

Computer Vision with Embedded Machine Learning

Object Detection Performance
Metrics

Confusion Matrix

		Predicted Label		
Actual Label		Ball	Dog	Toy
	Ball	205	10	1
	Dog	6	199	0
	Toy	9	17	223

Precision, Positive Predictive Value (PPV):
What proportion of positive predictions was actually correct?

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP} = \frac{\text{Green Box}}{\text{Green Box} + \text{Orange Box}} \\
 &= \frac{199}{199 + 27} = 0.881
 \end{aligned}$$

Confusion Matrix

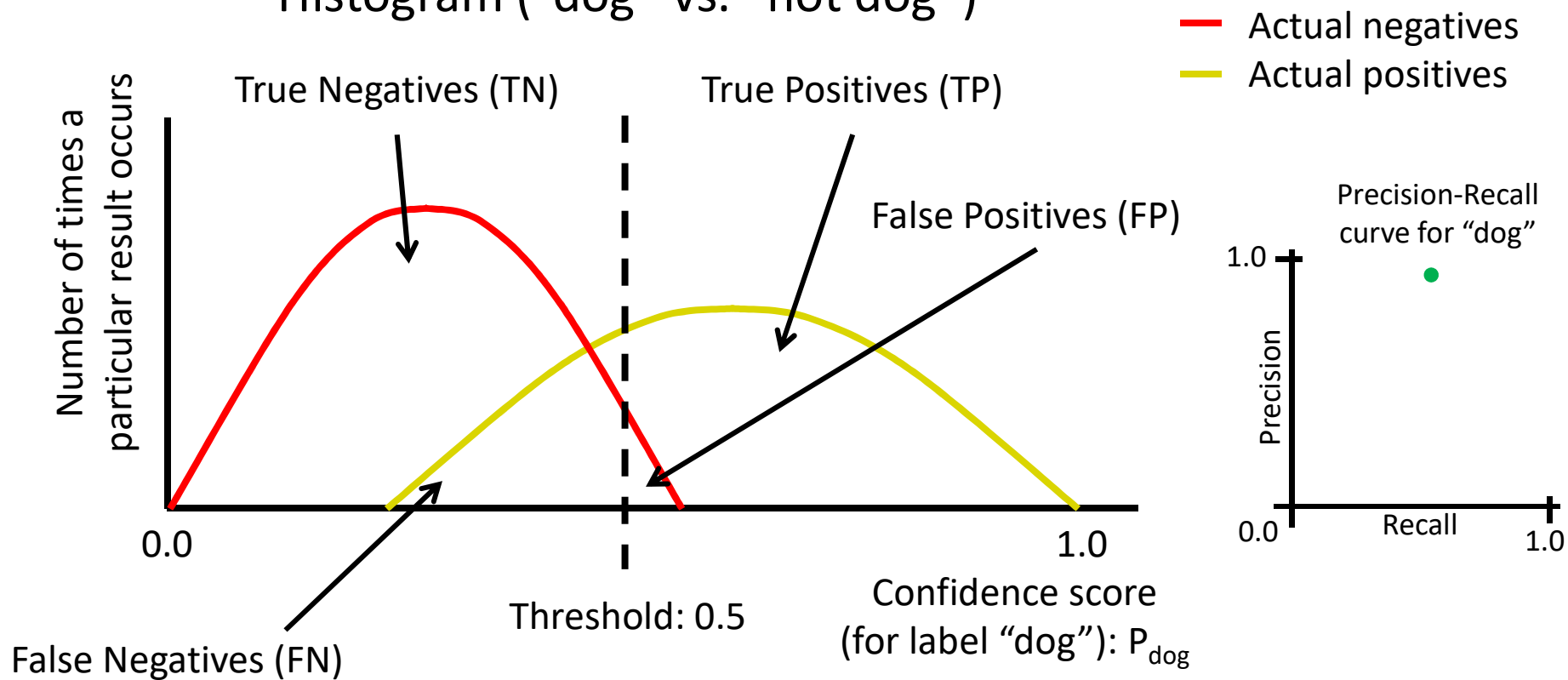
		Predicted Label		
Actual Label		Ball	Dog	Toy
	Ball	205	10	1
	Dog	6	199	0
	Toy	9	17	223

Recall, True Positive Rate (TPR):

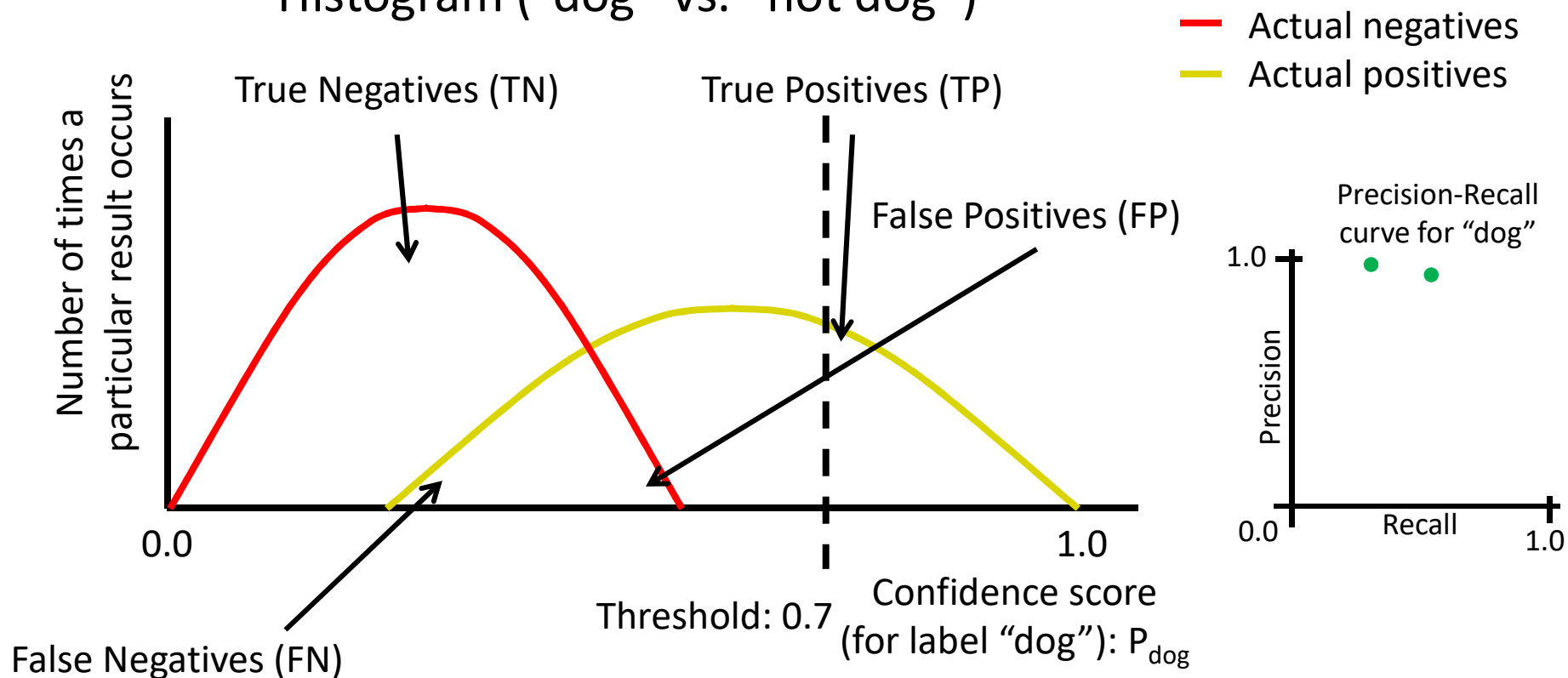
What proportion of actual positives was identified correctly?

$$\begin{aligned}
 \text{Recall} &= \frac{TP}{TP + FN} = \frac{\text{Green Box}}{\text{Green Box} + \text{Purple Box}} \\
 &= \frac{199}{199 + 6} = 0.971
 \end{aligned}$$

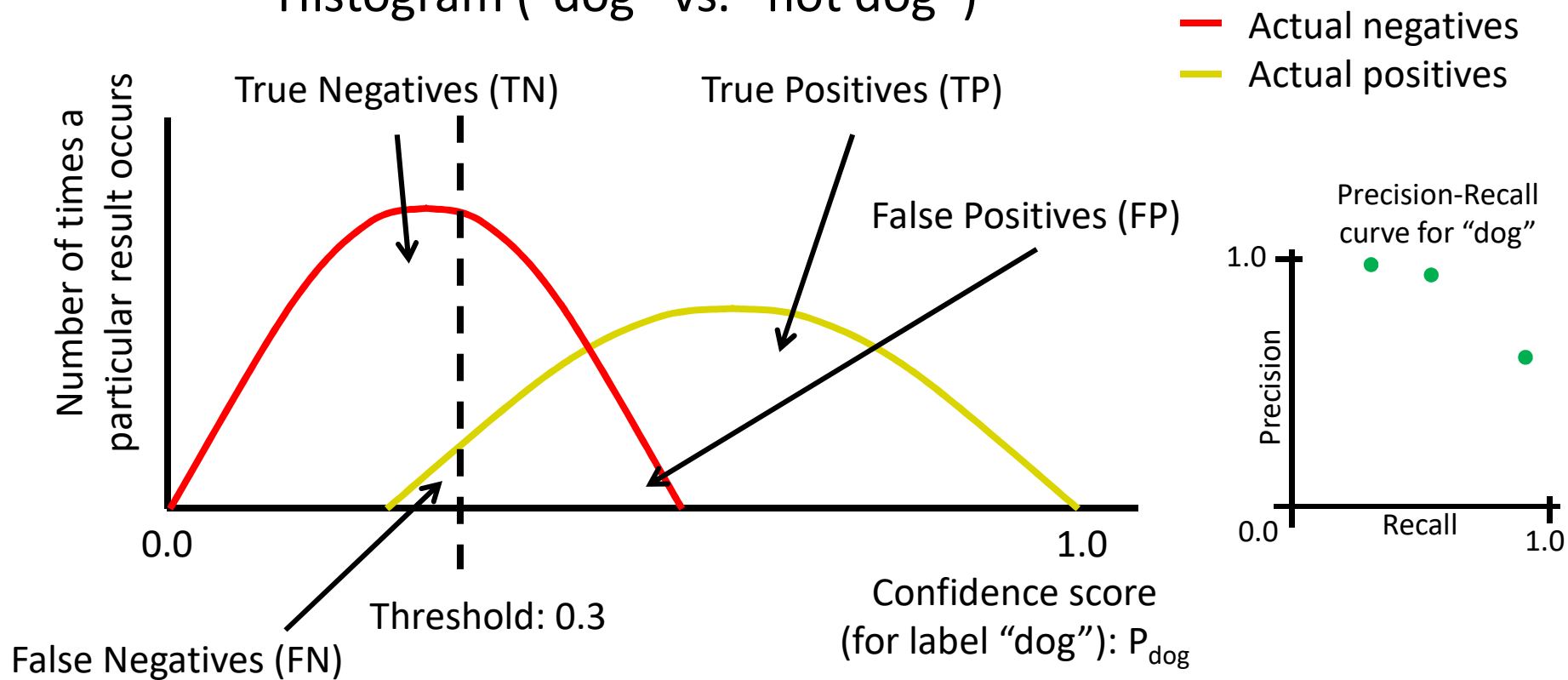
Histogram ("dog" vs. "not dog")



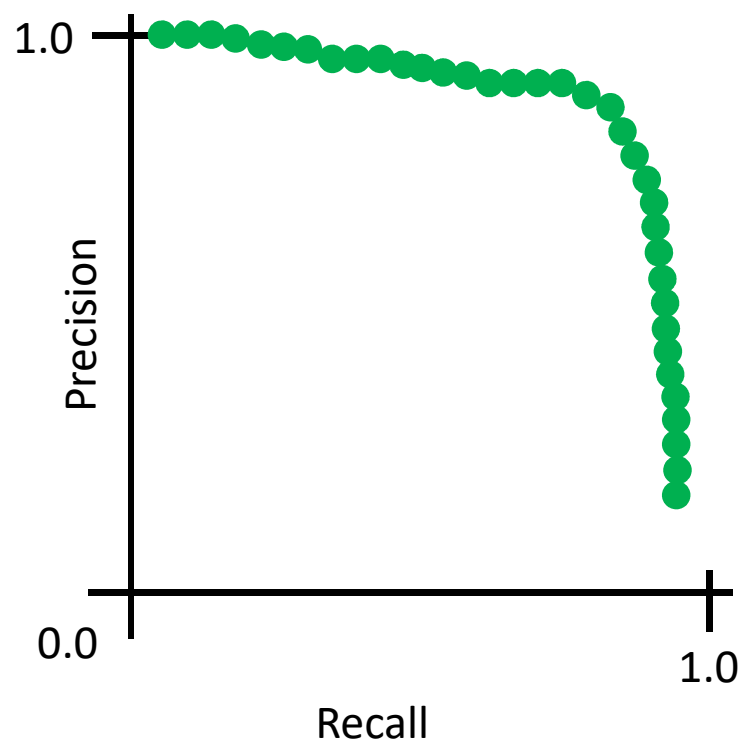
Histogram ("dog" vs. "not dog")



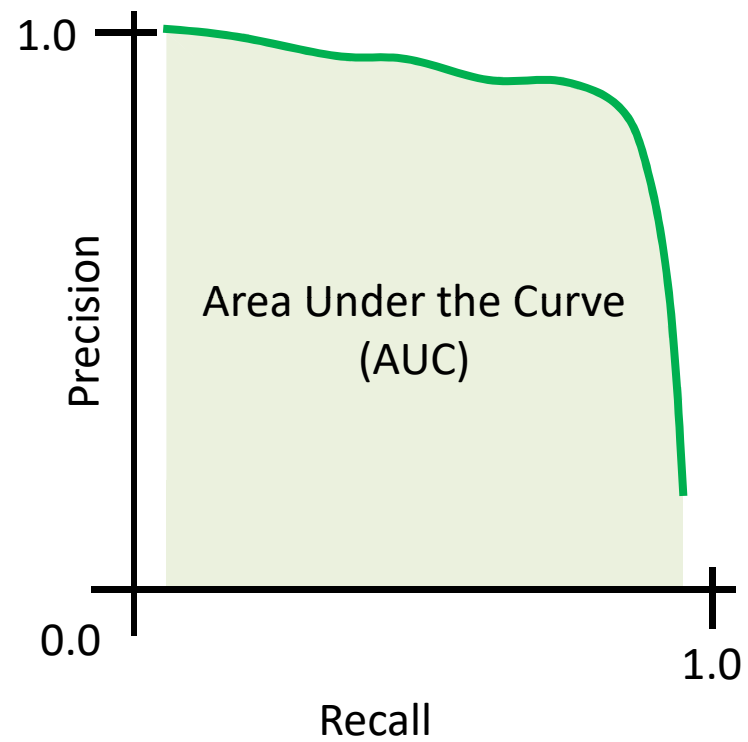
Histogram ("dog" vs. "not dog")



Precision-Recall curve for “dog”

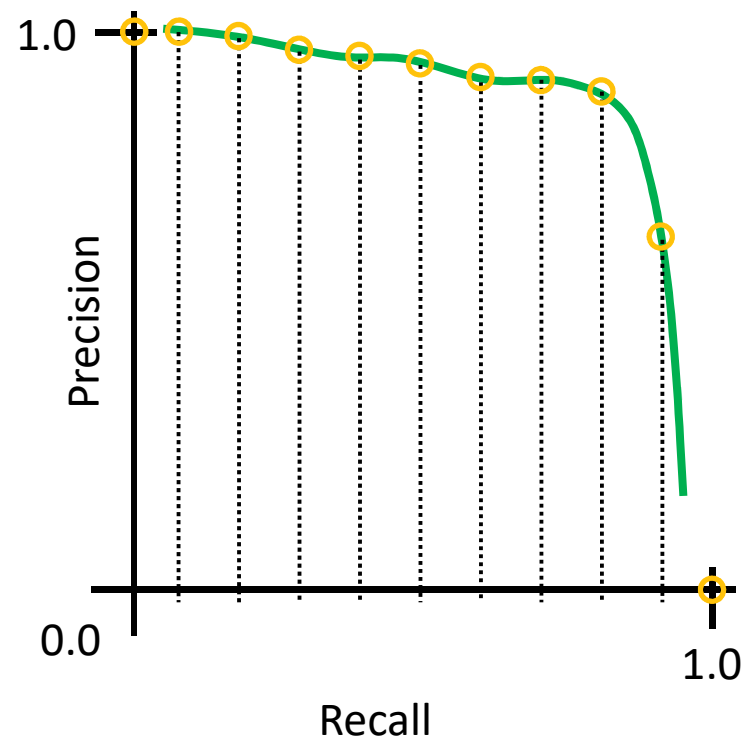


Average Precision (AP)



$$AP = \int_0^1 p(r) dr$$

Average Precision (AP) approximation

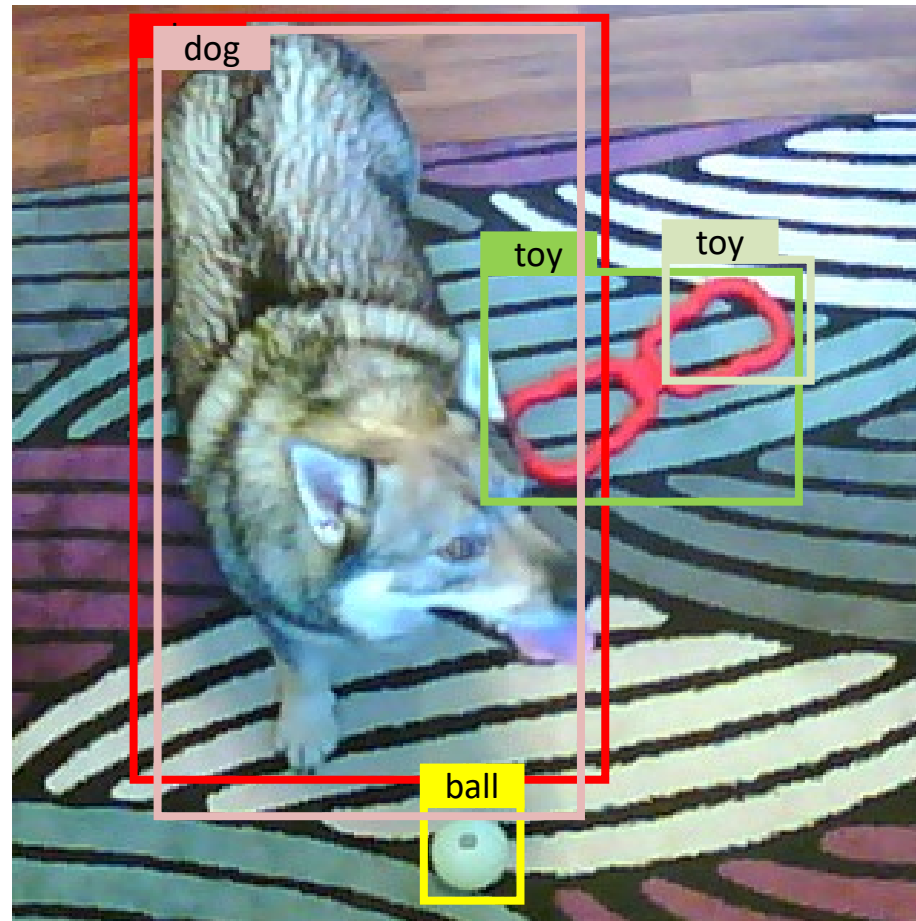


PASCAL Visual Object Classes
(VOC) Challenge (2009):

$$AP = \frac{1}{11} \sum_{r_i} p(r_i)$$

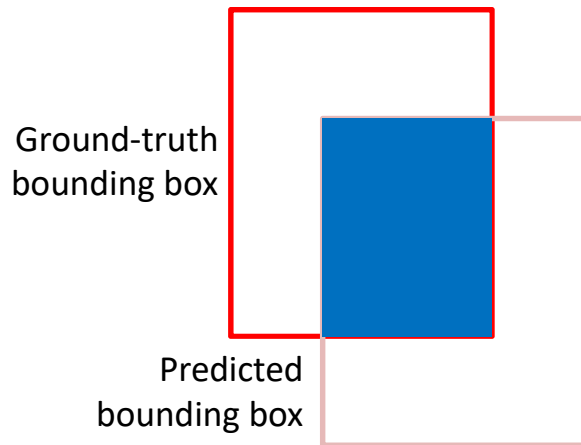
where $r_i = [0.0, 0.1, 0.2, \dots, 1.0]$

- **Object 1**
 - Class: dog
 - Bounding box
 - (x_1, y_1)
 - (w_1, h_1)
- **Object 2**
 - Class: toy
 - Bounding box
 - (x_2, y_2)
 - (w_2, h_2)
- **Object 3**
 - Class: ball
 - Bounding box
 - (x_3, y_3)
 - (w_3, h_3)



Intersection over Union (IoU)

Intersection: Area of overlap



Union: Area encompassed by both boxes



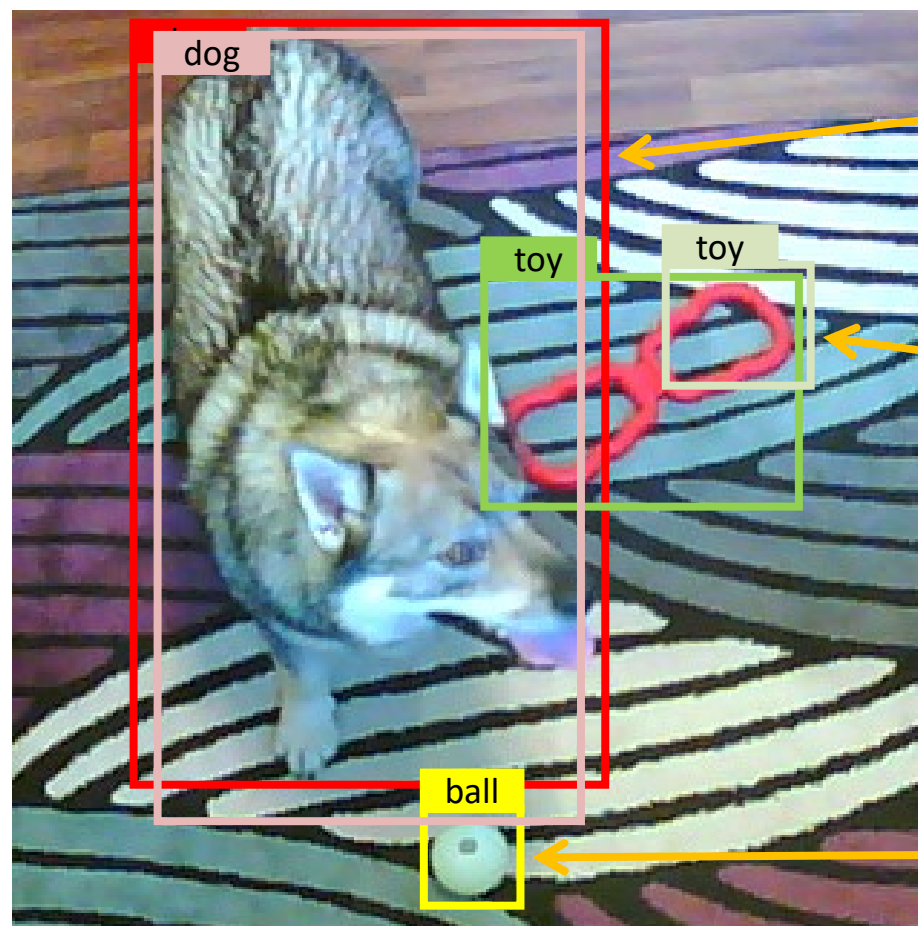
$$IoU = \frac{Intersection}{Union} = \frac{area\ of\ overlap}{area\ of\ union}$$

0.0 = complete miss

1.0 = perfect match

Object: IoU \geq 0.5

- **Object 1**
 - Class: dog
 - Bounding box
 - (x_1, y_1)
 - (w_1, h_1)
- **Object 2**
 - Class: toy
 - Bounding box
 - (x_2, y_2)
 - (w_2, h_2)
- **Object 3**
 - Class: ball
 - Bounding box
 - (x_3, y_3)
 - (w_3, h_3)



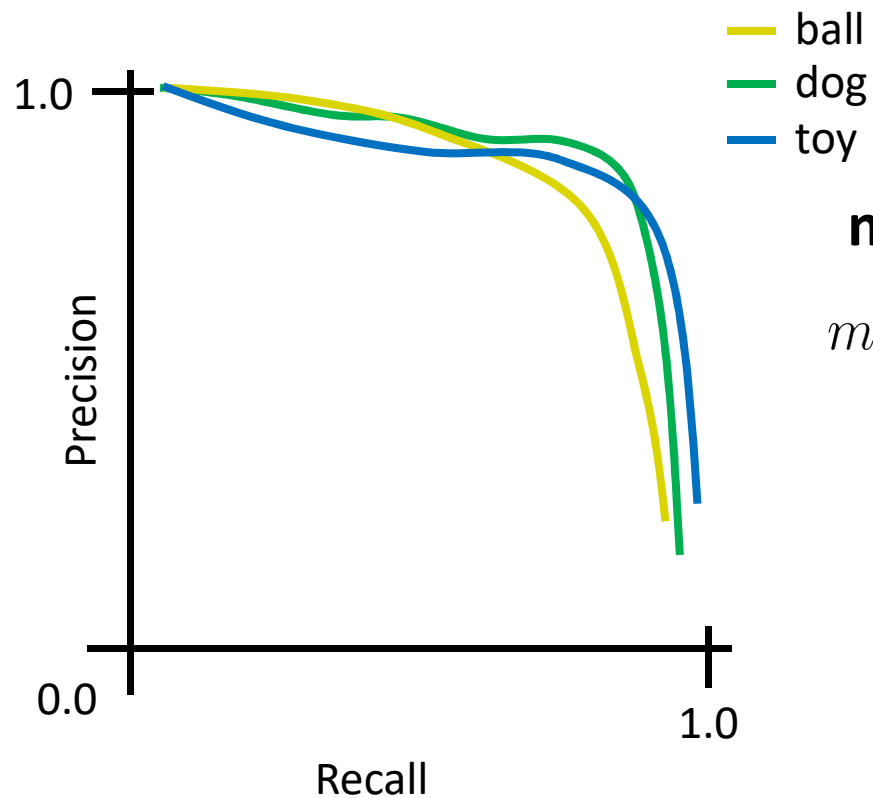
With $Th_{IoU} = 0.5$:

$IoU = 0.9$
True Positive (TP)

$IoU = 0.2$
False Positive (FP)

False Negative (FN)

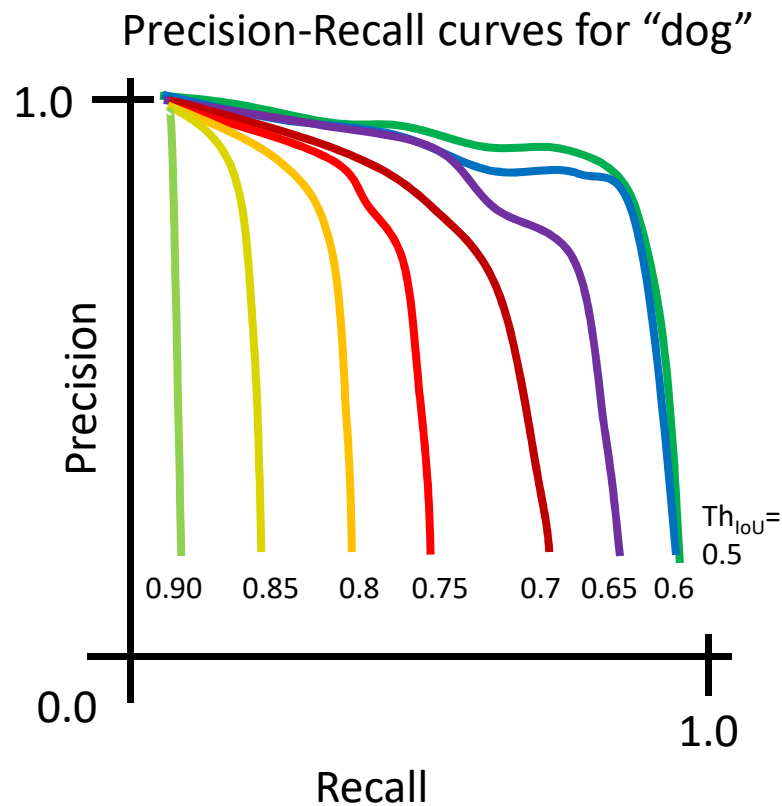
Precision-Recall curve (for IoU ≥ 0.5)



mean Average Precision (mAP)

$$mAP_{0.5} = \frac{1}{3}(AP_{ball} + AP_{dog} + AP_{toy})$$

mAP (for several IoU thresholds)



mean Average Precision (mAP)

$$mAP_{0.5} = \frac{1}{3}(AP_{ball} + AP_{dog} + AP_{toy})$$

COCO 2017 Challenge:

$$mAP = \frac{1}{10} \sum_i mAP_i$$

where $i = [0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95]$