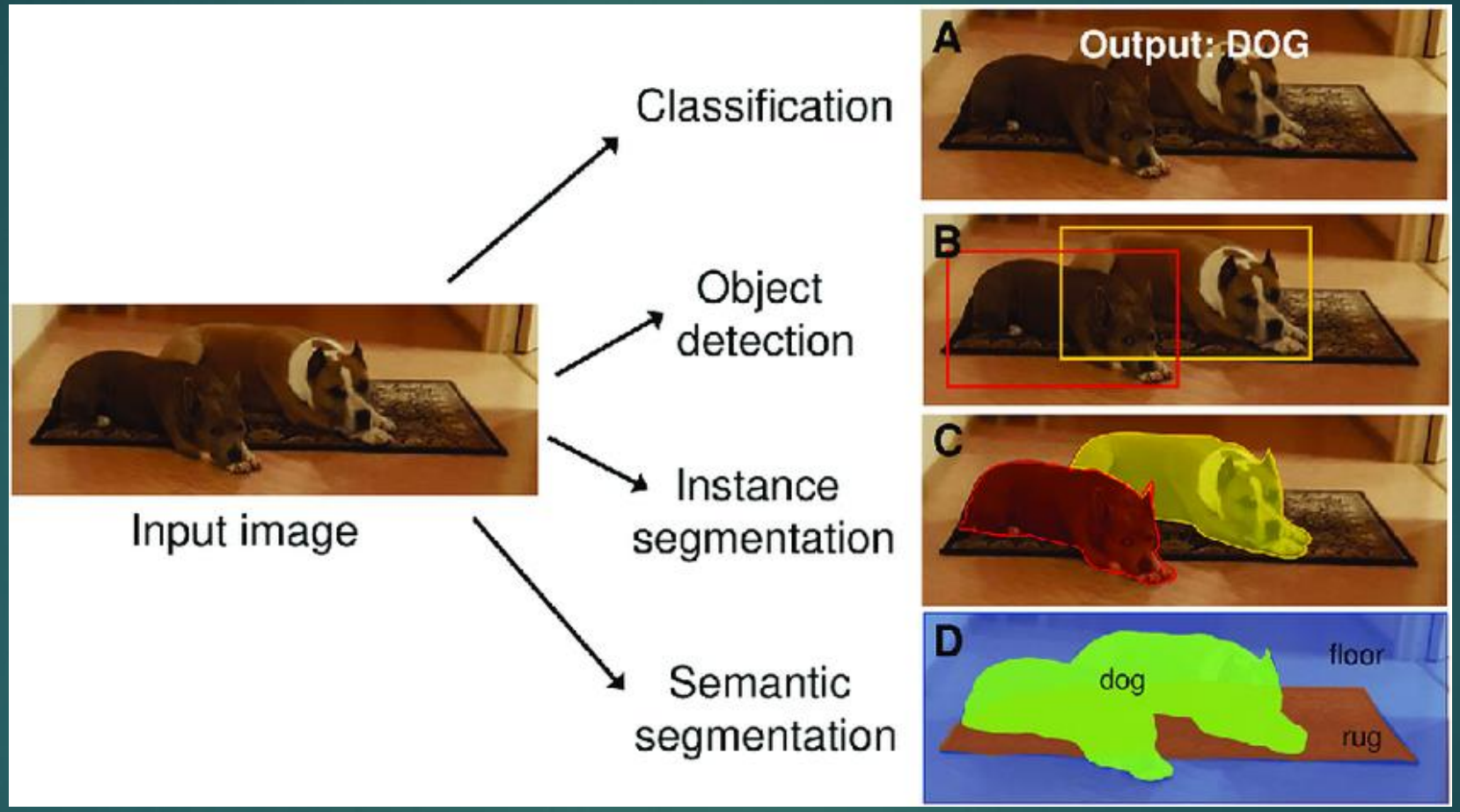
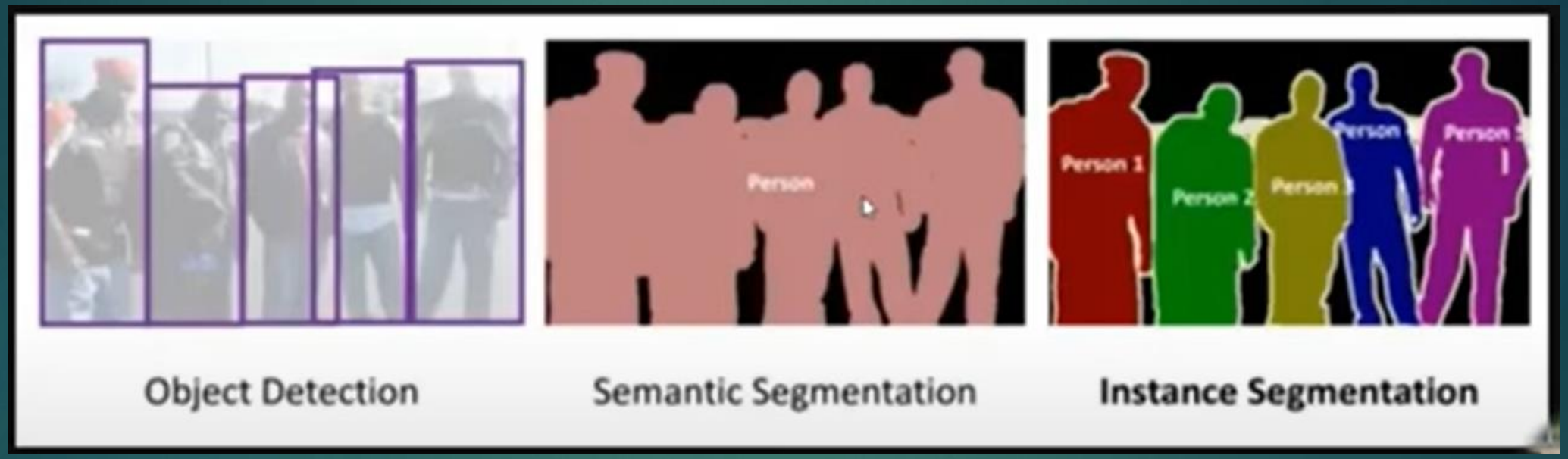
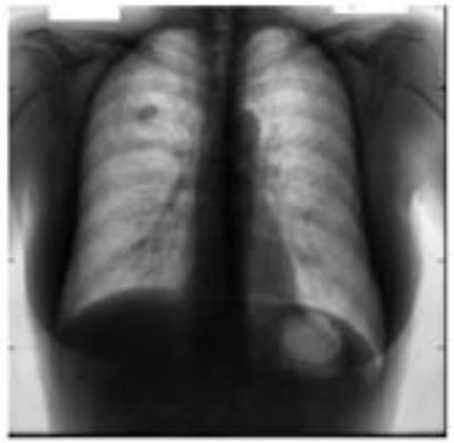
Image Segmentation With U-NET

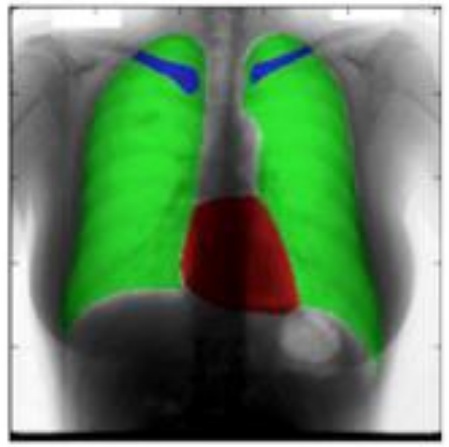
AMIN HEYDARI


Classification
Object detection
Instance segmentation
Semantic segmentation
Input image
A
Output: DOG
B
C
D
dog
floor
rug

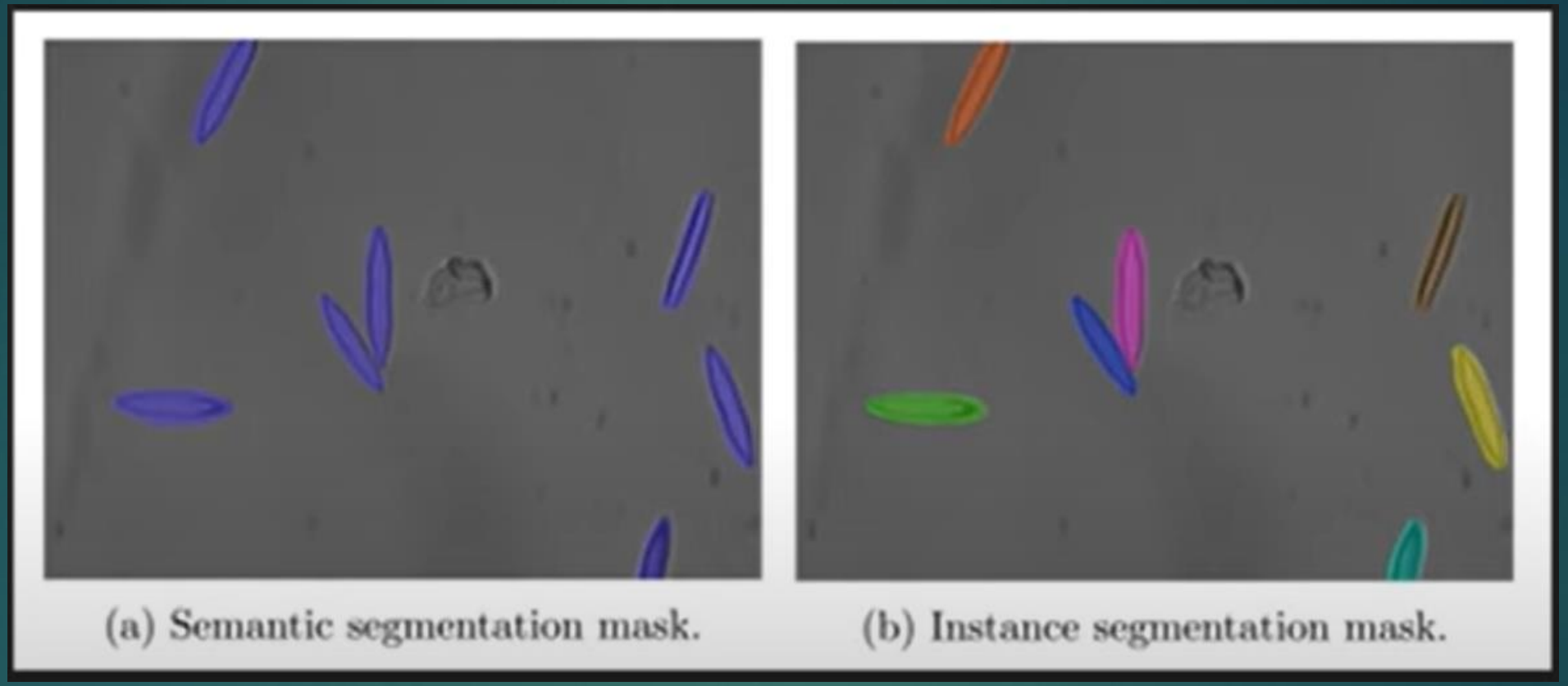

Person
Person 1
Person 2
Person
Object Detection
Semantic Segmentation
Instance Segmentation



Input Image



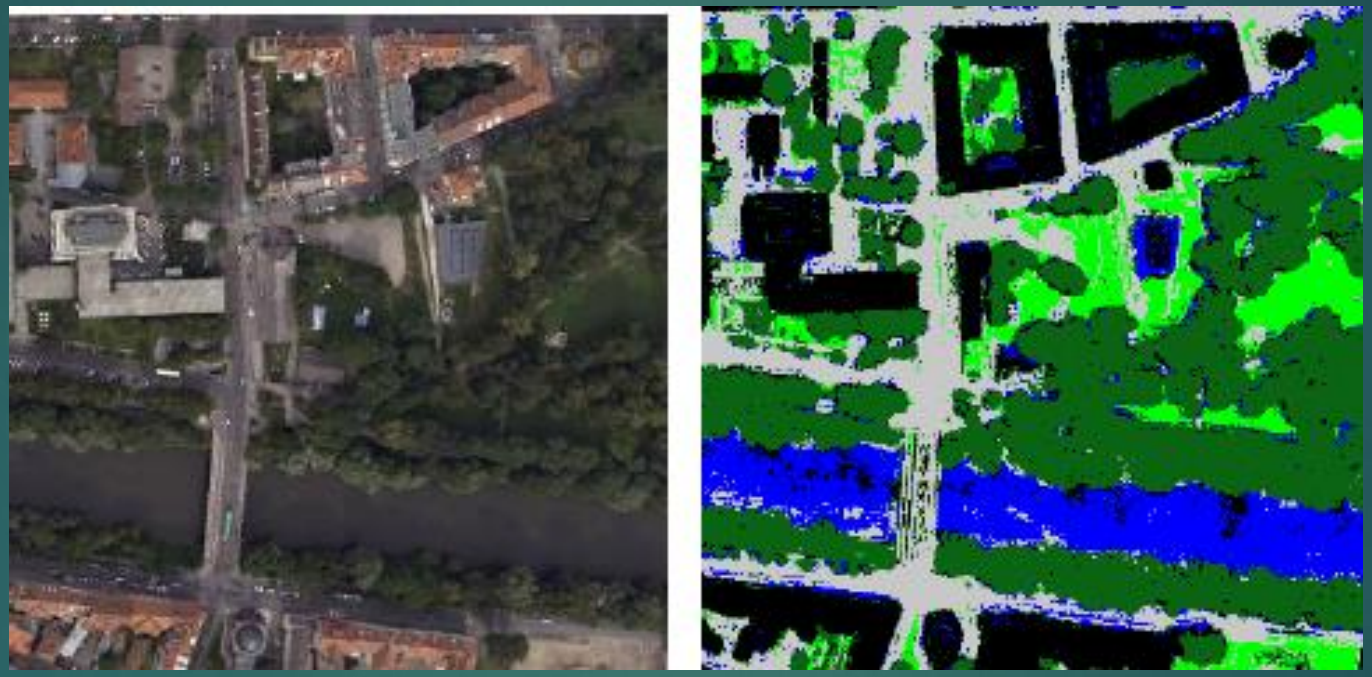
Segmented Image



(a) Semantic segmentation mask.

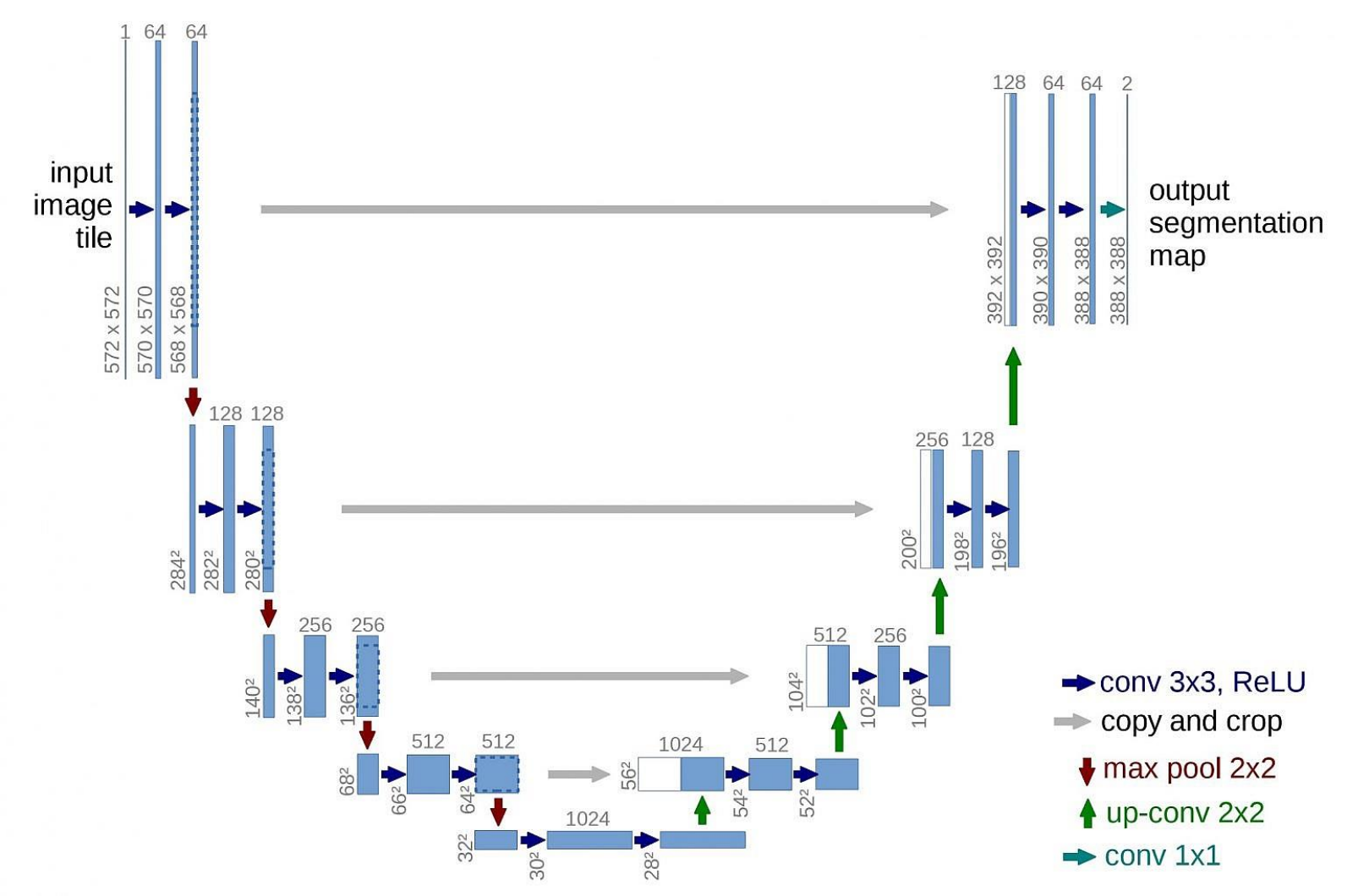
(b) Instance segmentation mask.


DeepLab V3 xception_cityscapes_trainfine (GTX980M) INPUT_SIZE=1539
Prediction time: 403ms (2.5 fps) AVG: 356ms (2.8 fps)
WX 86343



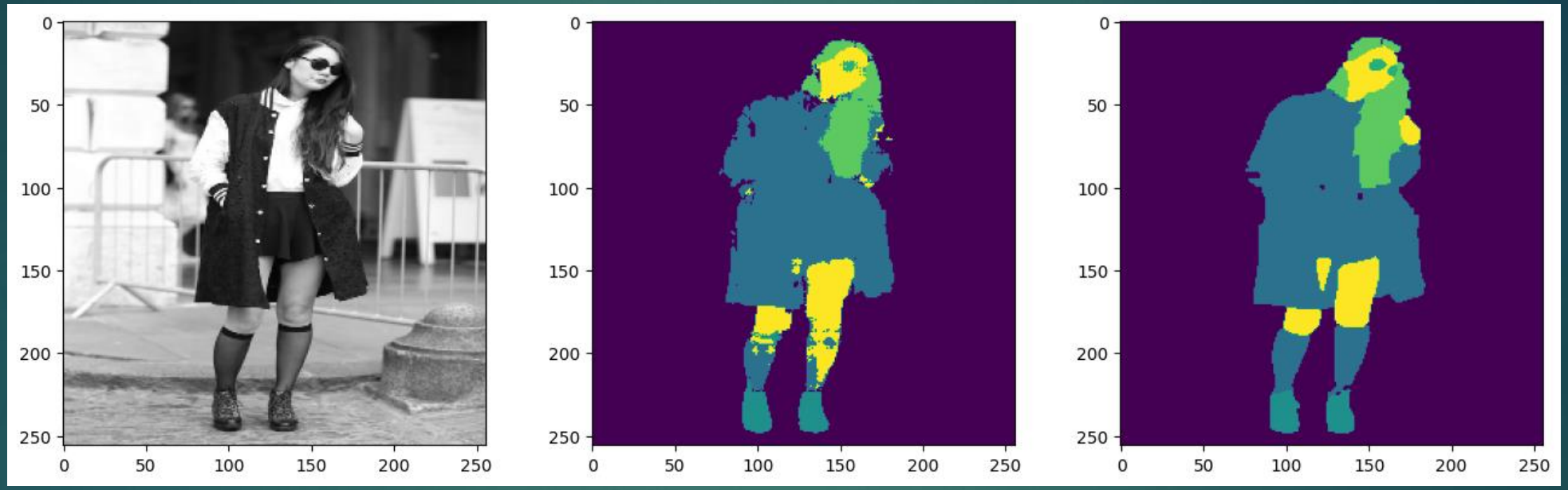
U-NET

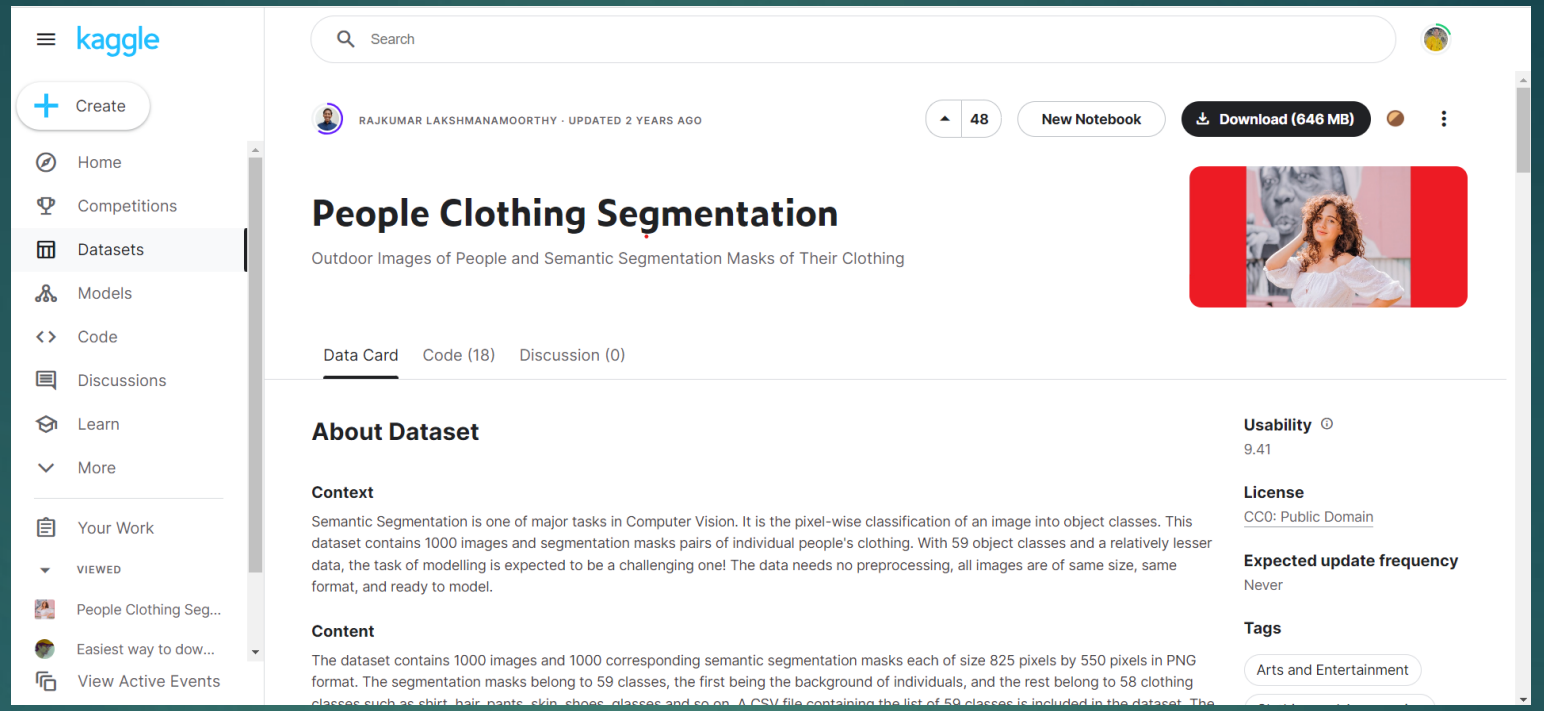
HTTPS://ARXIV.ORG/PDF/1505.04597.PDF


input image tile
output segmentation map
1
64
64
572 x 572
570 x 570
568 x 568
128
128
284²
282²
280²
256
256
140²
138²
136²
512
512
68²
66²
64²
1024
32²
30²
28²
1024
512
56²
54²
52²
512
256
104²
102²
100²
256
128
200²
198²
196²
128
64
64
2
392 x 392
390 x 390
388 x 388
388 x 388
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

People Clothing Segmentation

PROJECT


0
50
100
150
200
250
0
50
100
150
200
250
0
50
100
150
200
250
0
50
100
150
200
250
0
50
100
150
200
250
0
50
100
150
200
250



Dataset

https://www.kaggle.com/datasets/rajkumarl/people-clothing-segmentation