UCF Center for Research in Computer Vision

CAP5415 Computer Vision

Yogesh S Rawat

yogesh@ucf.edu

HEC-241

Instance Segmentation

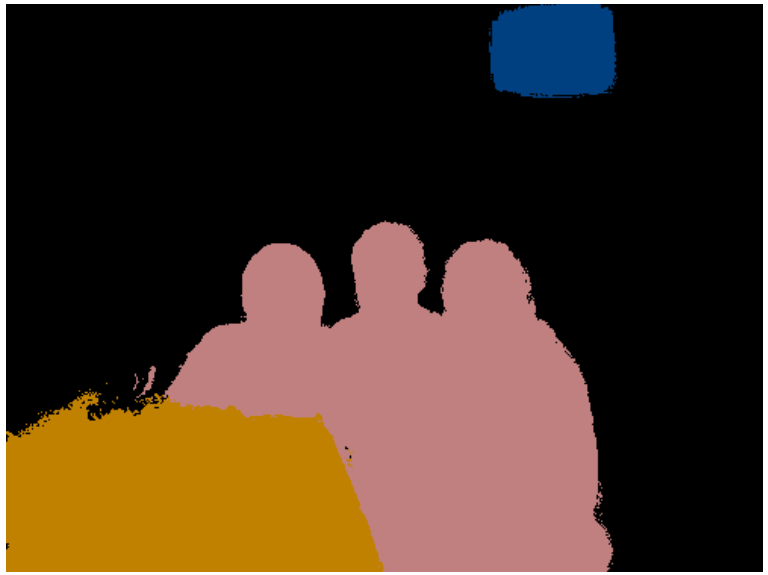
Lecture 18

Instance Segmentation

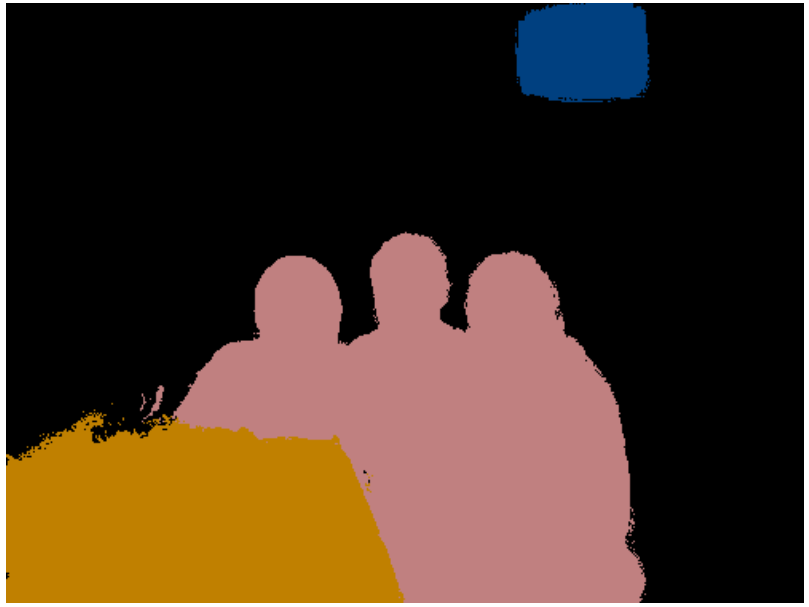
- Segment each instance of the same class separately.



Instance Segmentation

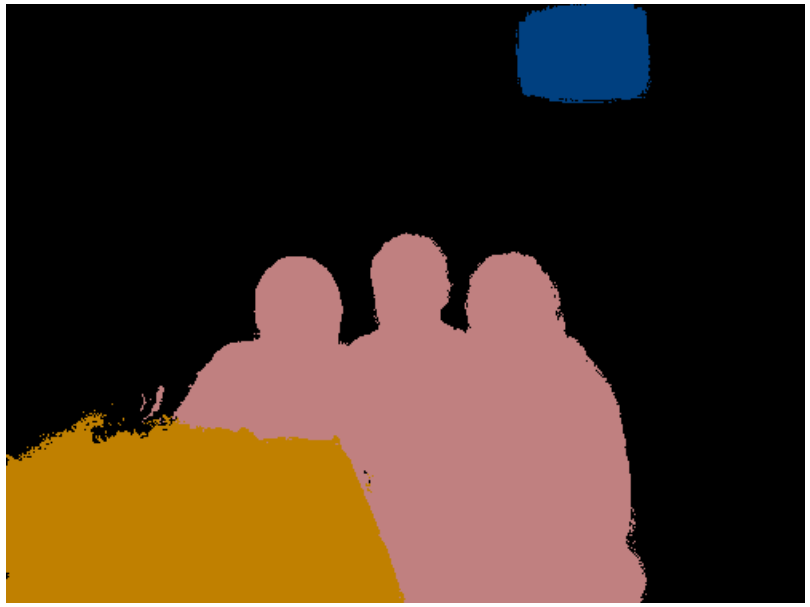


A simple solution

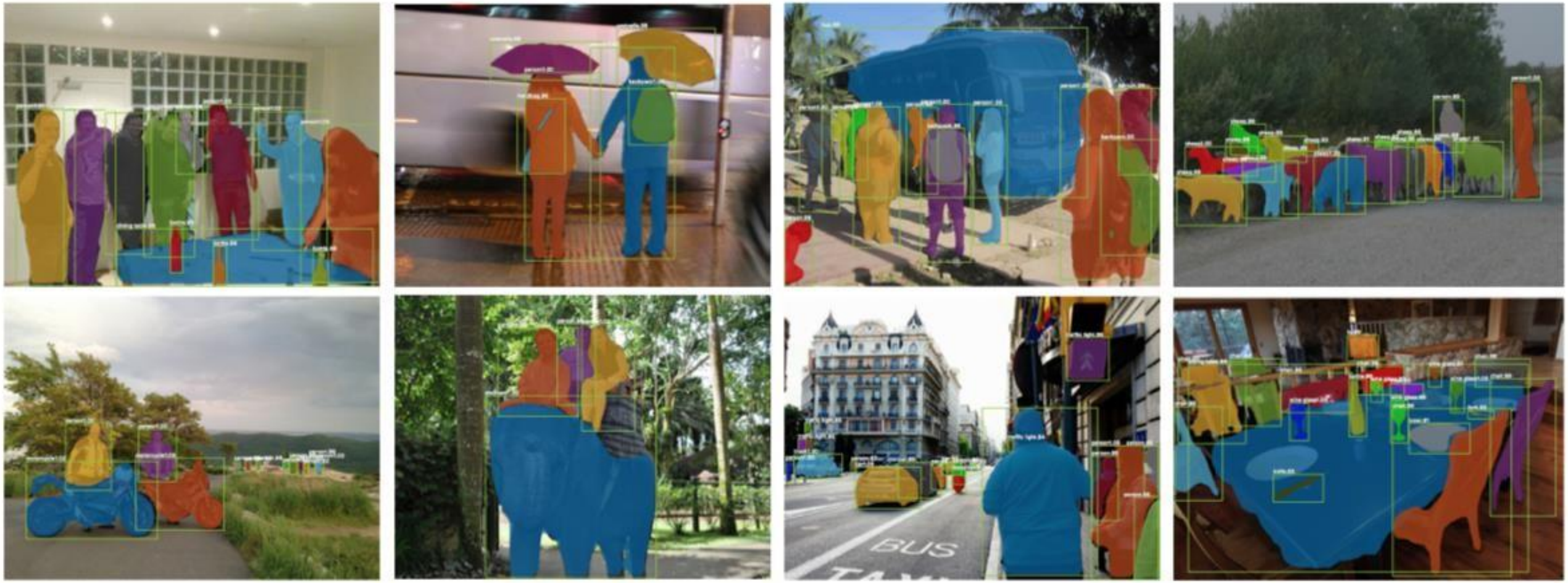


A simple solution

- Use results from object detection



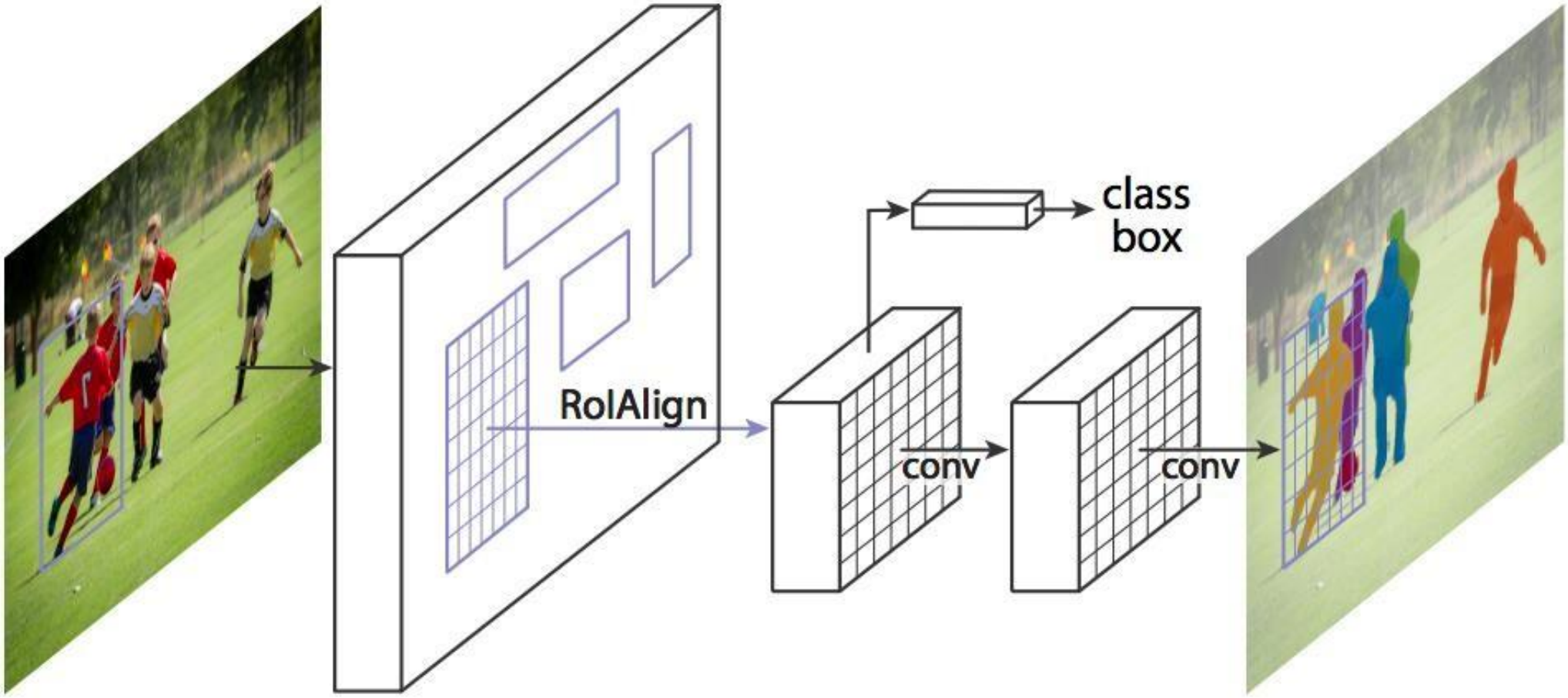
Mask - RCNN



K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017 (Best Paper Award)

Mask R-CNN

- Mask R-CNN = Faster R-CNN + FCN on RoIs

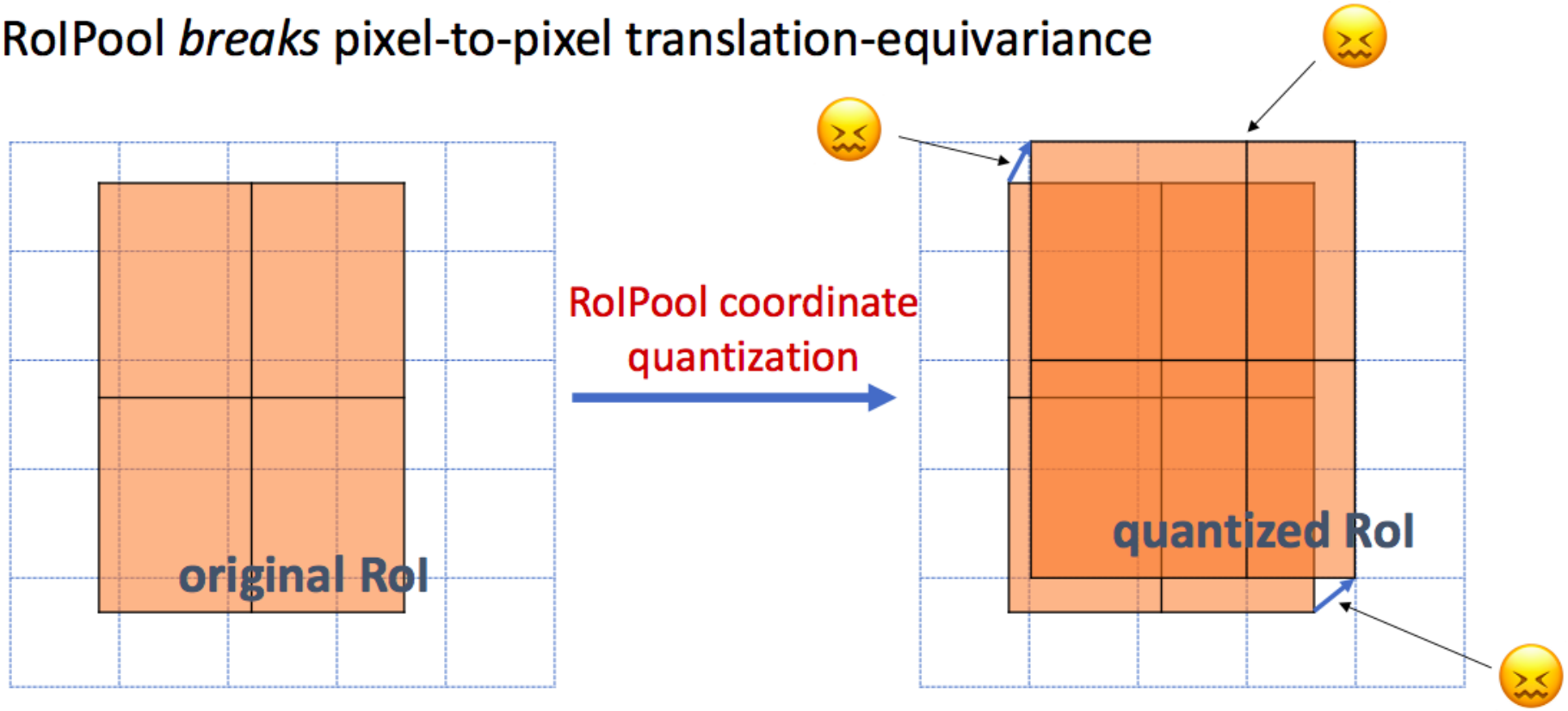


K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017 (Best Paper Award)

RoIAlign vs. RoIPool

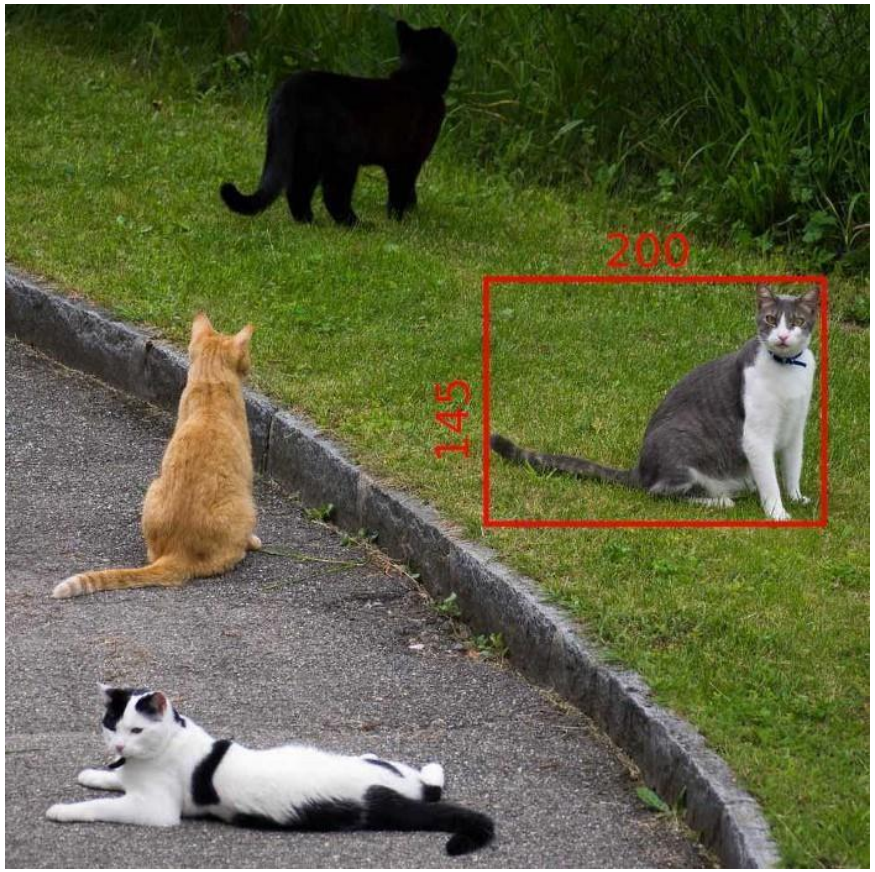
RoIPool: nearest neighbor quantization

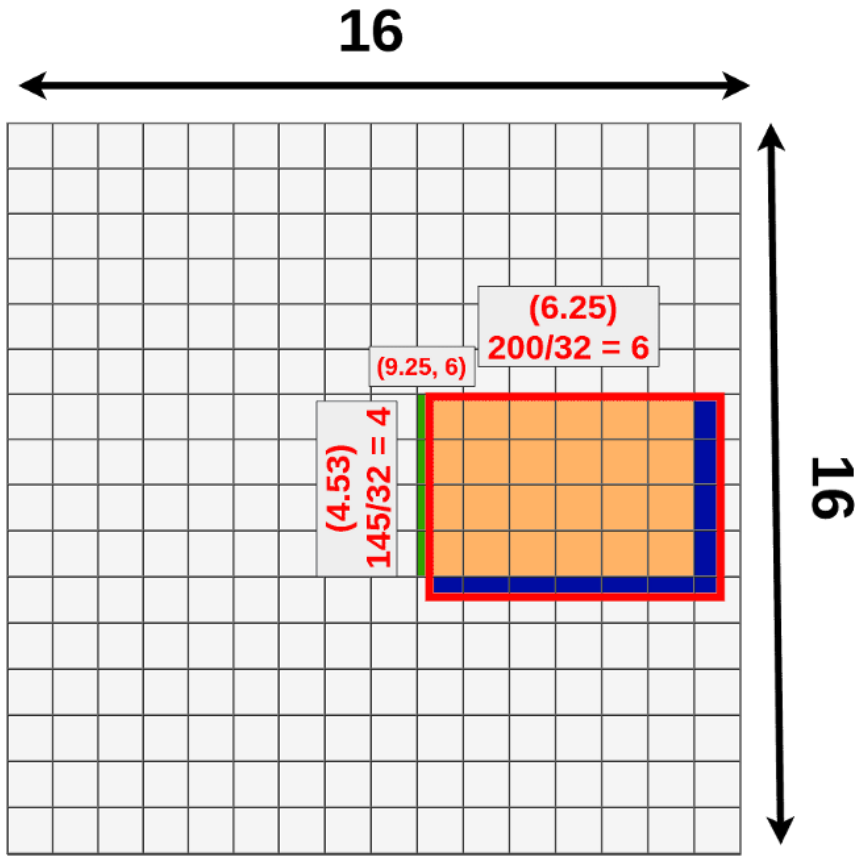
- RoIPool *breaks* pixel-to-pixel translation-equivariance



K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017 (Best Paper Award)

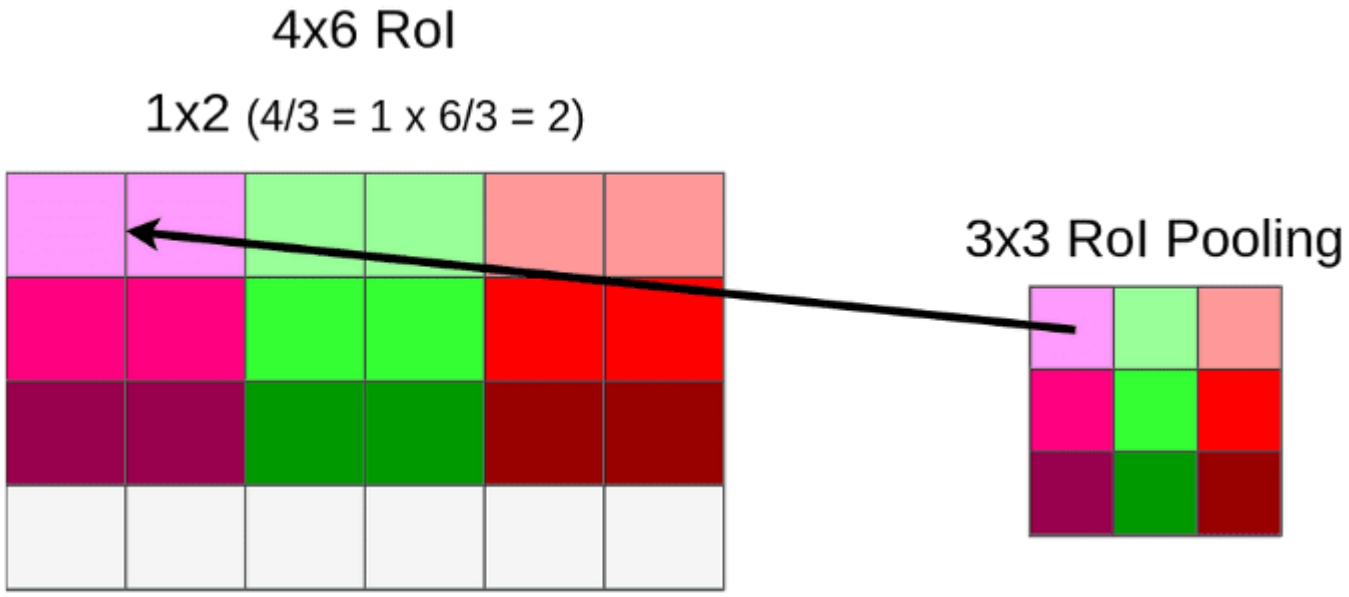
ROI-Pooling





https://tinyurl.com/y6xpm24d

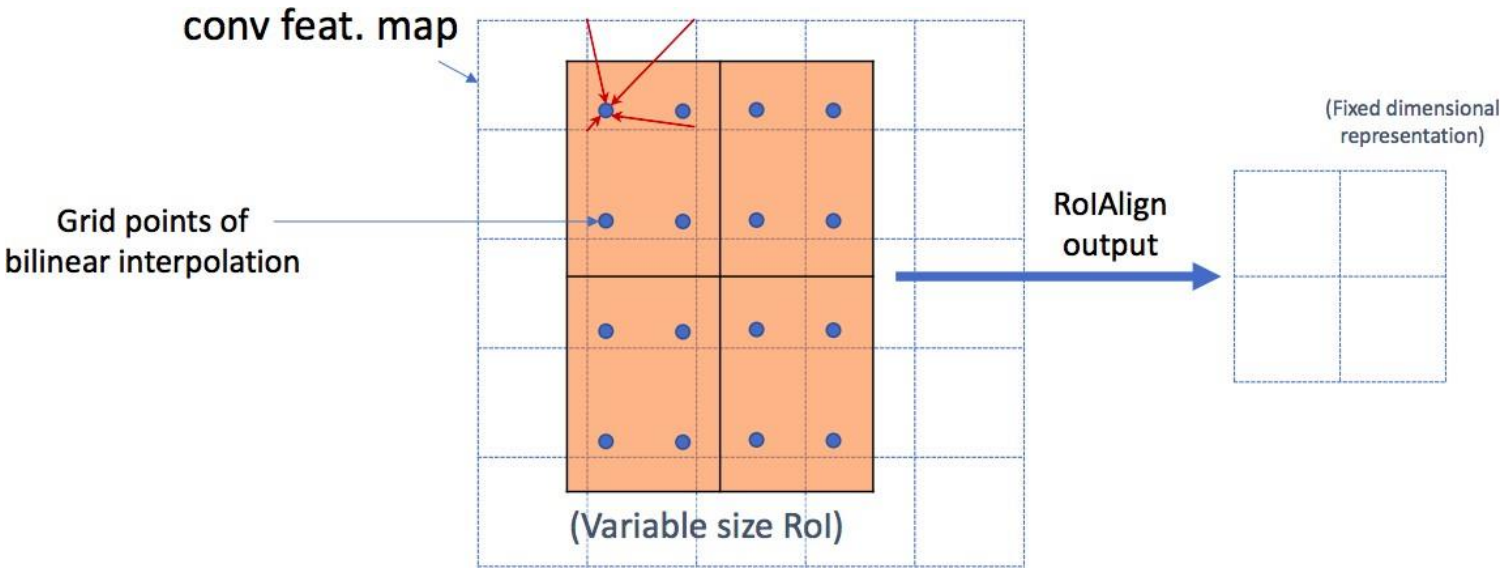
ROI-Pooling



https://tinyurl.com/y6xpm24d

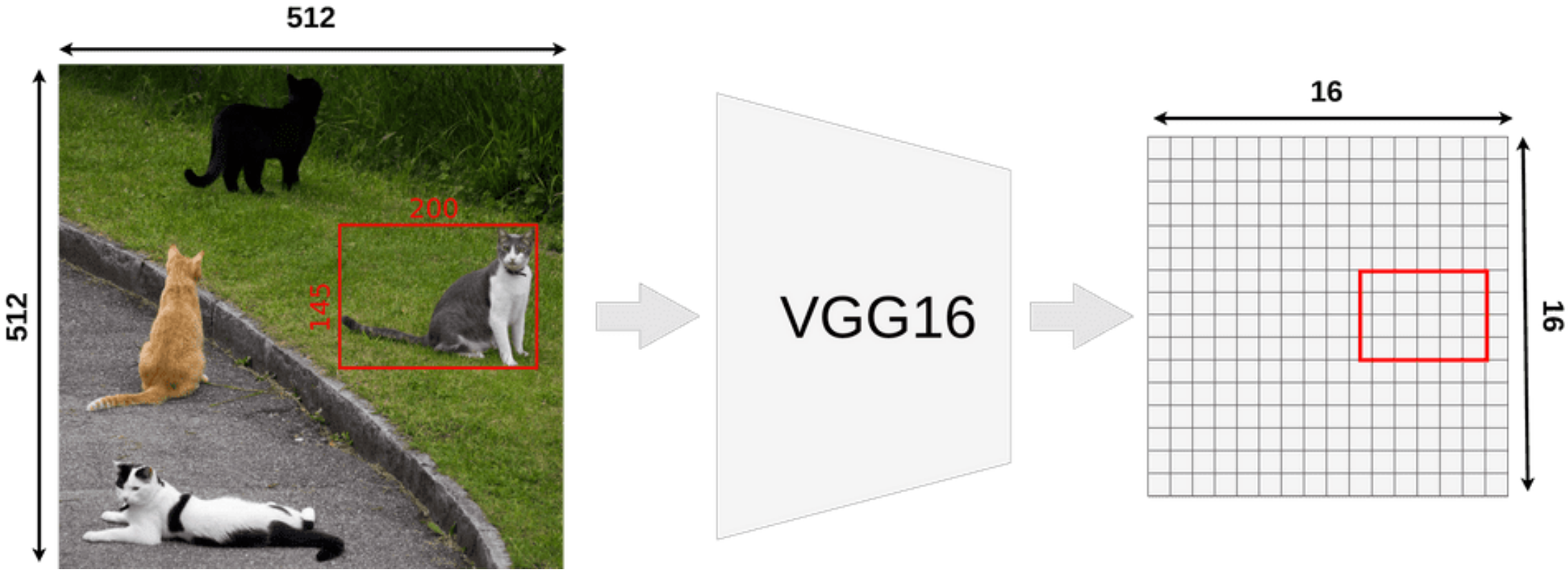
RoIAlign vs. RoIPool

- RoIPool: nearest neighbor quantization
- RoIAlign: bilinear interpolation



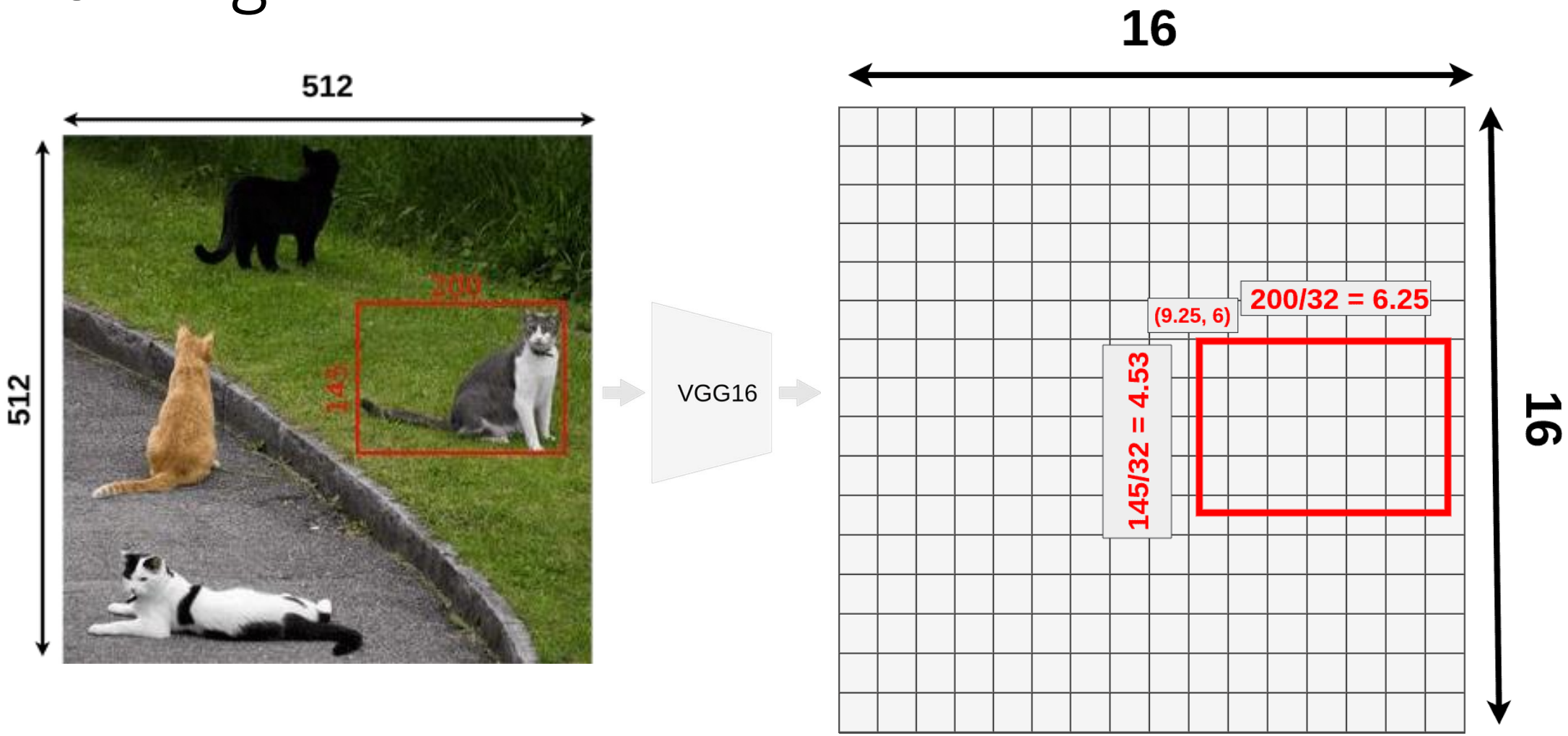
K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017 (Best Paper Award)

RoI-Align

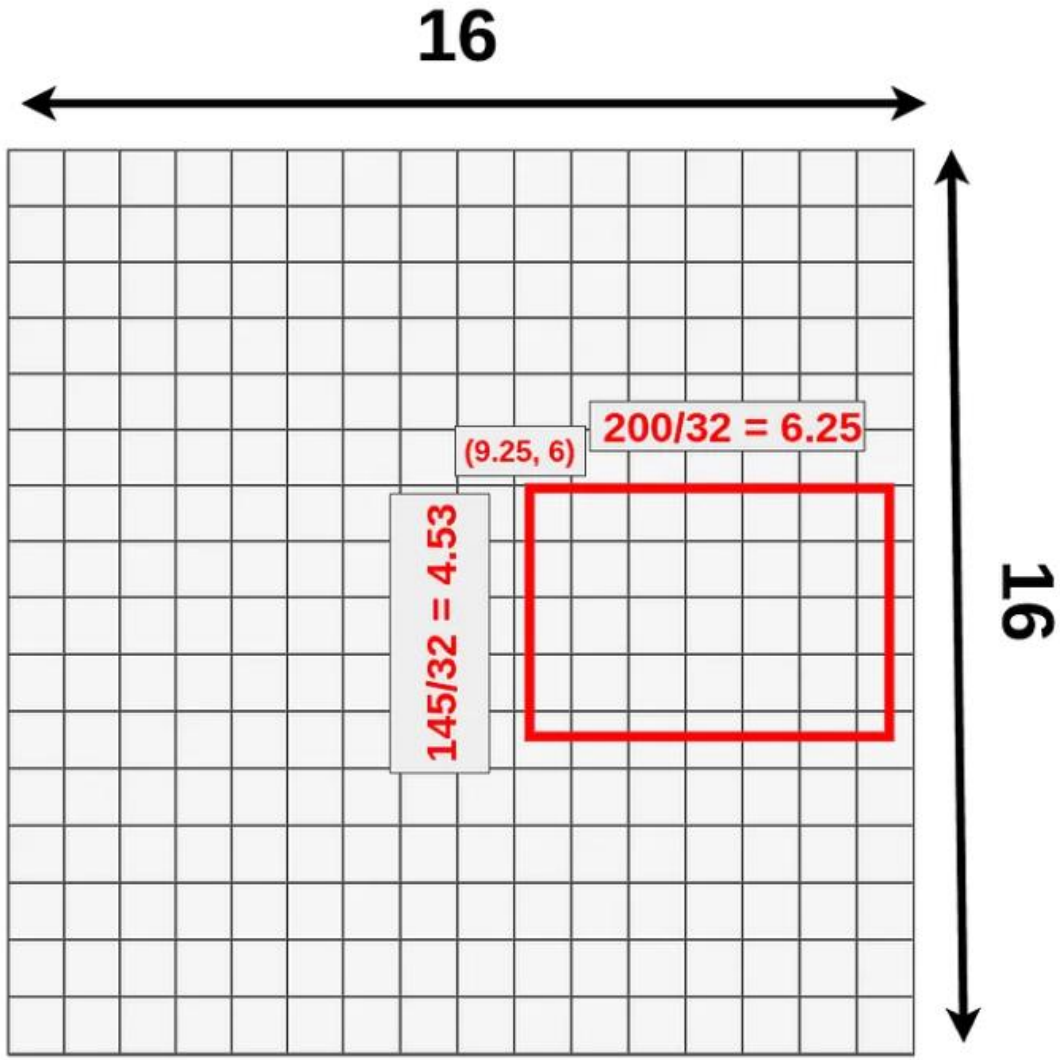


https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

RoI-Align

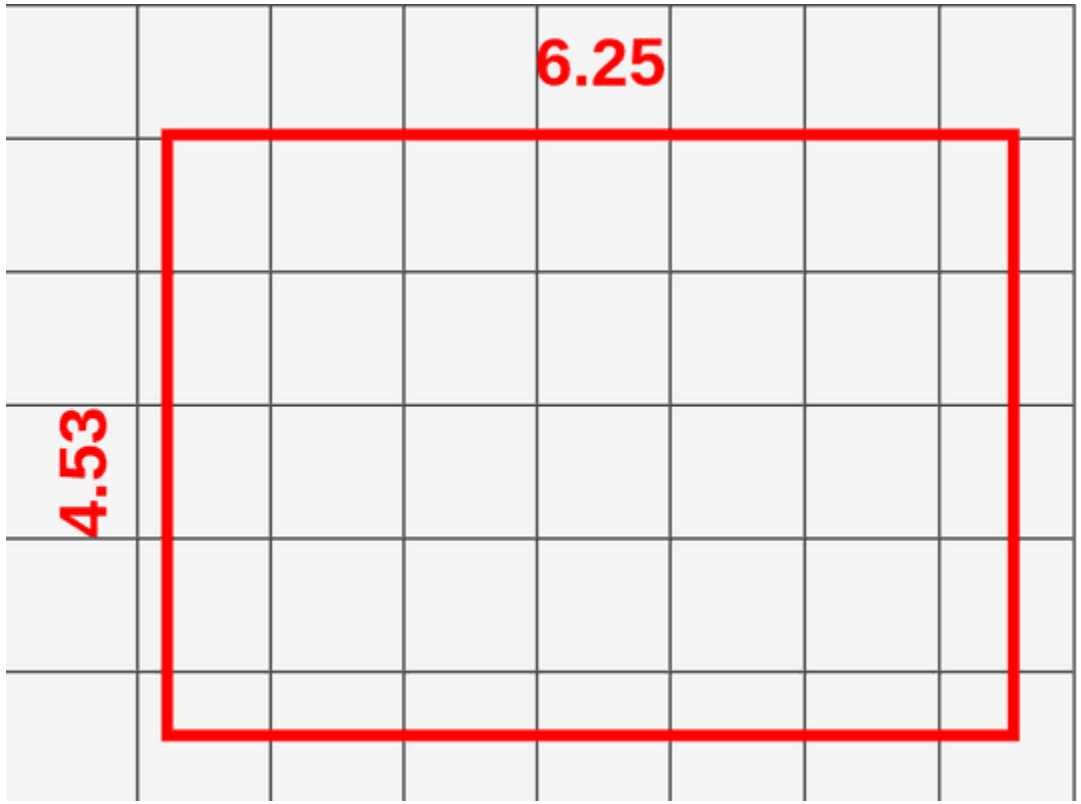


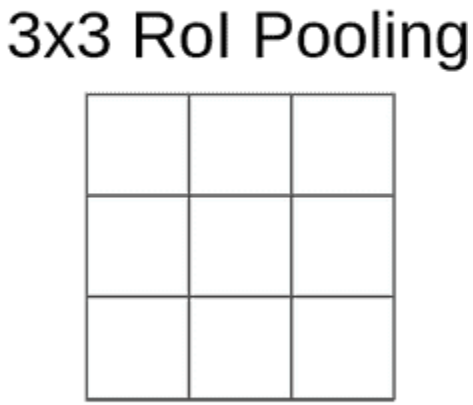
https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193



RoI placement

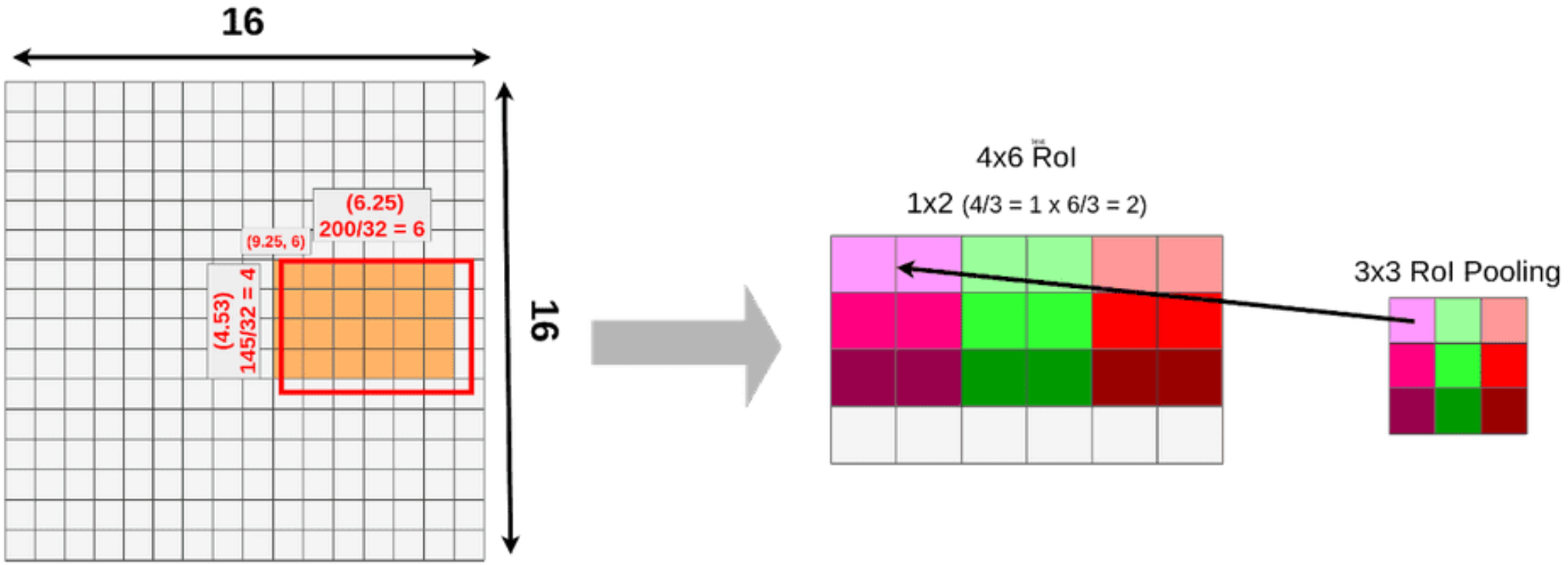
RoI-Pooling





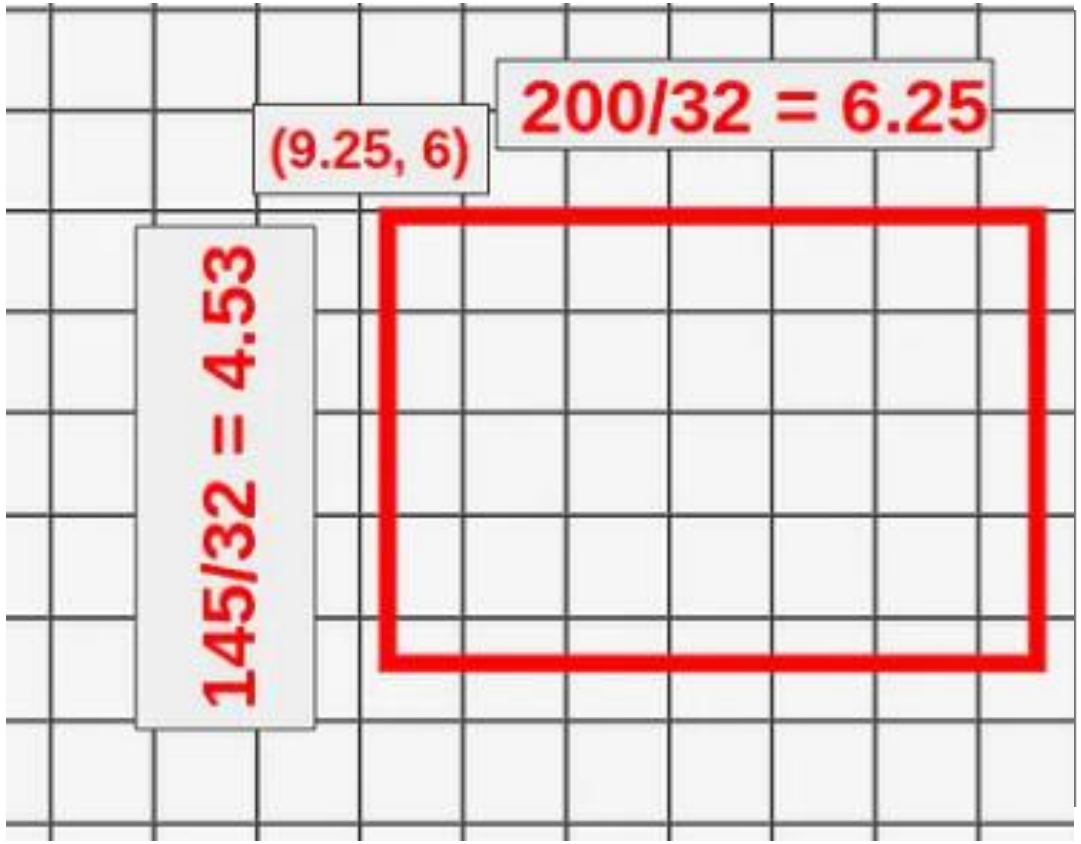
https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

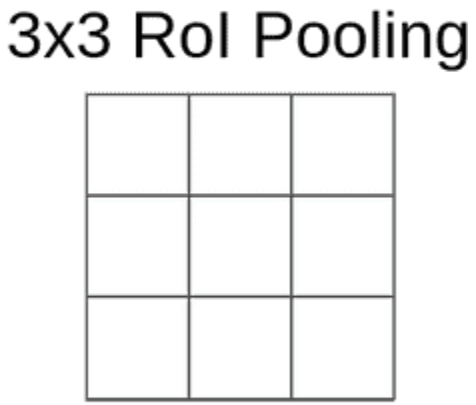
RoI-Pooling



https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

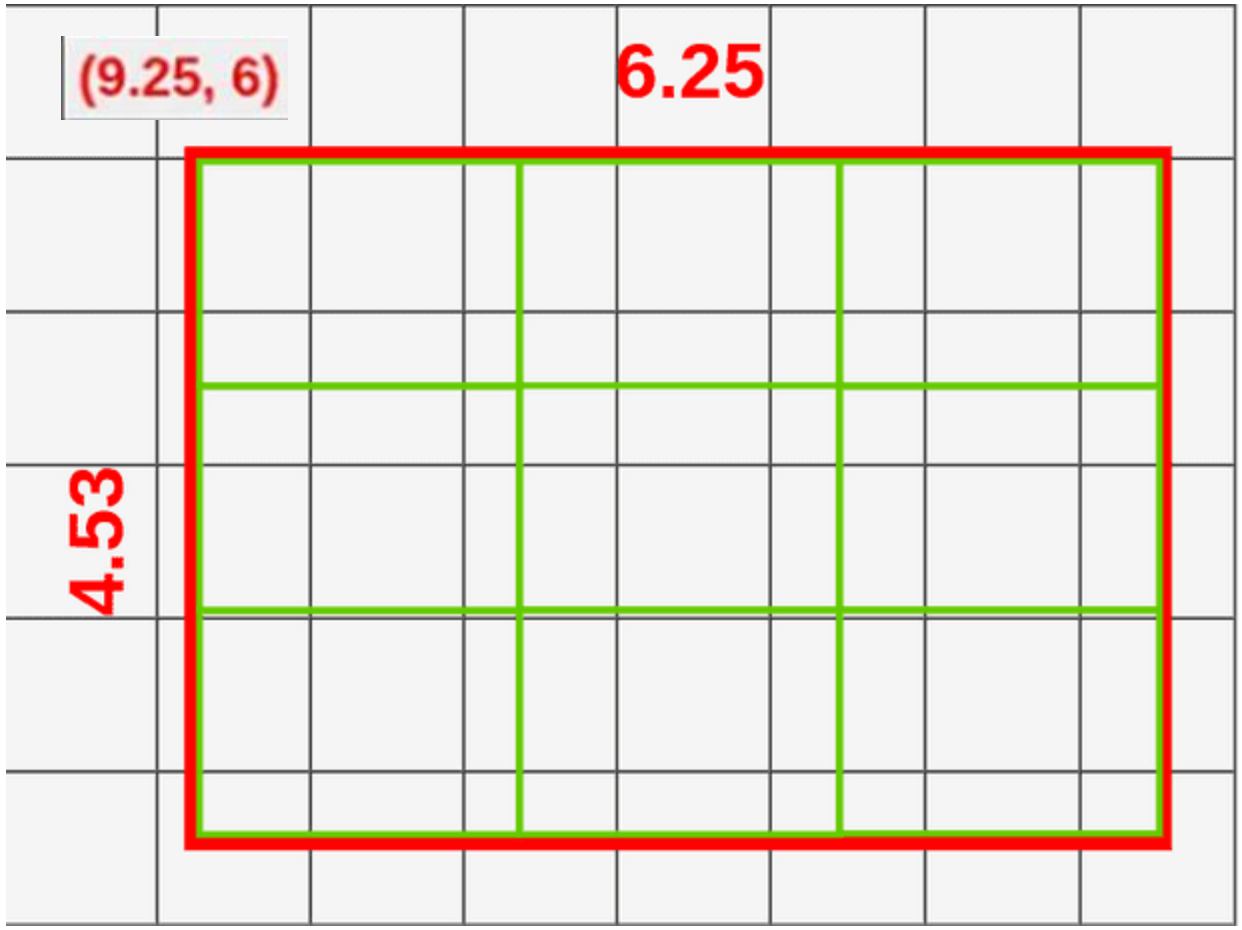
RoI-Align





https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

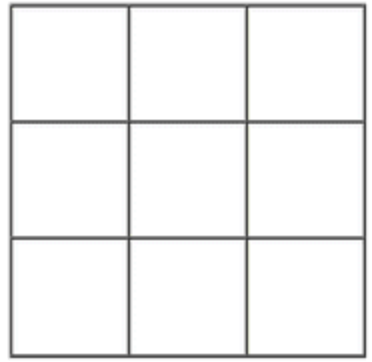
RoI-Align



Now divide the 3x3 region:

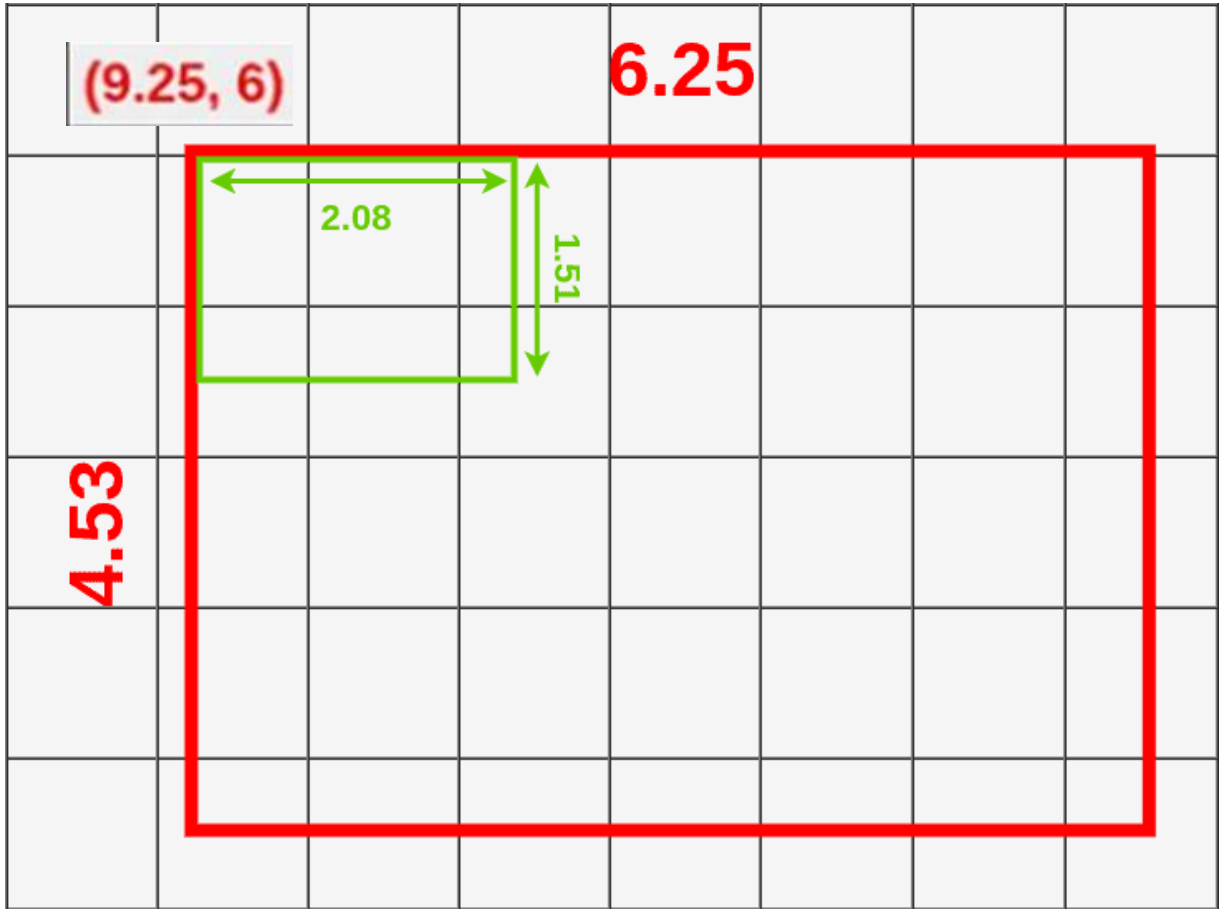
6.25 / 3 = 2.08
4.53 / 3 = 1.51

3x3 RoI Pooling



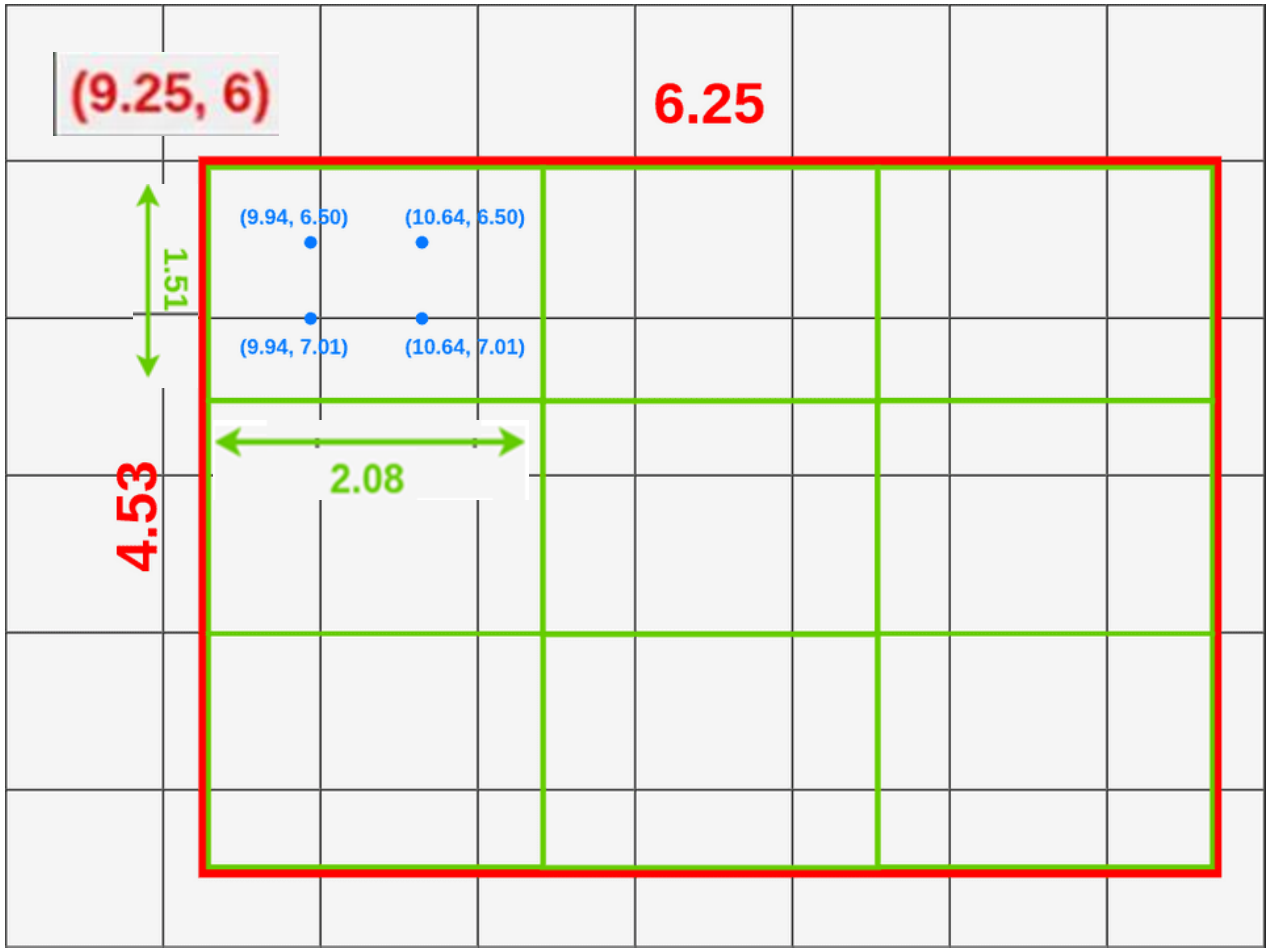
https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

RoI-Align



https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

RoI-Align



- Now further divide a 2.08 x1.25 into 3 x 3 equal part and ultimately find the coordinate of 4 points that further using **maxpool** operation help to find a single output of 3 x 3 required grid

In our case we're calculating first point (top left) coordinates like this:

- X = 9.25 + (2.08/3) * 1 = 9.94
- Y = 6 + (1.51/3) * 1 = 6.50

To calculate the second point (top right) we have to change only the X:

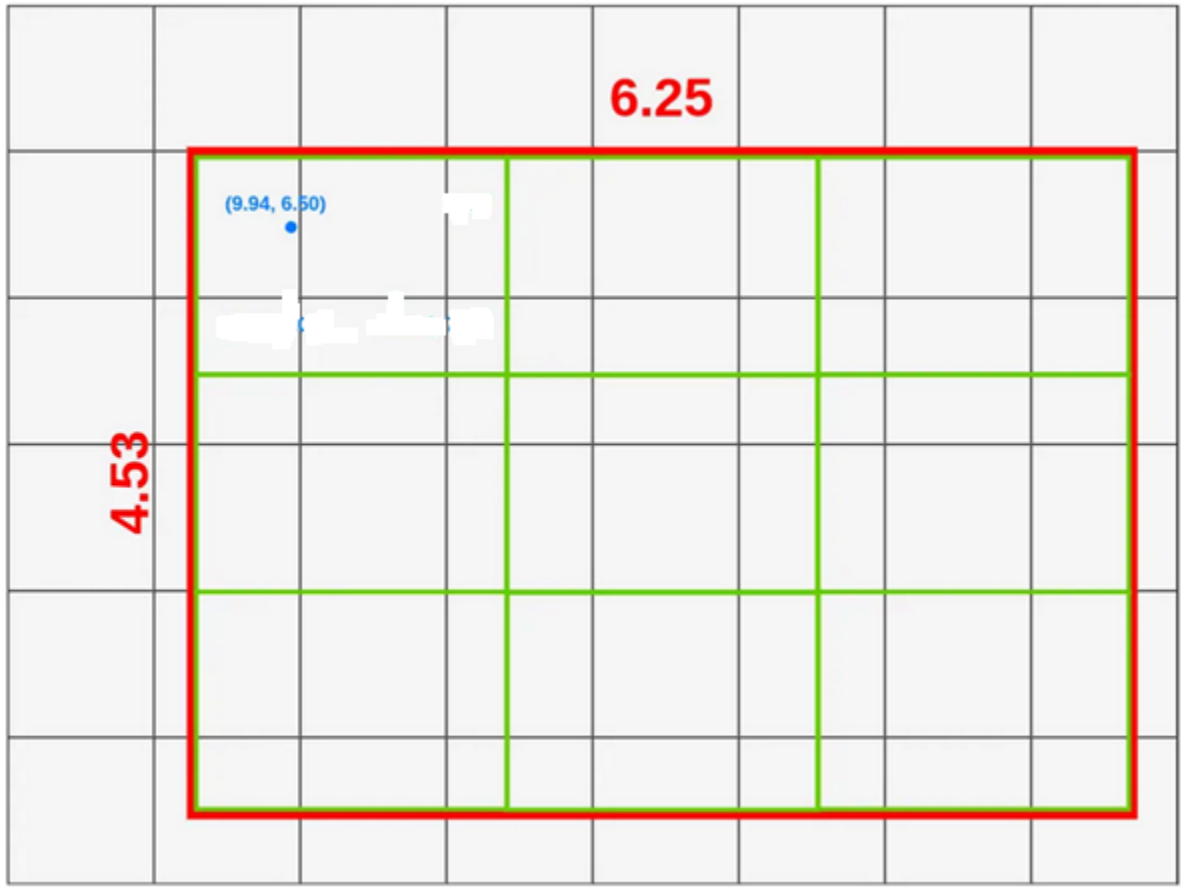
- X = 9.25 + (2.08/3) * 2 = 10.64
- Y = 6 + (1.51/3) * 1 = 6.50

- X = X_box + (2.08/3) * 1 = 9.94
- Y = Y_box + (1.51/3) * 2 = 7.01

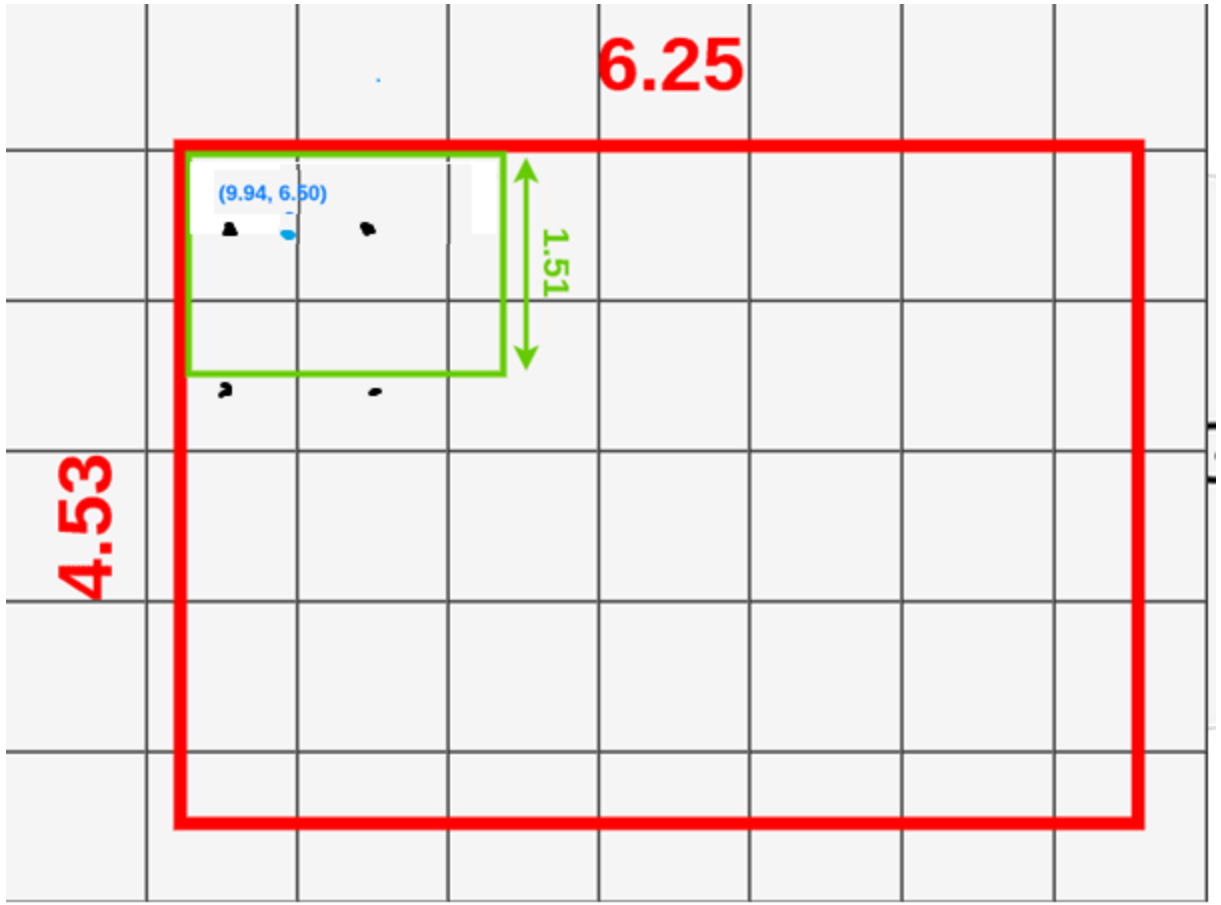
- X = X_box + (2.08/3) * 2 = 10.64
- Y = Y_box + (1.51/3) * 2 = 7.01

https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

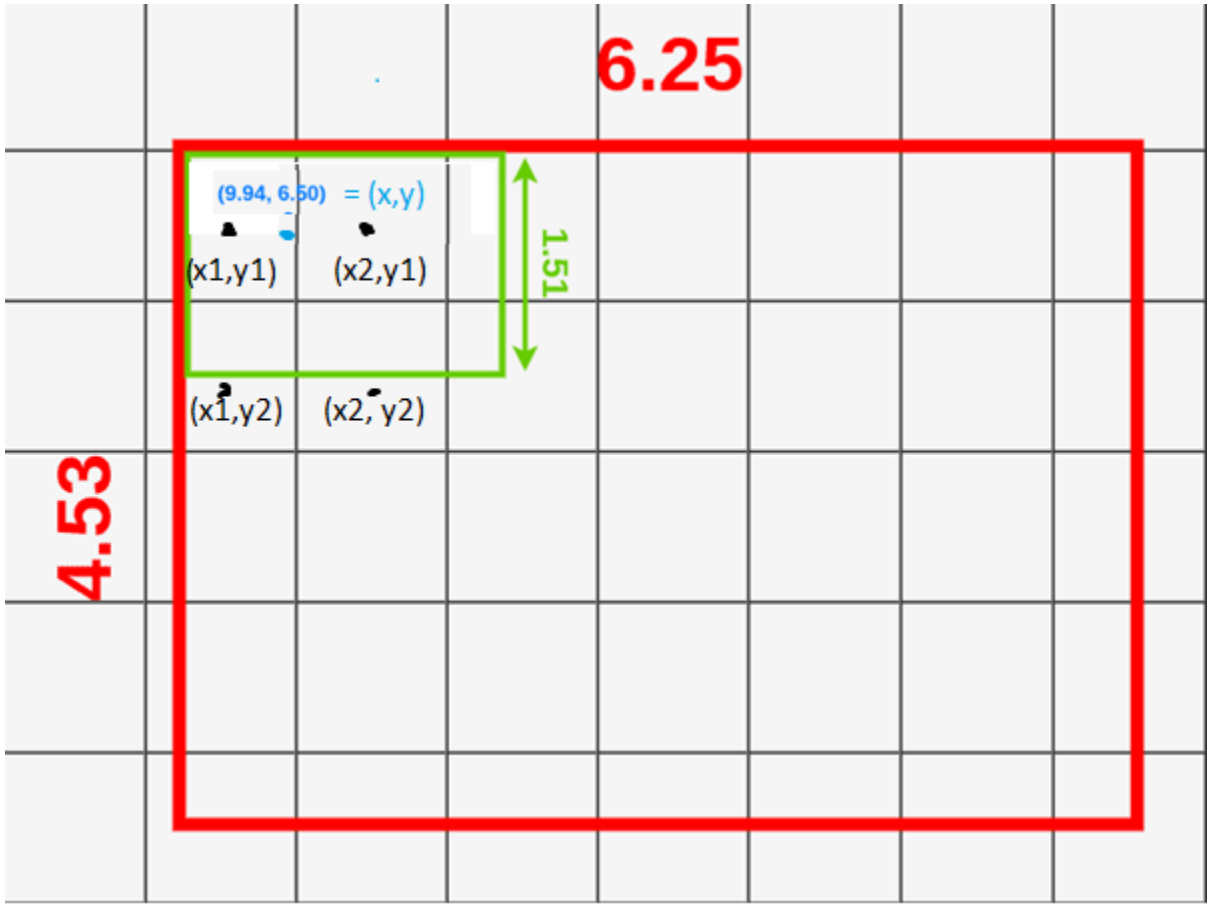
RoI-Align



Sampling points distribution

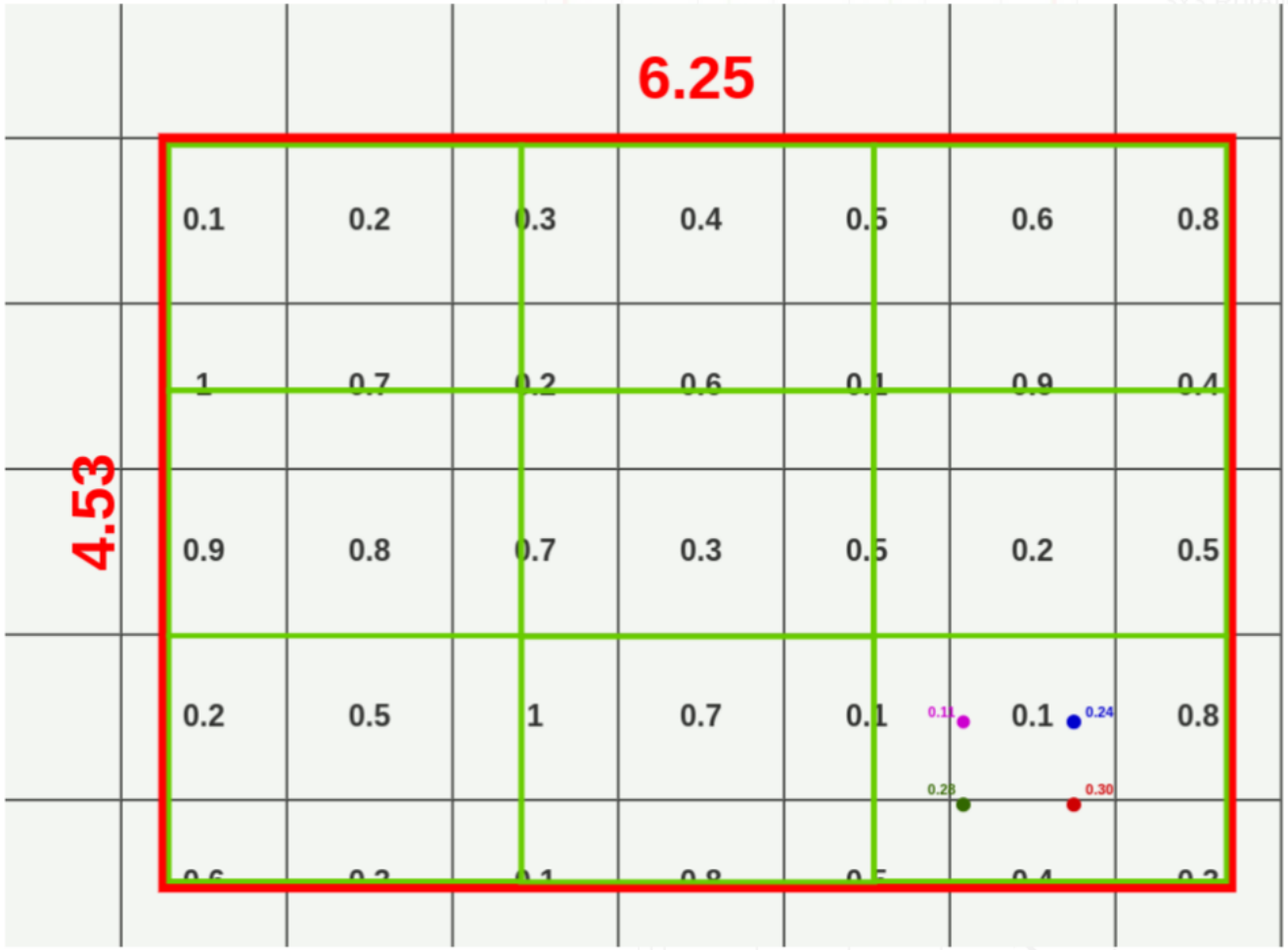


- Our point has coordinates **(9.44, 6.50)**
- Connecting it with 4 closest neighboring cells



(x, y) = (9.94, 6.50)

(x1, y1) = (9.5, 6.50),
(x2, y1) = (10.5, 6.50),
(x1, y2) = (9.5, 7.50),
(x2, y2) = (10.5, 7.50),



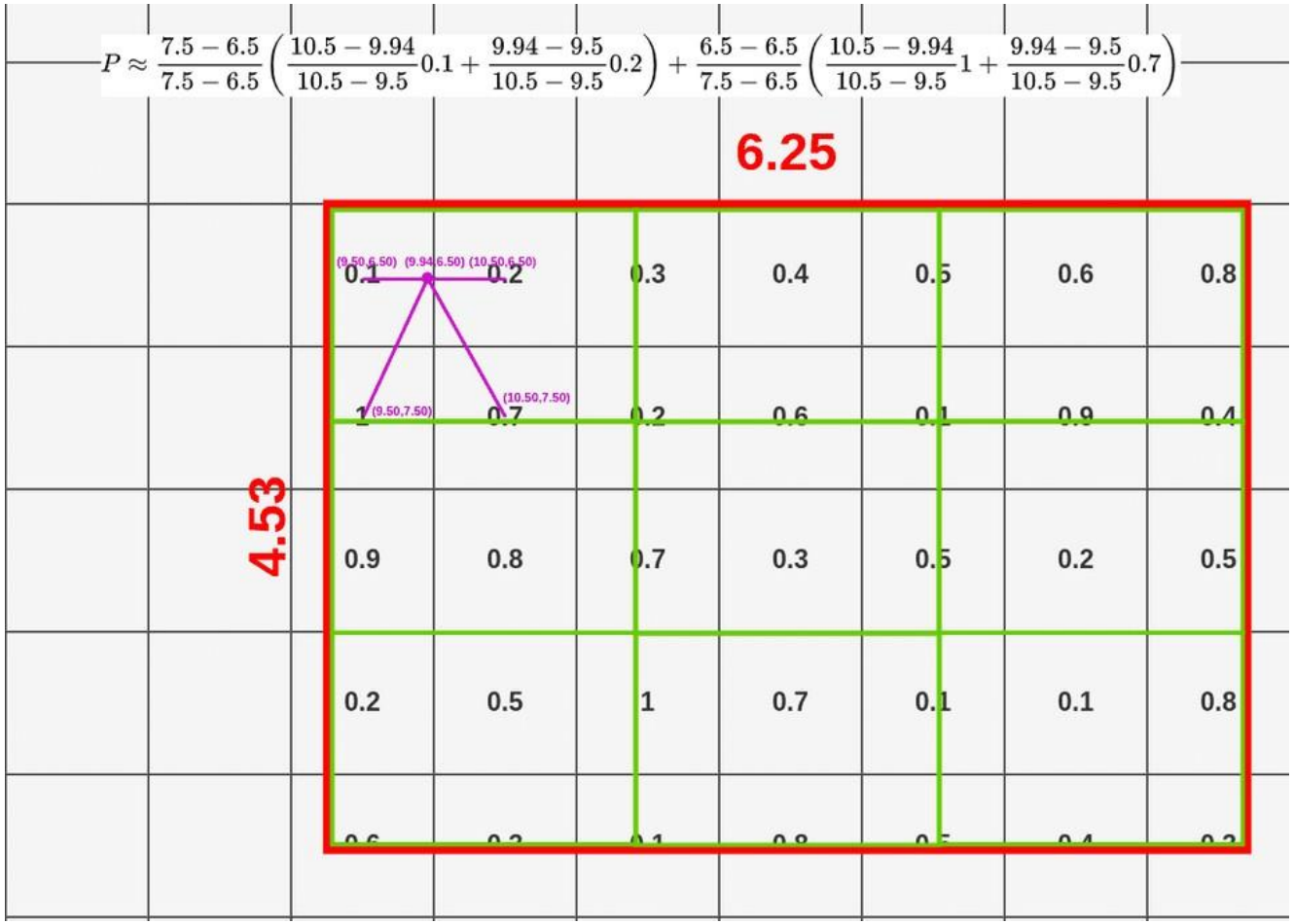
$Q_{11} = 0.1, Q_{21} = 0.2$

$Q_{12} = 1, \quad Q_{22} = 0.7$

RoI-Align

$$P \approx \frac{y_2 - y}{y_2 - y_1}\left(\frac{x_2 - x}{x_2 - x_1}Q_{11} + \frac{x - x_1}{x_2 - x_1}Q_{21}\right) + \frac{y - y_1}{y_2 - y_1}\left(\frac{x_2 - x}{x_2 - x_1}Q_{12} + \frac{x - x_1}{x_2 - x_1}Q_{22}\right)$$

Bilinear Interpolation equation



(x, y) = (9.94, 6.50)
(x1, y1) = (9.5, 6.50), (x2, y1) = (10.5, 6.50),
(x1, y2) = (9.5, 7.50), (x2, y2) = (10.5, 7.50),

$Q_{11} = 0.1$, $Q_{21} = 0.2$
$Q_{12} = 1$, $Q_{22} = 0.7$

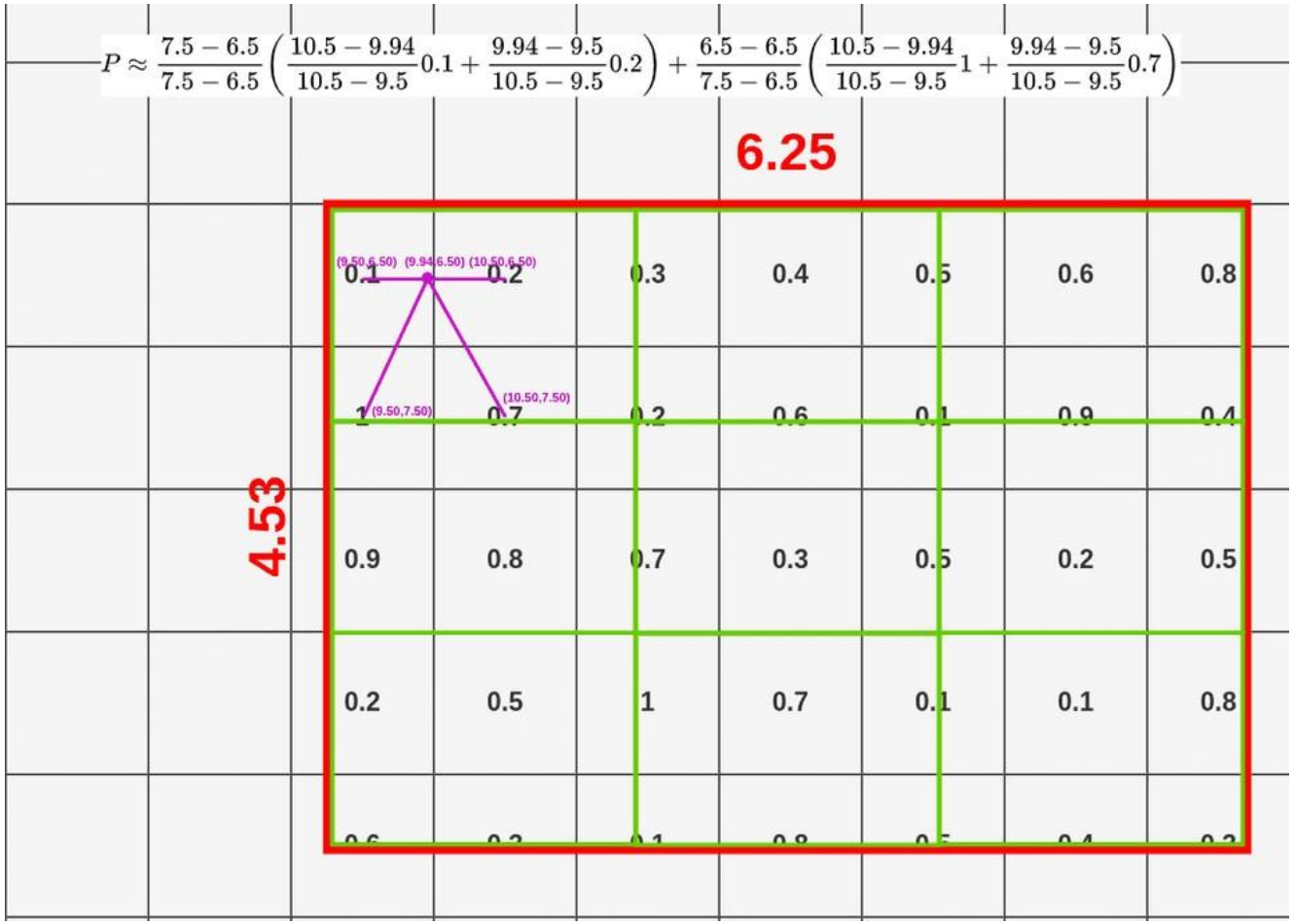
P(1,1)= P(9.94, 6.50)
= (7.5 – 6.5)/(7.5 – 6.5) ((10.5- 9.94)/ (10.5- 9.5) * 0.1 + (9.94 – (10.5- 9.5) * 0.2)) + (6.5 – 6.5)/(7.5 – 6.5)(((10.5 – 9.94)/10.5- 9.5 + ((9.94- 9.5) /(10.5-9.5)) * 0.7)
= 0.14

https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

RoI-Align

$$P \approx \frac{y_2 - y}{y_2 - y_1}\left(\frac{x_2 - x}{x_2 - x_1}Q_{11} + \frac{x - x_1}{x_2 - x_1}Q_{21}\right) + \frac{y - y_1}{y_2 - y_1}\left(\frac{x_2 - x}{x_2 - x_1}Q_{12} + \frac{x - x_1}{x_2 - x_1}Q_{22}\right)$$

Bilinear Interpolation equation

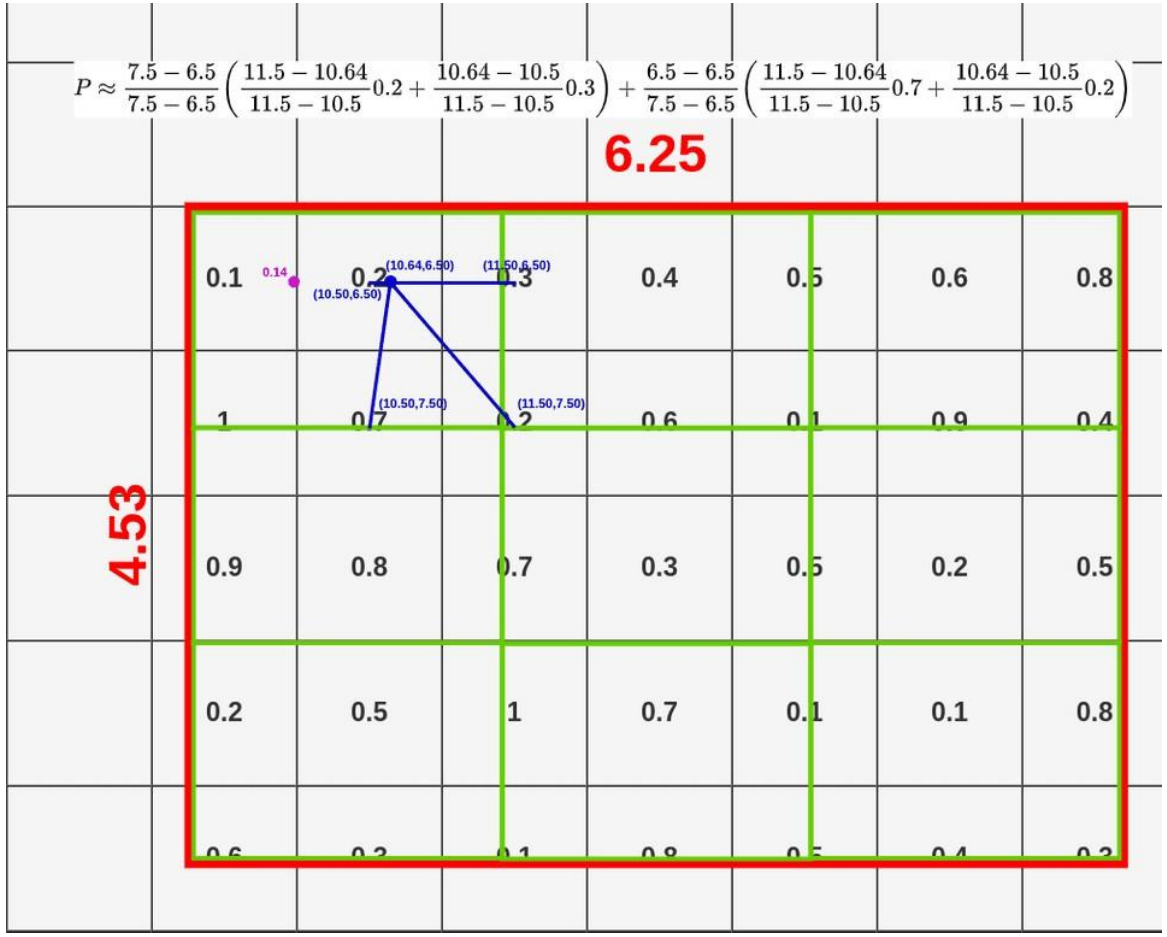


When you take the first point from our box, you're connecting with closest neighboring cells (exactly to the middle).
In this case, our point has coordinates **(9.44, 6.50)**.

Closest middle of the cell in top-left direction is **(9.50, 6.50)** (would be (9.50, 5.50) if our point was only 0.01 higher on the grid). Then we have to select a bottom-left point and the closes one is **(9.50, 7.50)**. Following the same rule, we're selecting **(10.50, 6.50)** and **(10.50, 7.50)** as top-right and bottom-right points. Above the RoI, you could see the whole calculation to get the value for the first point is **0.14**

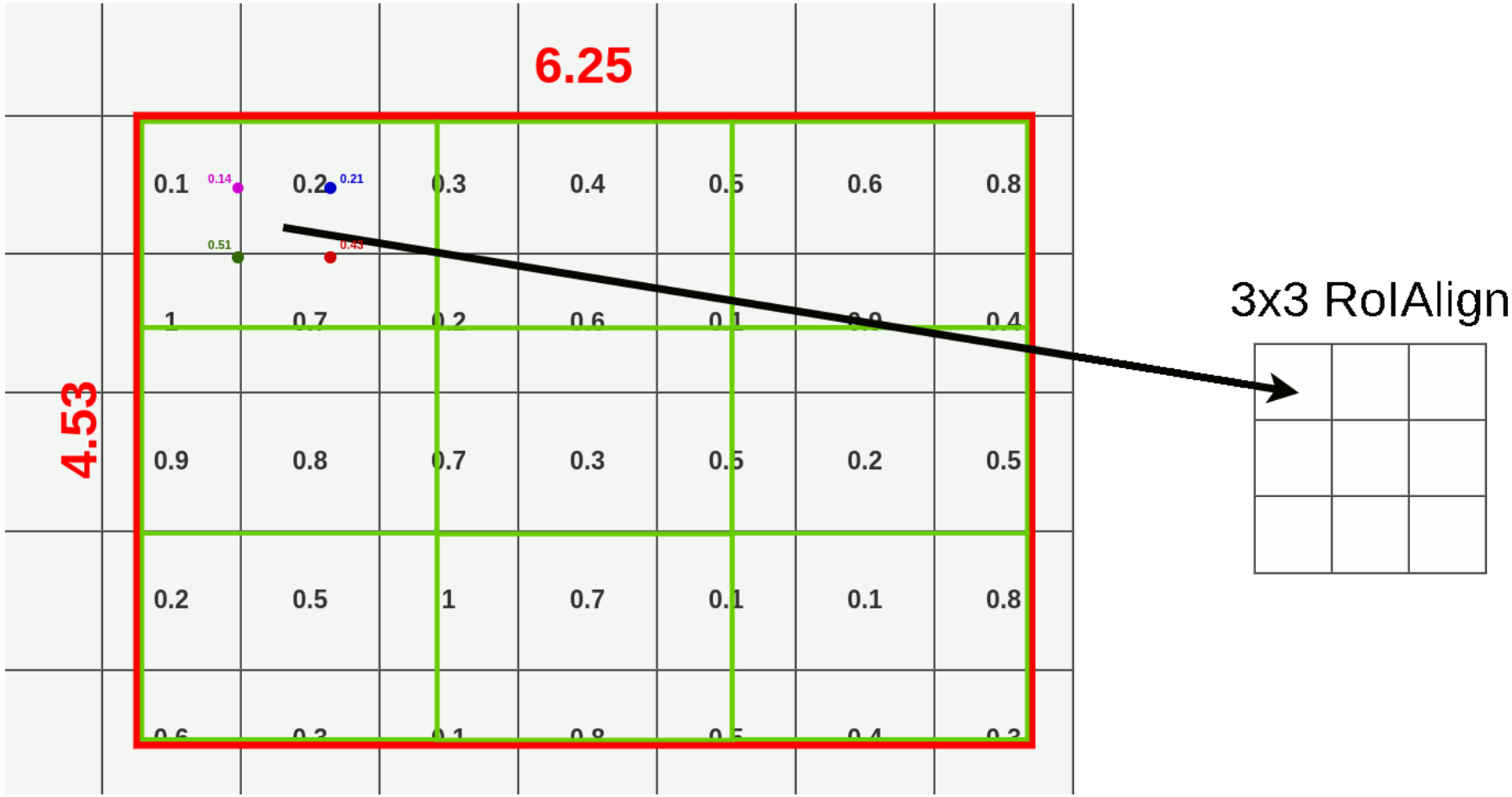
https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

RoI-Align



https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

RoI-Align



https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

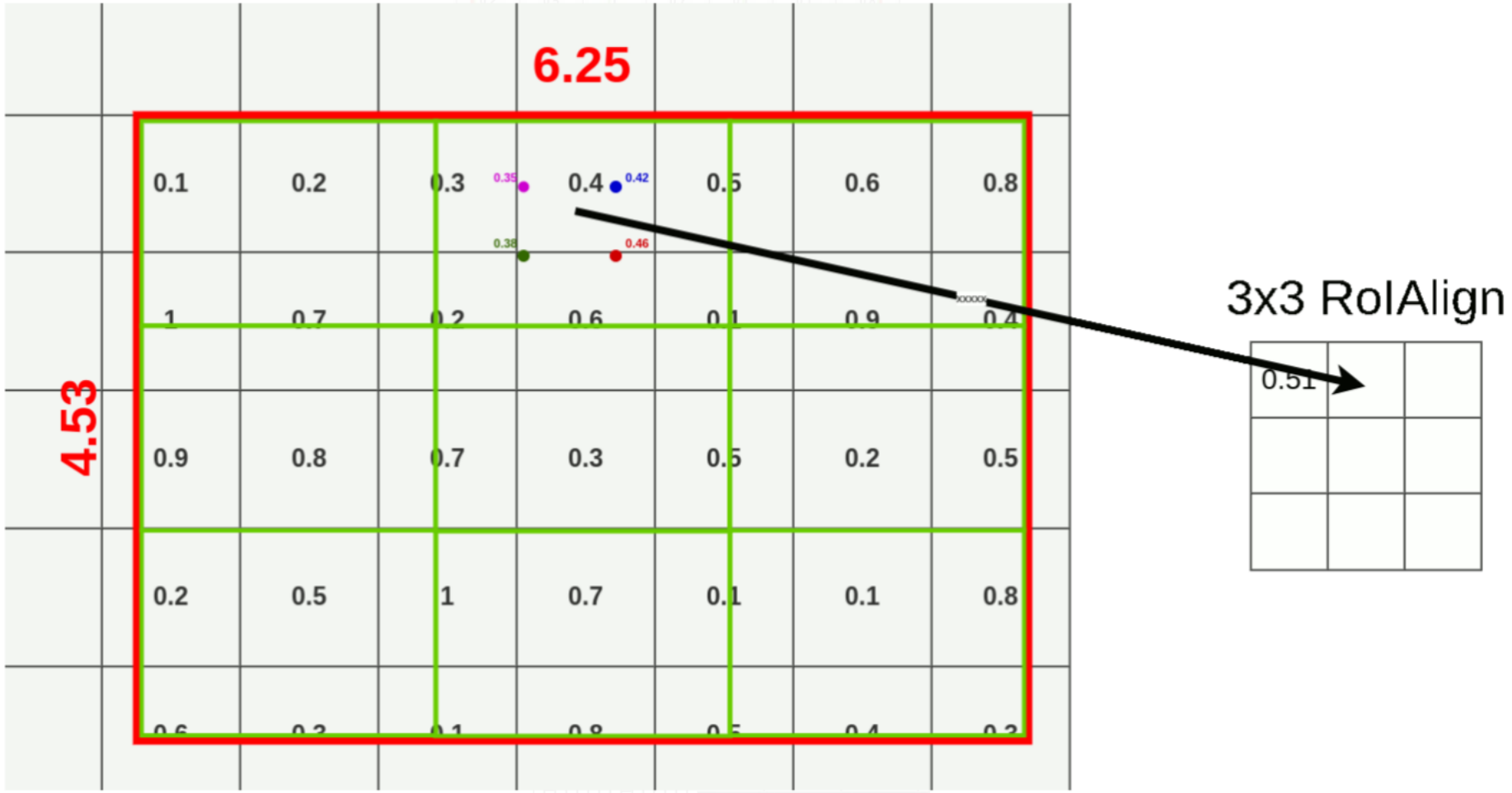
RoI-Align

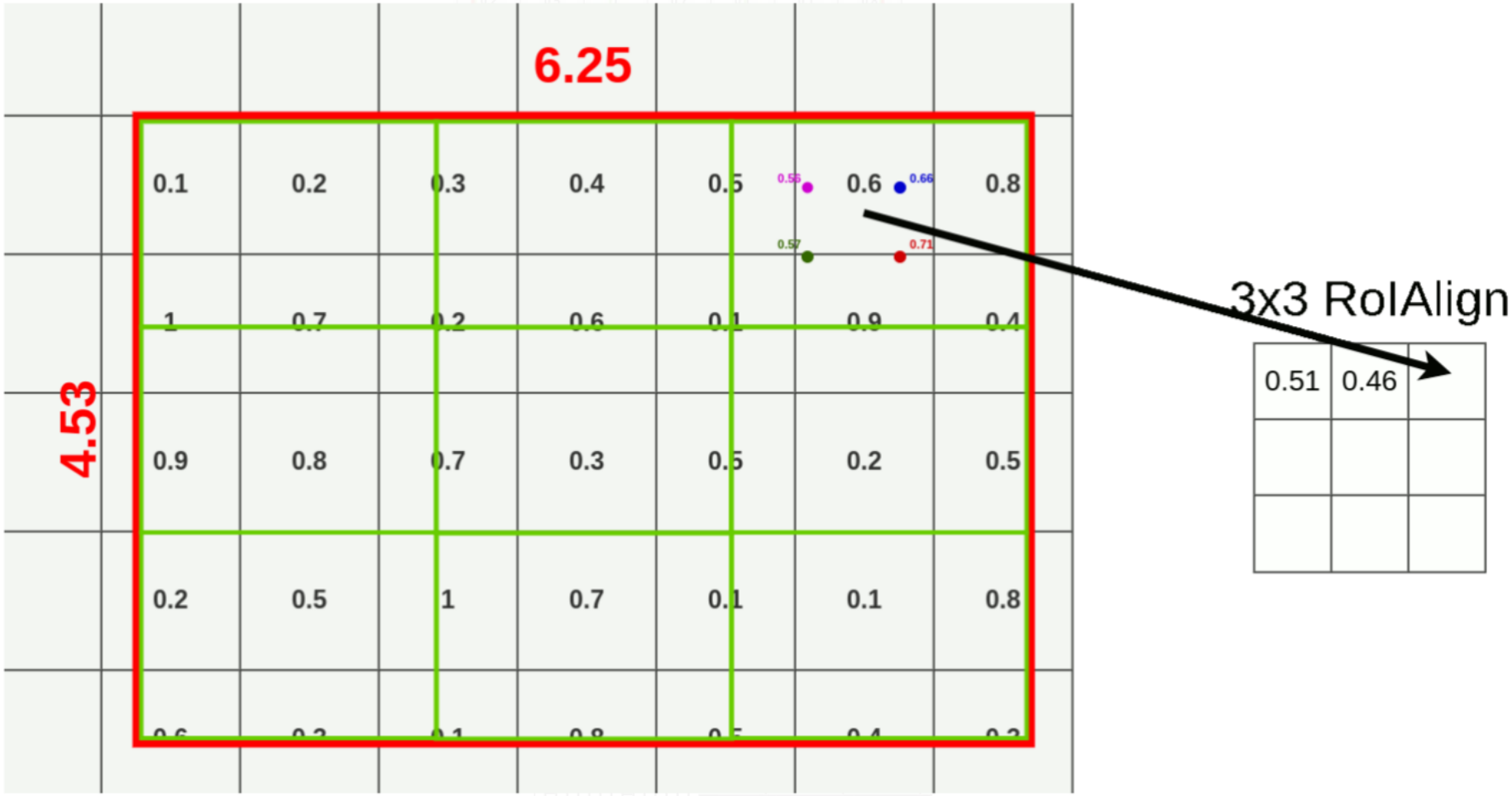
1x1 = MAX(0.14, 0.21, 0.51, 0.43) = 0.51

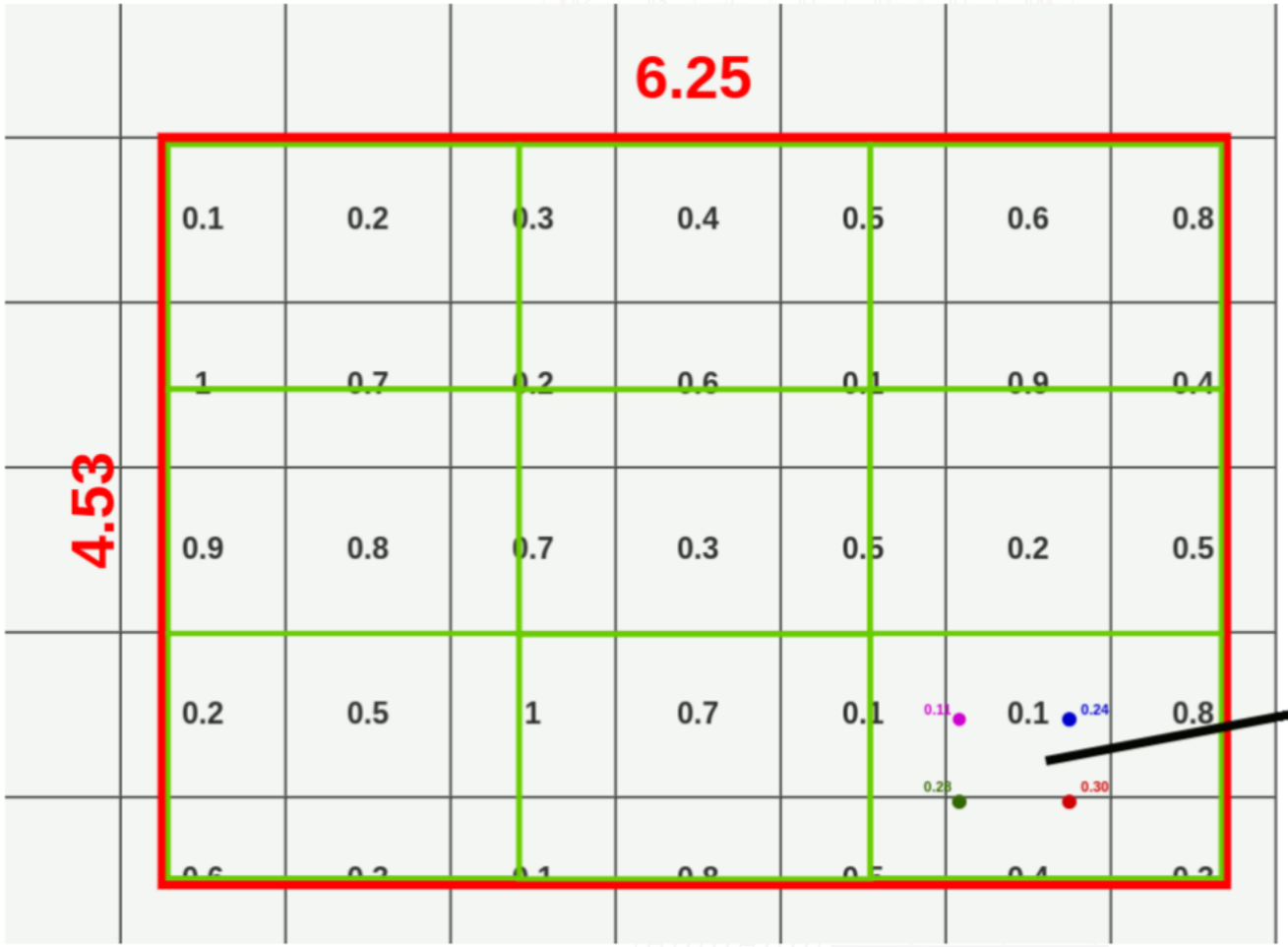
3x3 RoIAlign

0.51		

https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

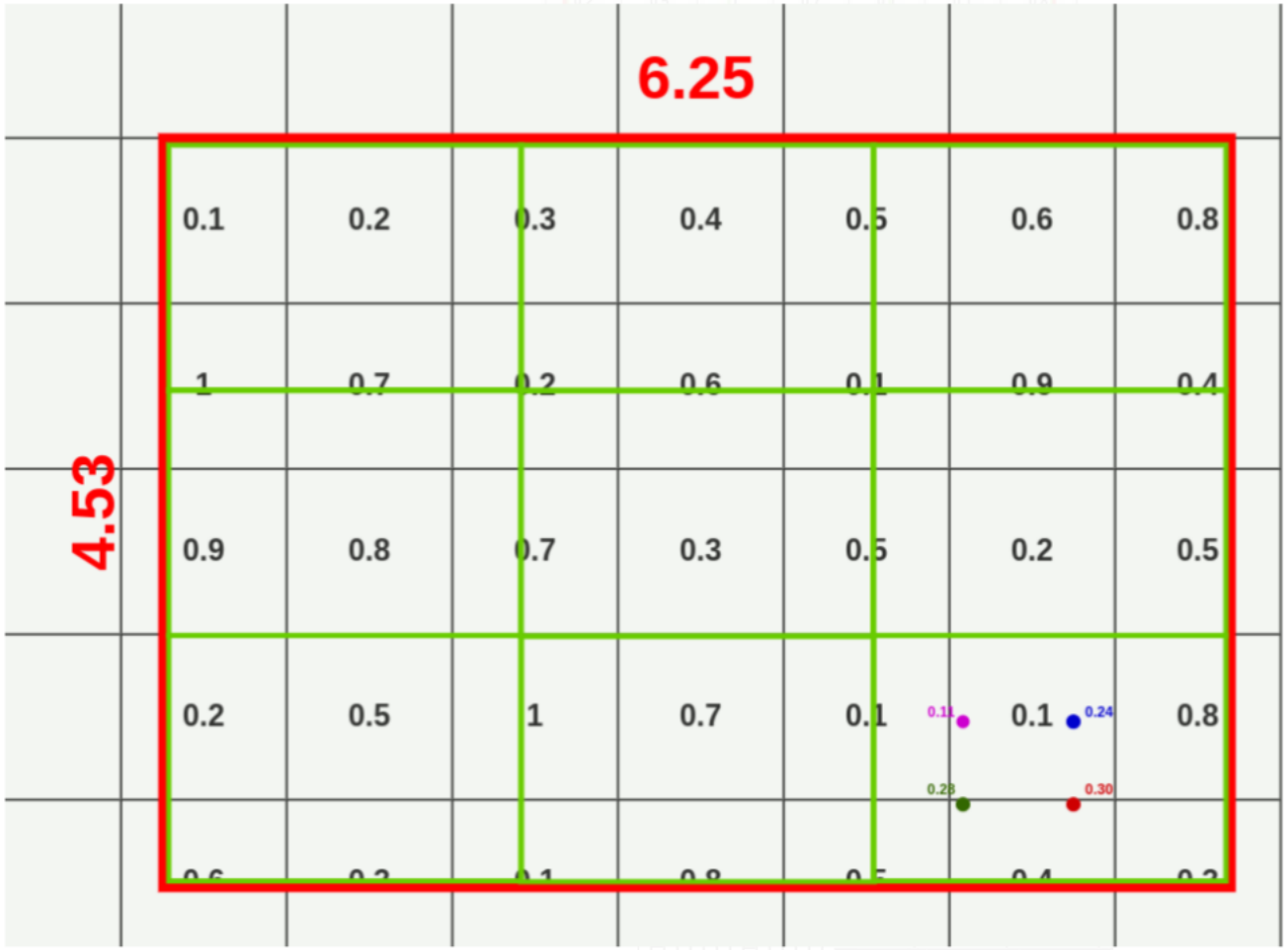

6.25
4.53
0.1 0.2 0.3 0.4 0.5 0.6 0.8
0.35
0.42
0.38
0.46
1 0.7 0.2 0.6 0.1 0.9 0.4
0.9 0.8 0.7 0.3 0.5 0.2 0.5
0.2 0.5 1 0.7 0.1 0.1 0.8
0.6 0.3 0.1 0.8 0.5 0.4 0.3
3x3 RoIAlign
0.51


6.25
4.53
0.1 0.2 0.3 0.4 0.5 0.6 0.8
1 0.7 0.2 0.6 0.1 0.9 0.4
0.9 0.8 0.7 0.3 0.5 0.2 0.5
0.2 0.5 1 0.7 0.1 0.1 0.8
0.55
0.66
0.57
0.71
3x3 RoIAlign
0.51 0.46



3x3 RoIAlign

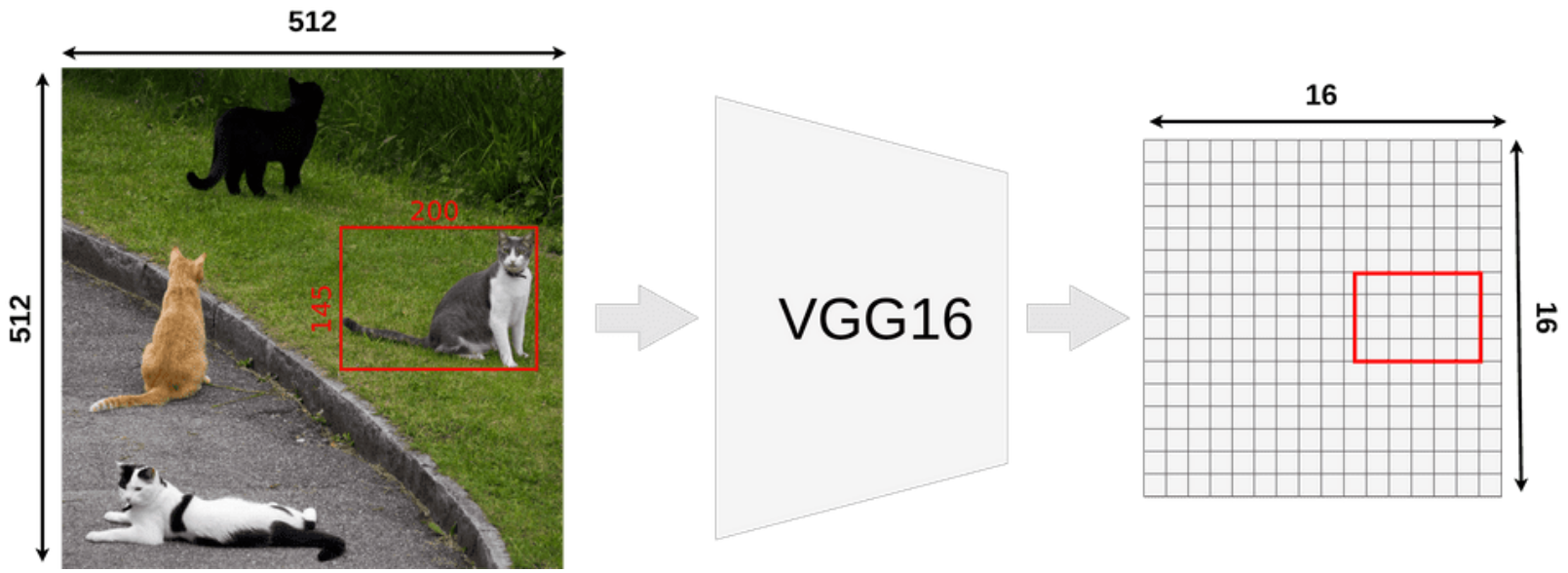
0.51	0.46	0.71
0.86	0.50	0.52
0.56	0.83	

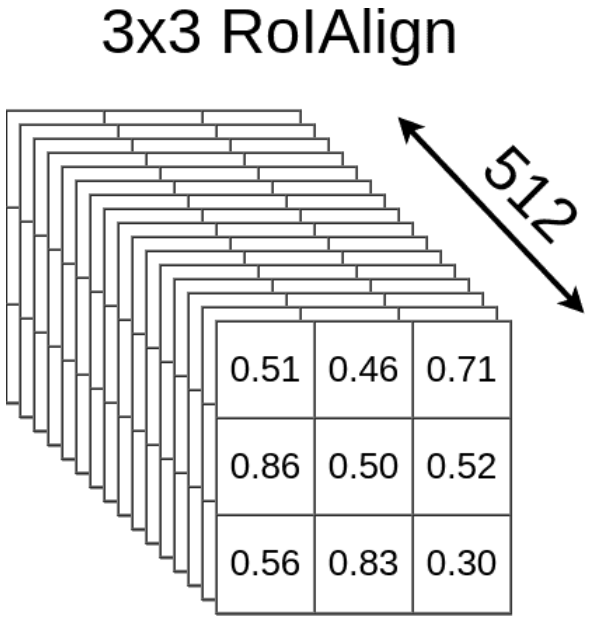


3x3 RoIAlign

0.51	0.46	0.71
0.86	0.50	0.52
0.56	0.83	0.30

RoI-Align

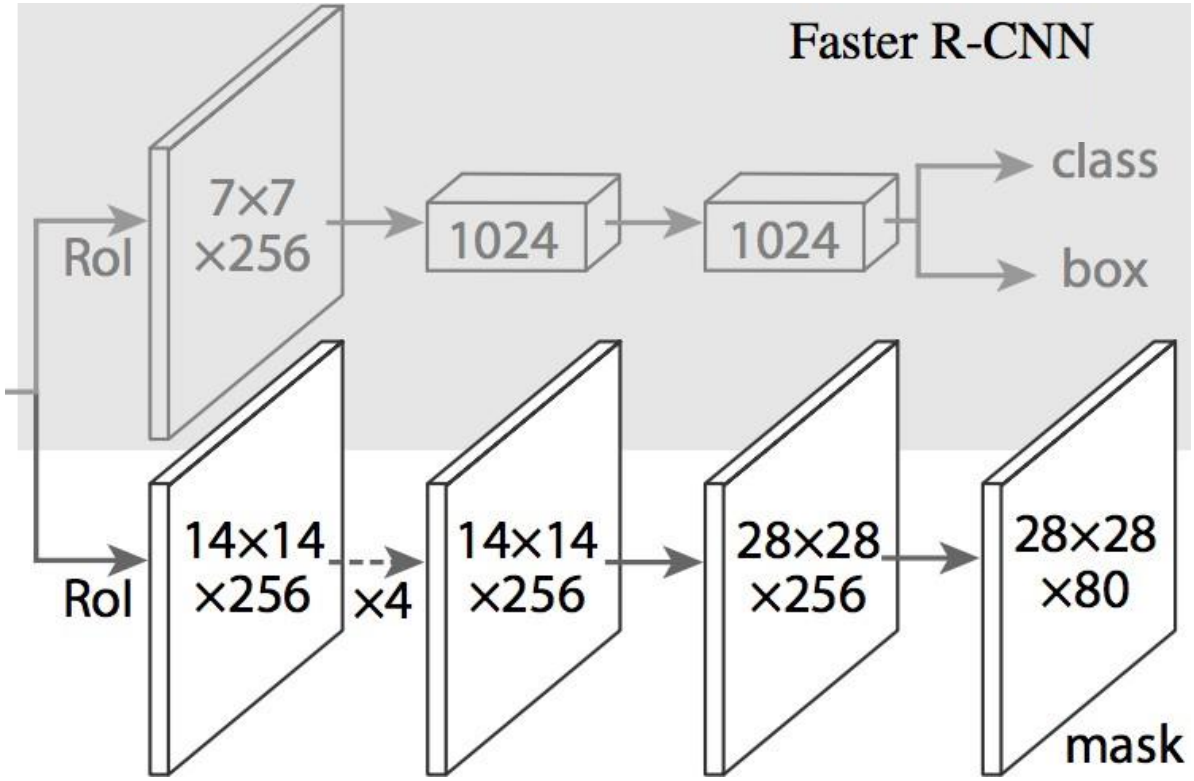




https://towardsdatascience.com/understanding-region-of-interest-part-2-roi-align-and-roi-warp-f795196fc193

Mask R-CNN

- From RoIAlign features, predict class label, bounding box, and segmentation mask

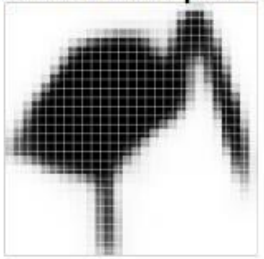


K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017 (Best Paper Award)

Mask R-CNN

Validation image with box detection shown in red

28x28 soft prediction



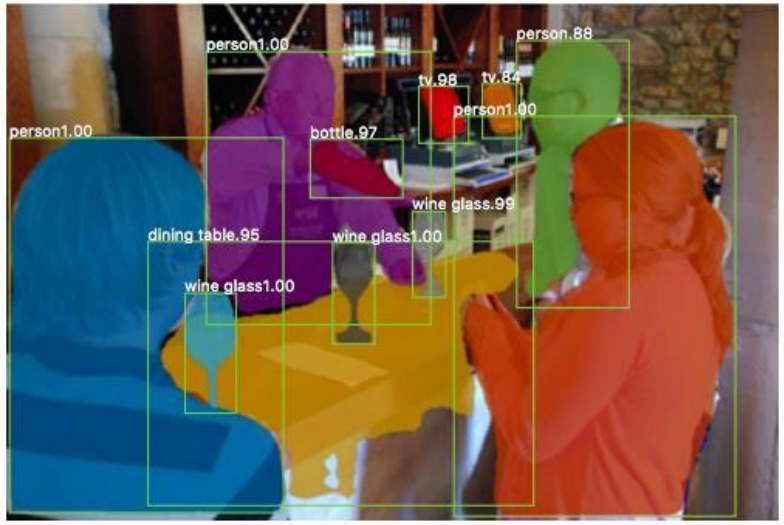
Resized Soft prediction

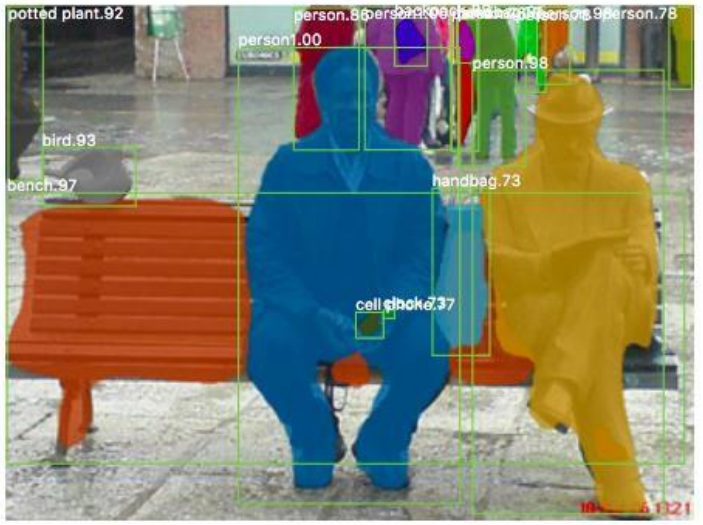
Final mask

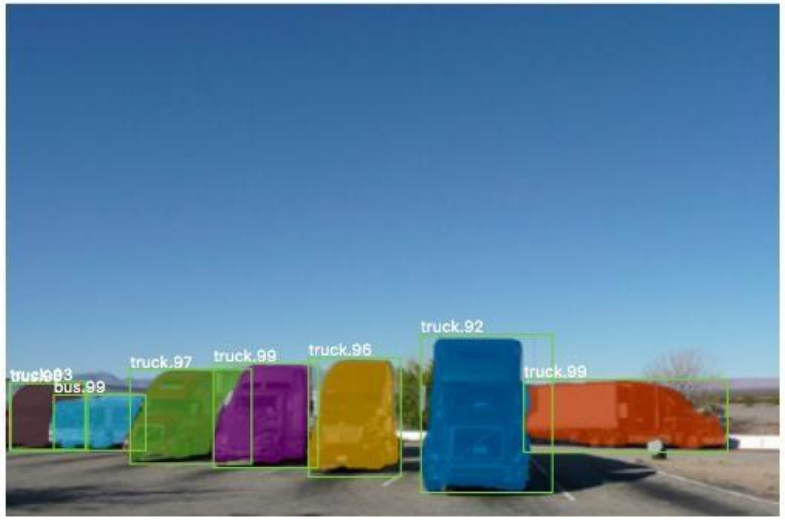
K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN, ICCV 2017 (Best Paper Award)

Example results

Example results







Questions?

Sources for this lecture include materials from works by Sedat Ozer, Ulas Bagci, and Svetlana Lazebnik