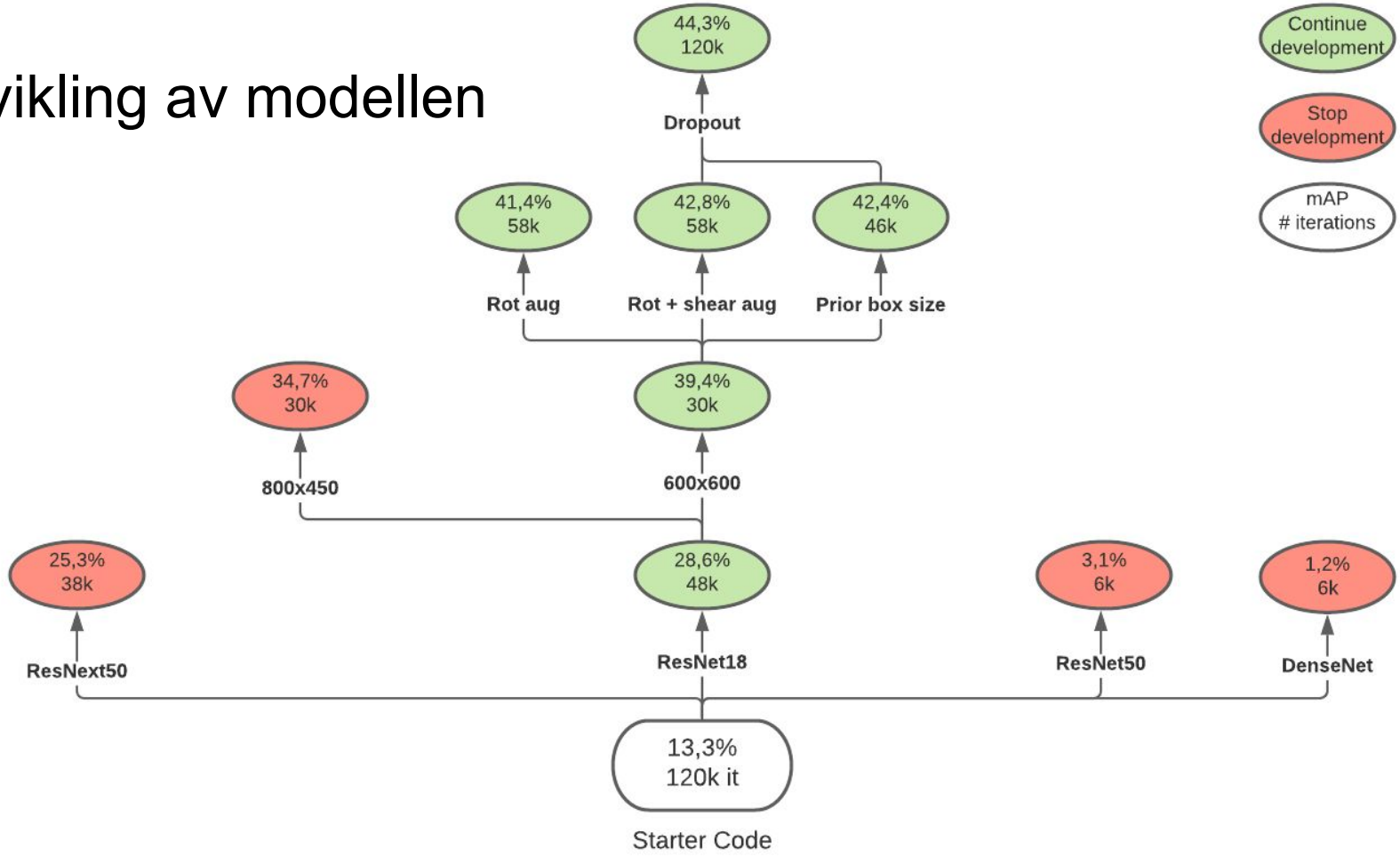
Datasyn og dyp læring

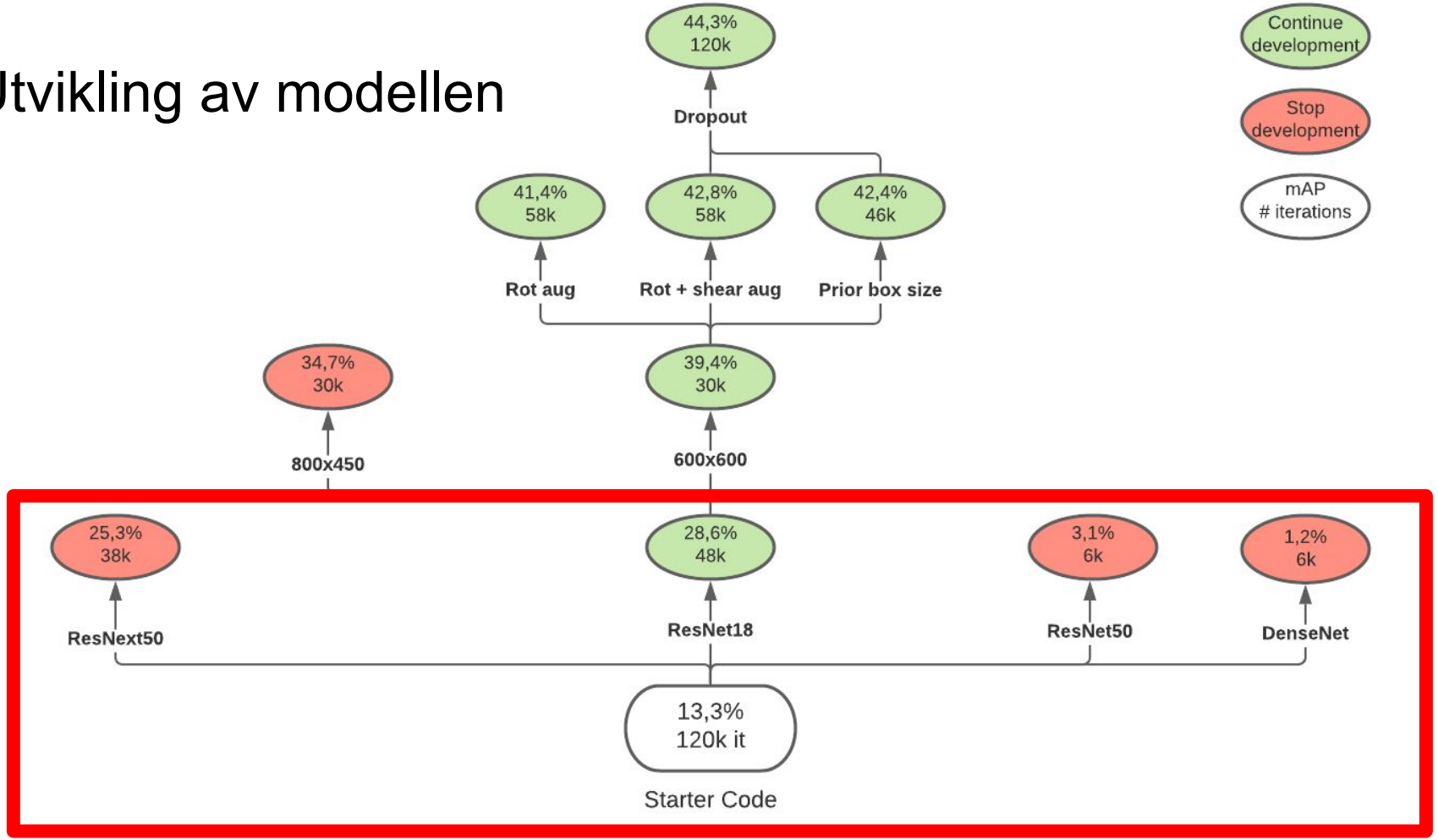
RDD

Gruppe 120

Utvikling av modellen



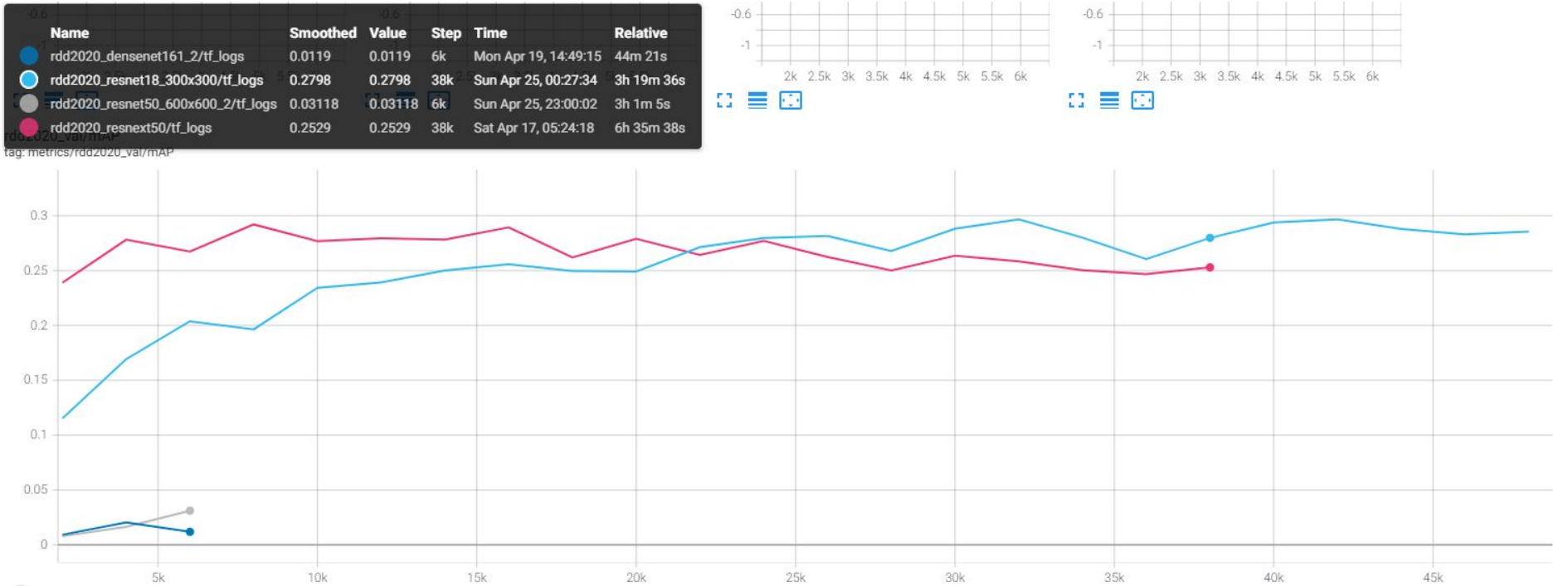
Utvikling av modellen



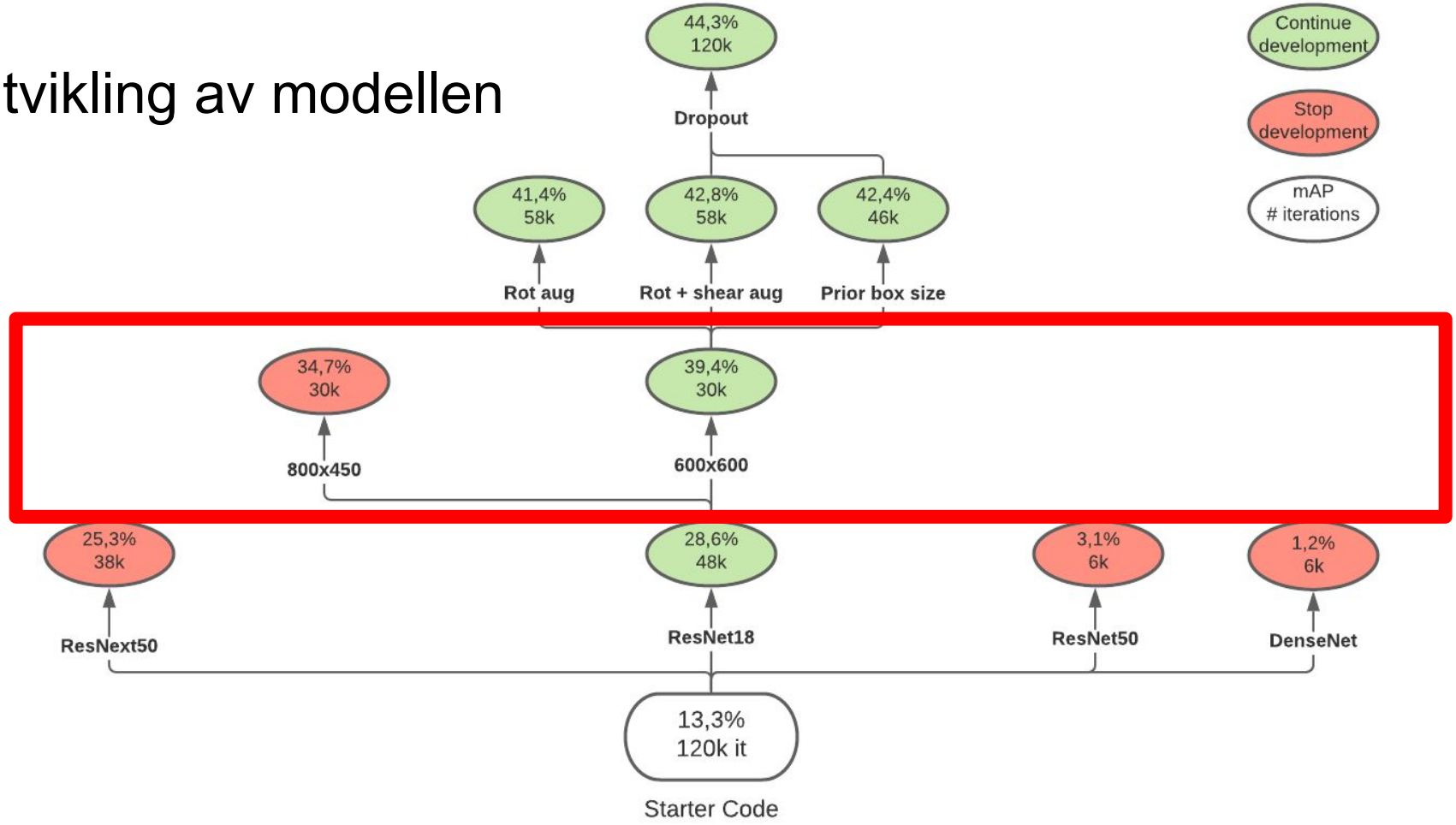
Utvikling av modellen

- **Backbone: residual networks (fra torchvision)**
  - Unngå accuracy saturation selv om vi bruker dype nettverk.
  - Prøvde ulike antall lag: ResNet18, ResNeXt50, ResNet50

Backbones


Name Smoothed Value Step Time Relative
rdd2020_densenet161_2/tf_logs 0.0119 0.0119 6k Mon Apr 19, 14:49:15 44m 21s
rdd2020_resnet18_300x300/tf_logs 0.2798 0.2798 38k Sun Apr 25, 00:27:34 3h 19m 36s
rdd2020_resnet50_600x600_2/tf_logs 0.03118 0.03118 6k Sun Apr 25, 23:00:02 3h 1m 5s
rdd2020_resnext50/tf_logs 0.2529 0.2529 38k Sat Apr 17, 05:24:18 6h 35m 38s
tag: metrics/rdd2020_val/mAP
0.3
0.25
0.2
0.15
0.1
0.05
0
5k
10k
15k
20k
25k
30k
35k
40k
45k

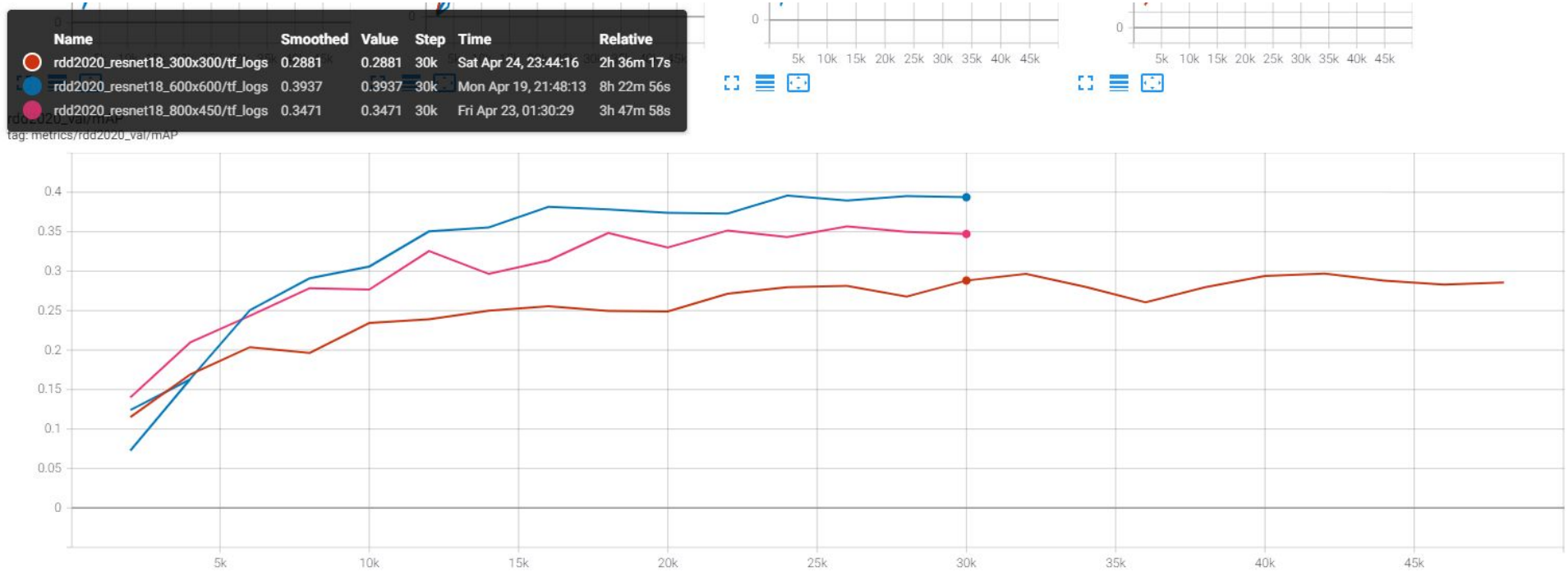
Utvikling av modellen



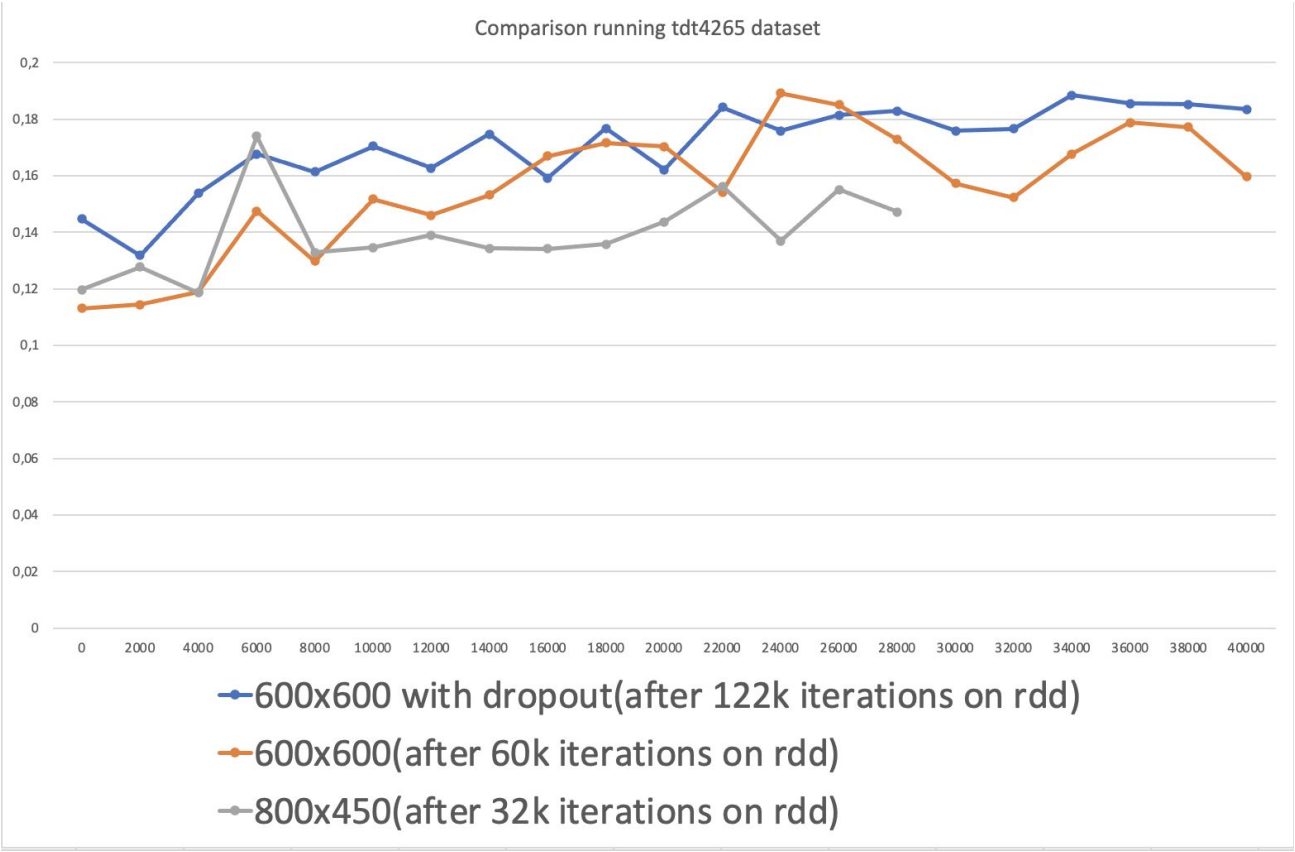
Utvikling av modellen

- Backbone: residual networks (fra torchvision)
  - Unngå accuracy saturation selv om vi bruker dype nettverk.
  - Prøvde ulike antall lag: ResNet18, ResNeXt50, ResNet50
- **Inputdimensjoner**
  - Høyere oppløsning på input gir modellen mer informasjon å basere prediksjoner på.

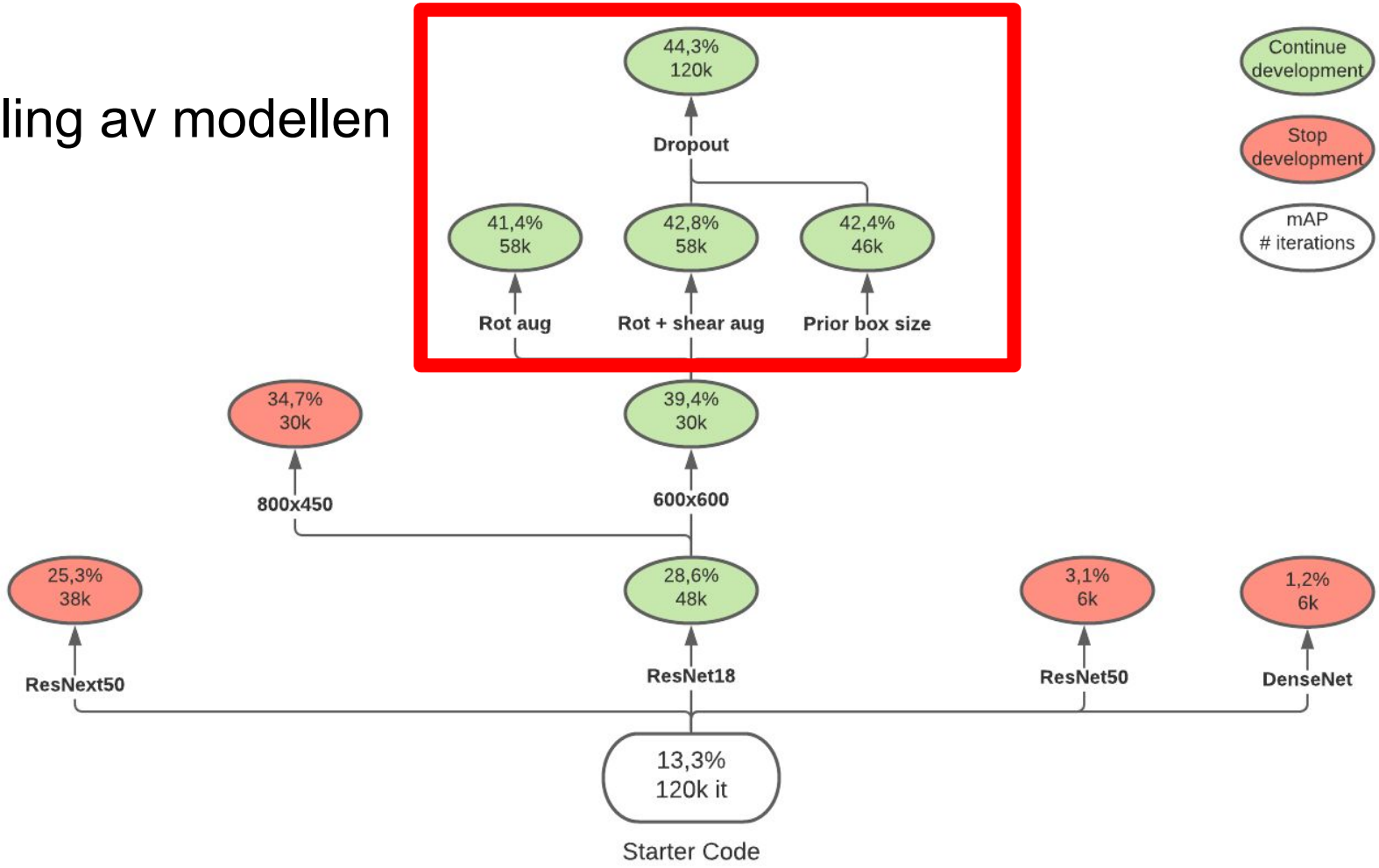
Input-oppløsning RDD



Input-oppløsning TDT4265



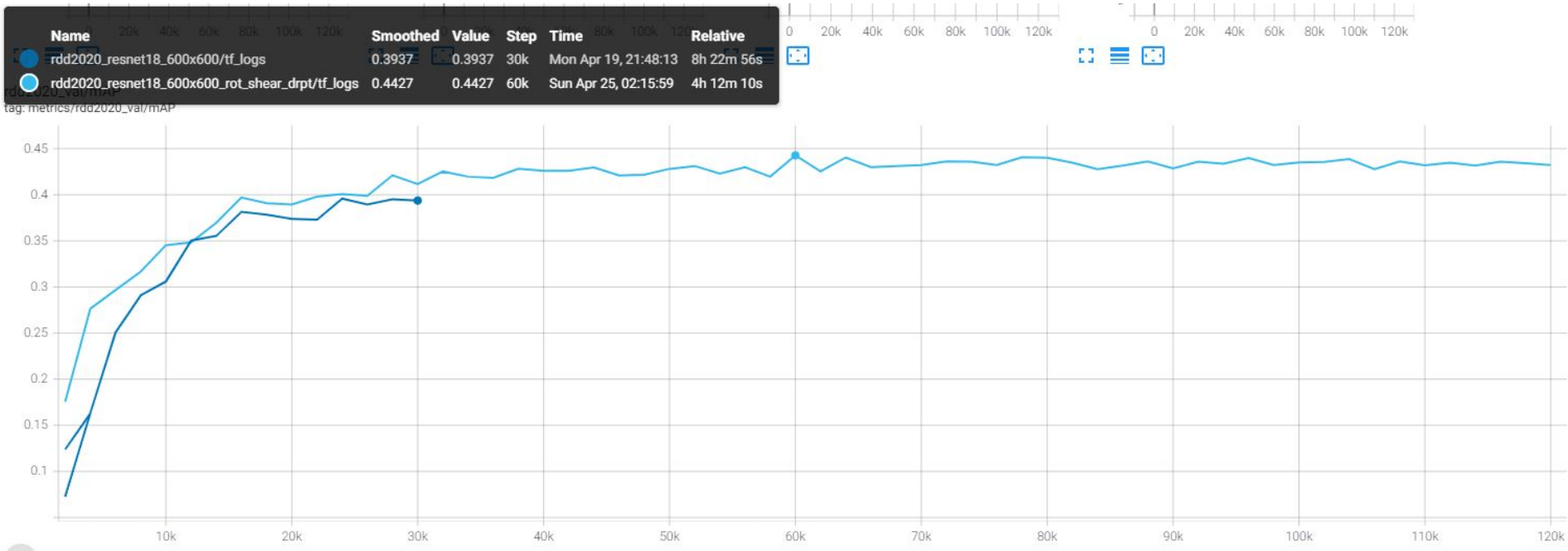
Utvikling av modellen



Utvikling av modellen

- Backbone: residual networks (fra torchvision)
  - Unngå accuracy saturation selv om vi bruker dype nettverk.
  - Prøvde ulike antall lag: ResNet18, ResNeXt50, ResNet50
- Inputdimensjoner
  - Høyere oppløsning på input gir modellen mer informasjon å basere prediksjoner på.
- **Siste endringer**
  - Prior Box sizes
  - 10% dropout i de fire siste lagene
  - Input Normalization
  - Data Augmentation
    - RandomMirror og RandomSampleCrop
    - Random rotation (+/- 10 grader)
    - Random shear

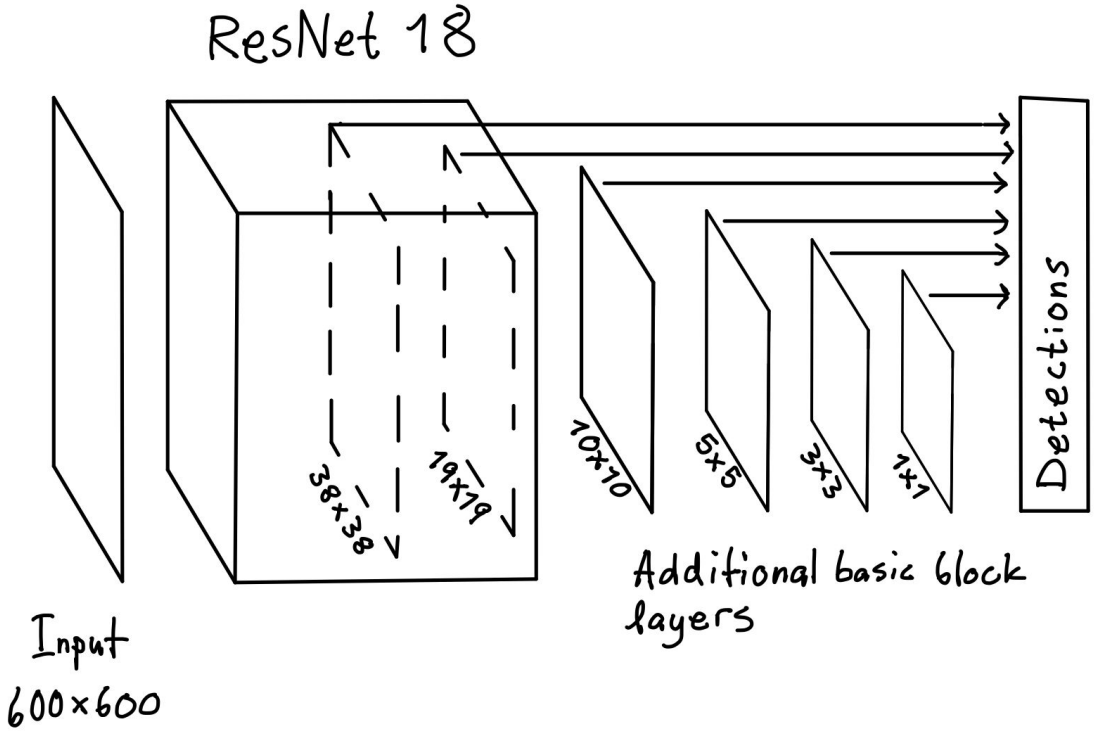
Siste endringer



Utvikling av modellen

- Input normalization
  - Raskere trening når inputs normaliseres til en standard skala.
- SSD: multi-scale feature maps for deteksjon
  - Seks feature maps med oppløsning på 38x38, 19x19, 10x10, 5x5, 3x3 og 1x1.
    - 38x38 og 19x19 fra ResNet18
    - Resten fra Basic Blocks lagt til ResNet18
  - Oppdager objekter i forskjellige størrelser.
- Dropout i de fire siste Basic Block lagene

Endelig modell



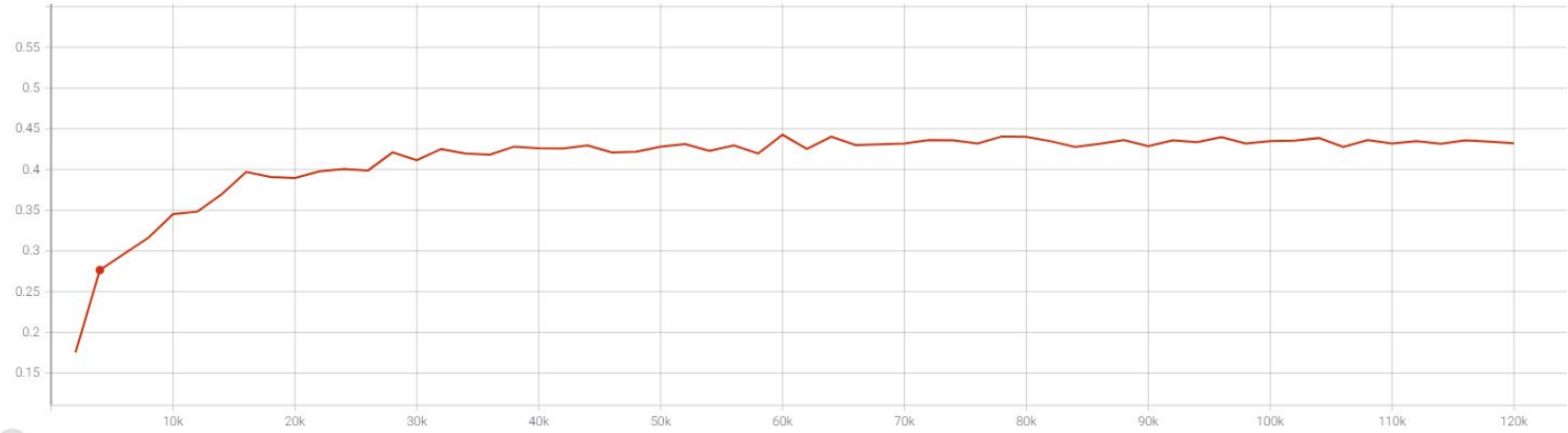
Trening av den endelige modellen

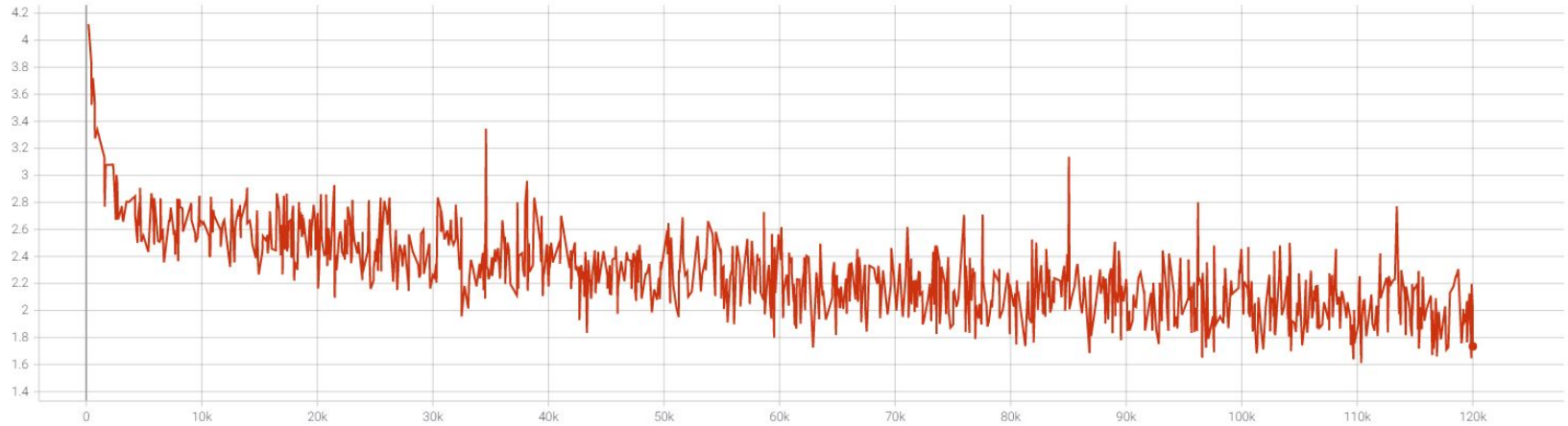
- Pre-train på ImageNet
- Totalt 8 timer trening på RDD2020
- 4 timer trening på TDT4265

Resultater

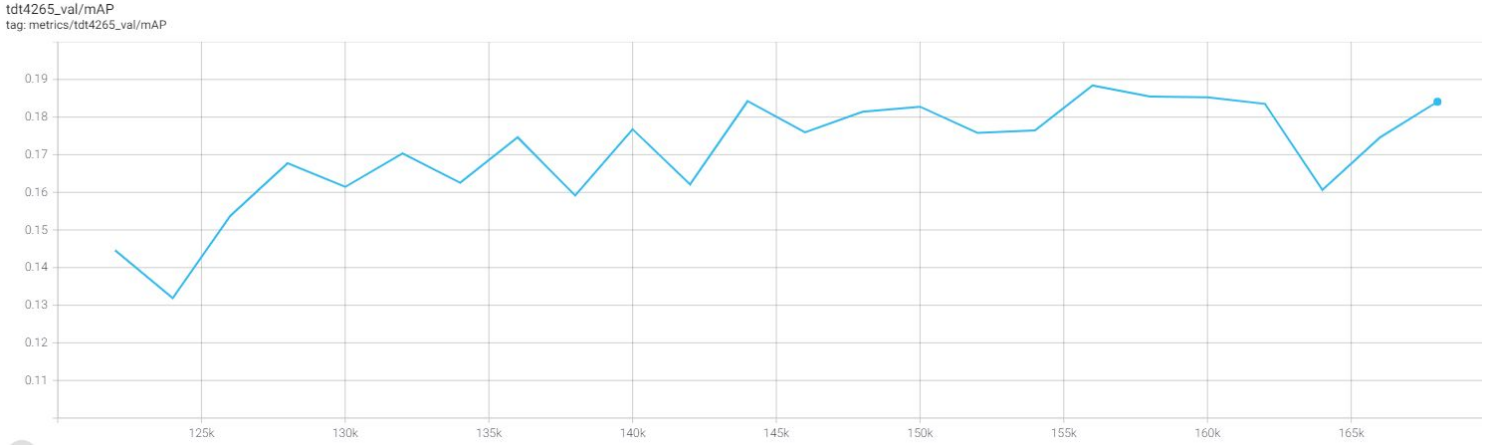
- RDD mAP: 44.27 %
- TDT4265 mAP: 18.84 %
- Leaderboard Full mAP: 0.031

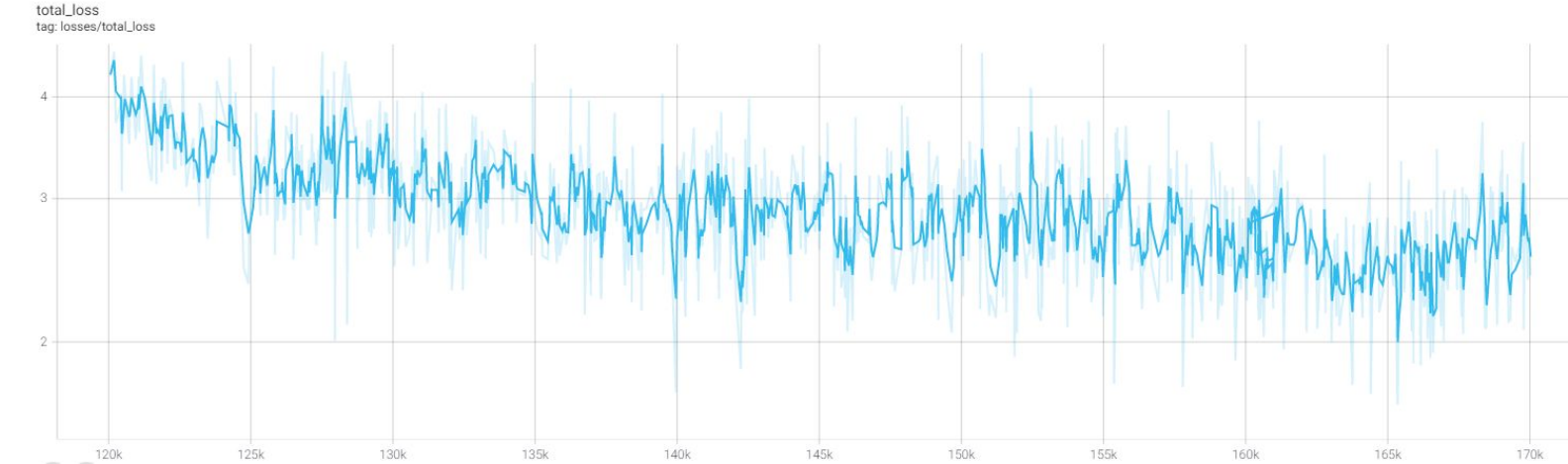
Resultater - Total loss og mAP på RDD2020 datasett





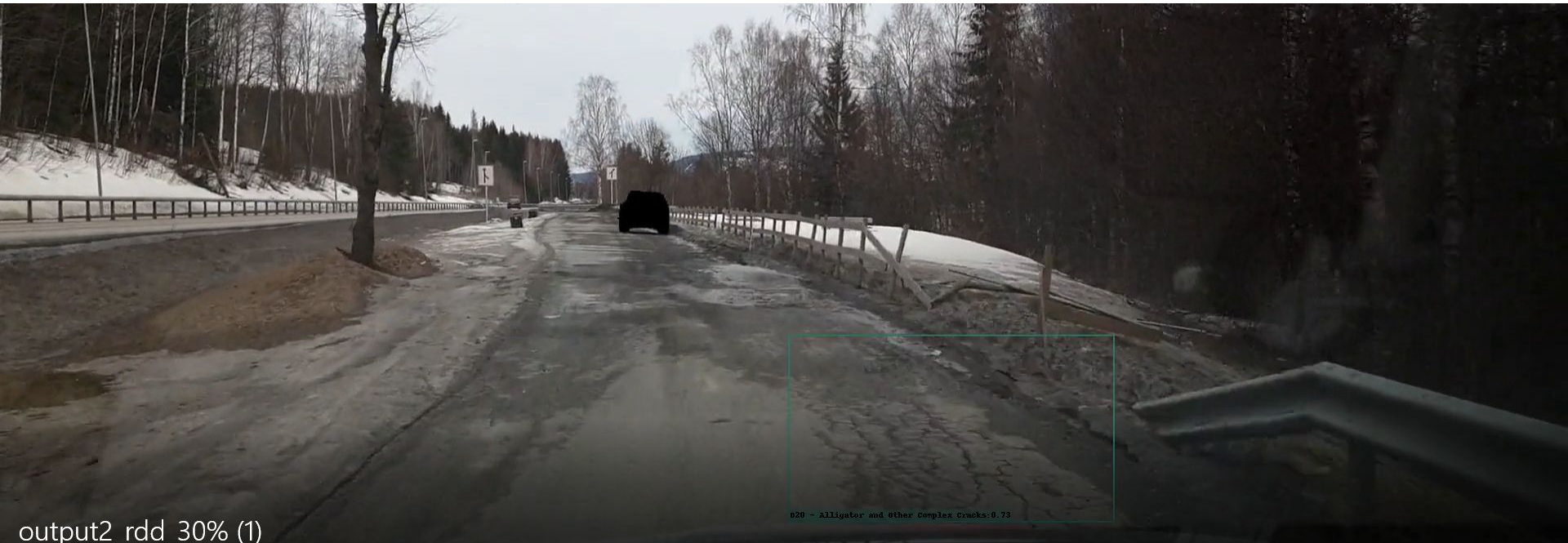
Resultater - Total loss og mAP på TDT4265 datasett





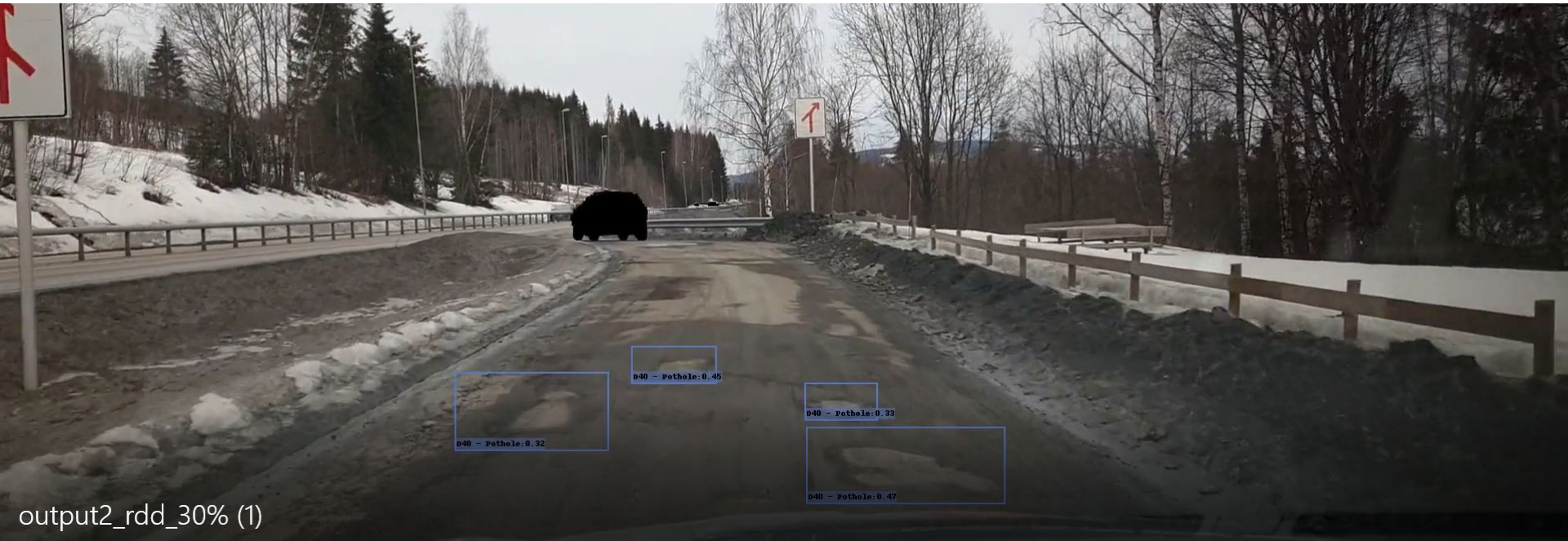
Resultater

AP D20 - Alligator Cracks and Complex Cracks: 61.18%



Resultater

AP D40 - Potholes: 50.58%



output2_rdd_30% (1)

Resultater - Video

Resultater - Video

Resultater


D20 - Alligator and Other Complex Cracks 0.35
D00 - Linear Longitudinal Crack 0.53
D00 - Linear Longitudinal Crack 0.34

Runtime Analysis

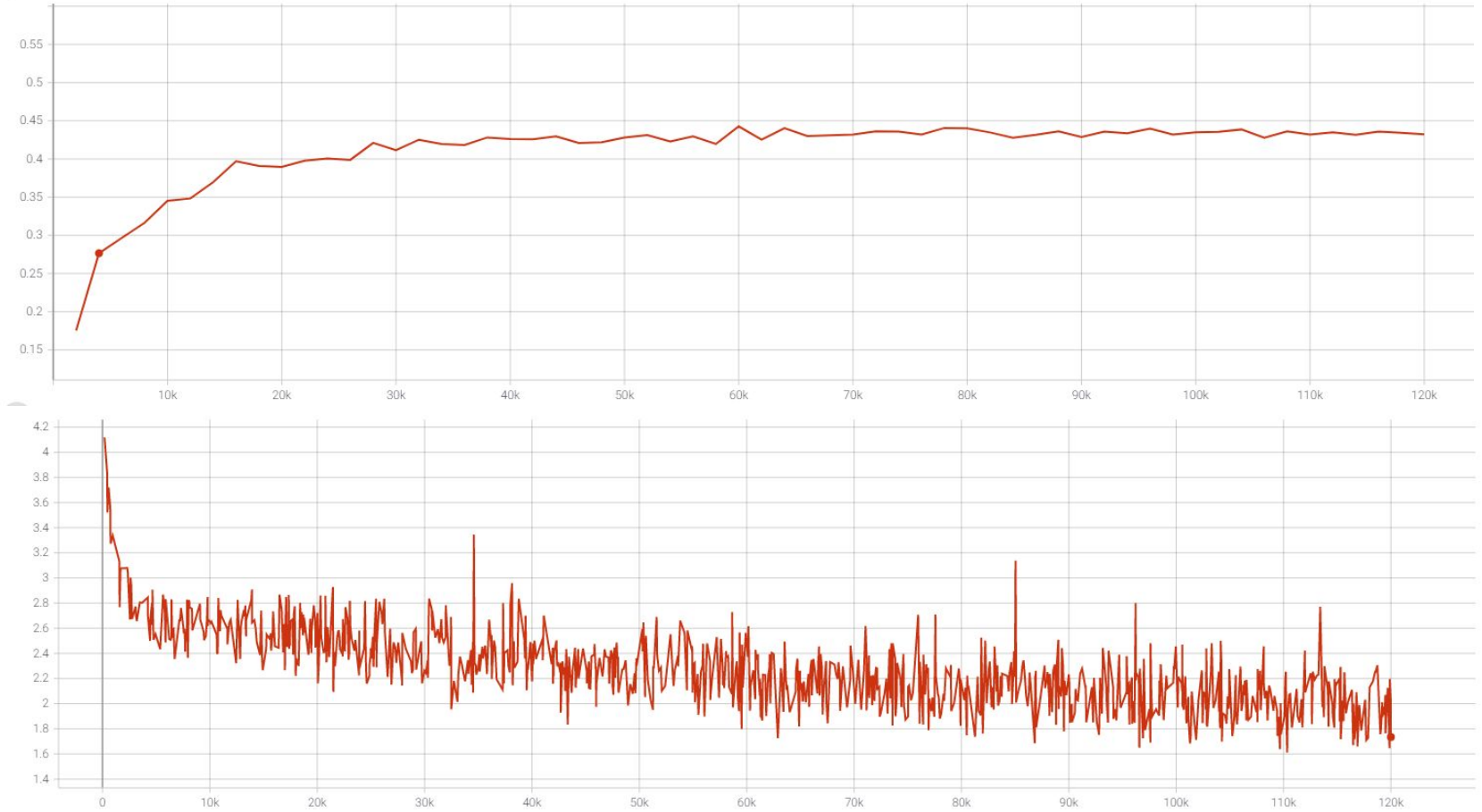
- Final model: **15-17 FPS**

- Efforts to improve runtime:
  - 18 layers istedet for 50 layers
    - ResneXt50: 9-10 FPS
  - Oppløsning på feature maps
  - If we had more time:
    - Pruning

Diskusjon - Hva vi ville gjort annerledes / videre arbeid

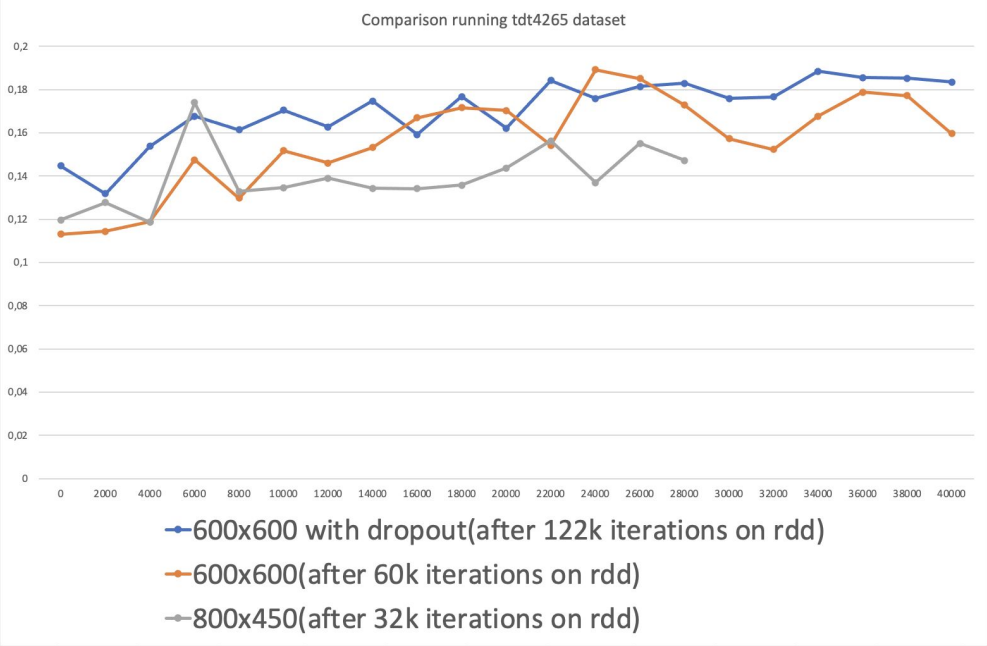
- mAP flater ut tidlig, mens loss synker
  - Mest sannsynlig overfitting
  - Eksperimentere mer med data augmentation og dropout
  - Motvirke overfitting -> dypere modeller
- Inception
- Fryse lag

RDD2020 Final Model total_loss and mAP



Diskusjon - uventede resultater/ metoder som ikke virket

- Dårlig ytelse med input-oppløsning 800x450
  - Dårlig trening på RDD2020 -> dårlig ytelse på TDT4265?



Gruppedynamikk