

R-CNN & Fast R-CNN

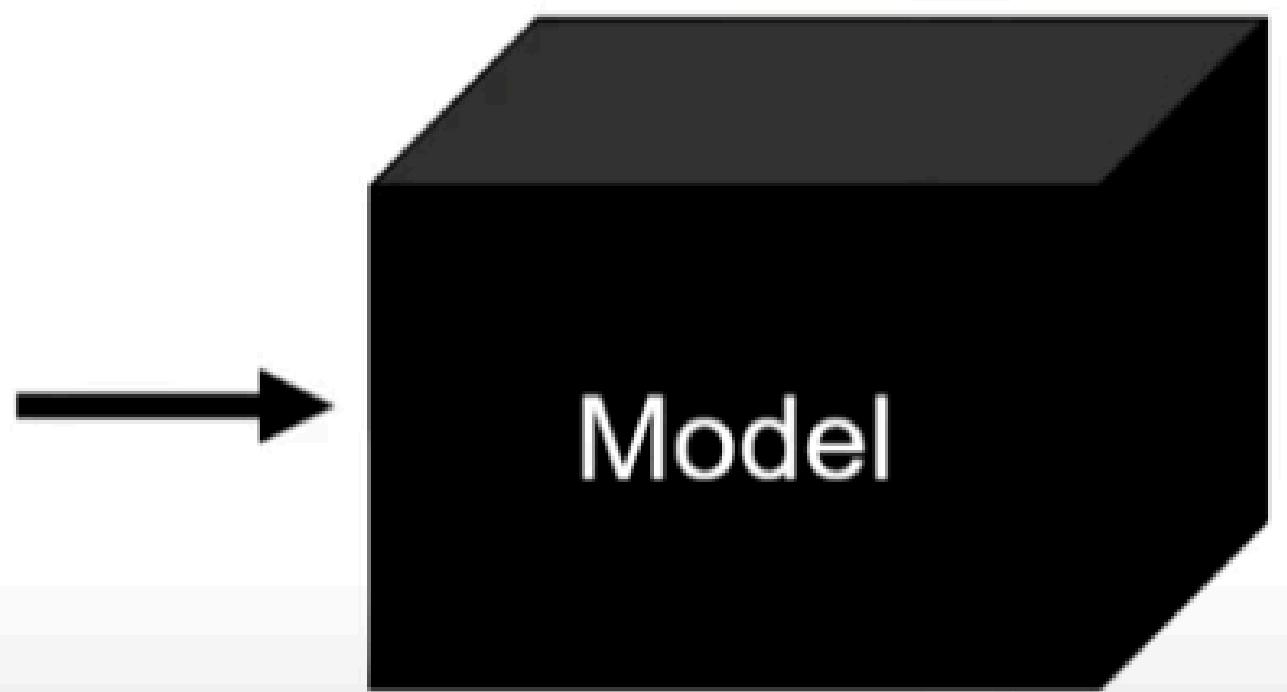
Everything You Need to Know

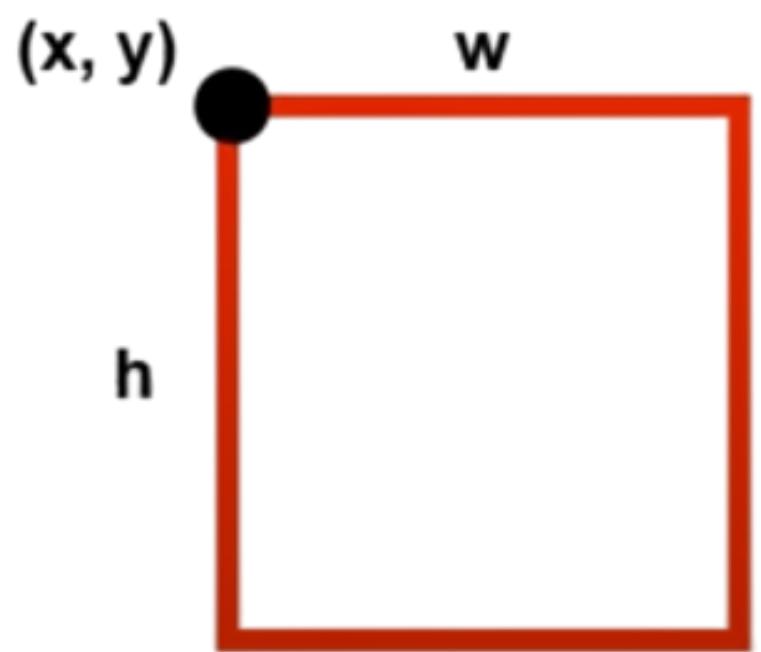
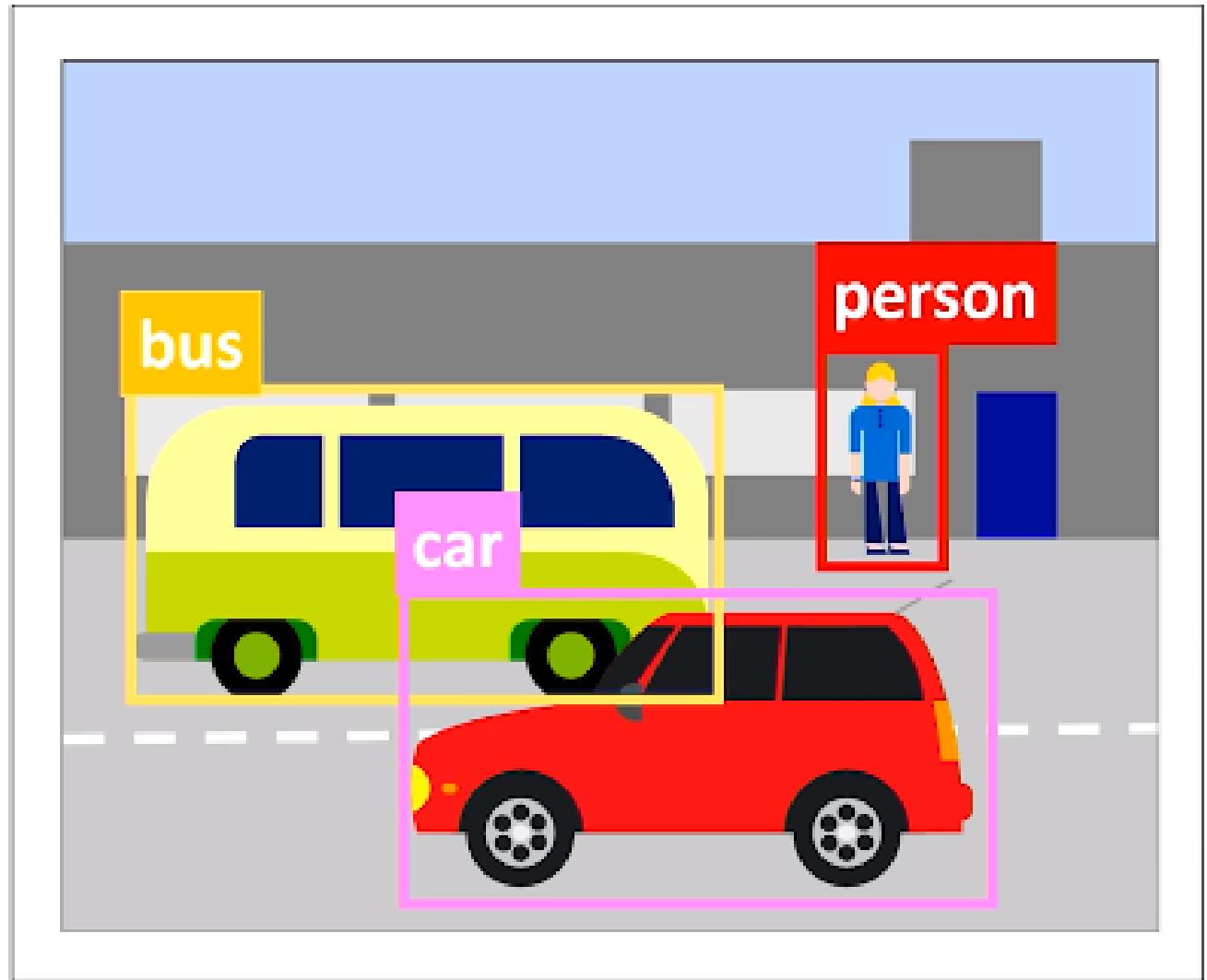
for

Object Detection

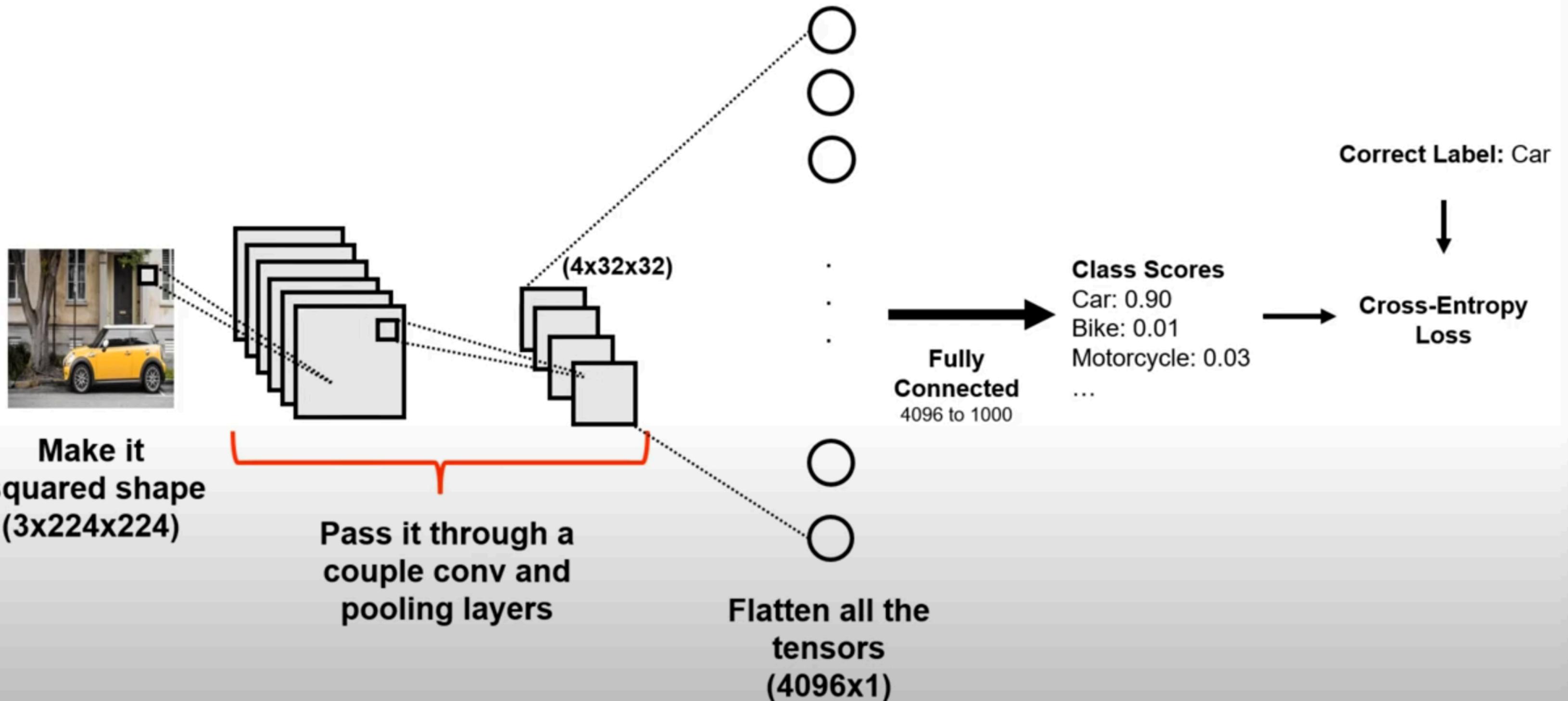
R-CNN:
**(Region-based Convolutional Neural
Networks)**

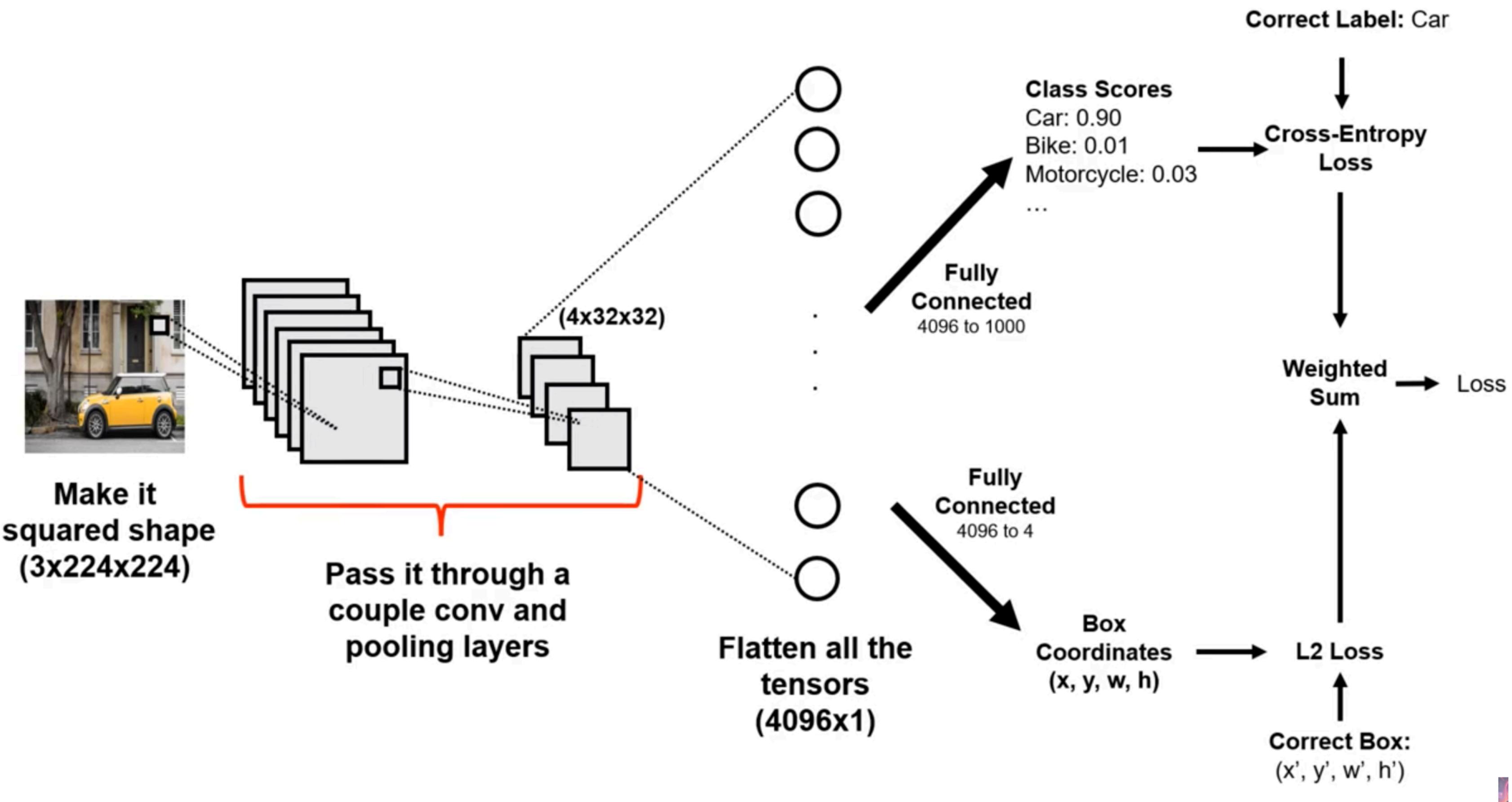
Our Goal in OD





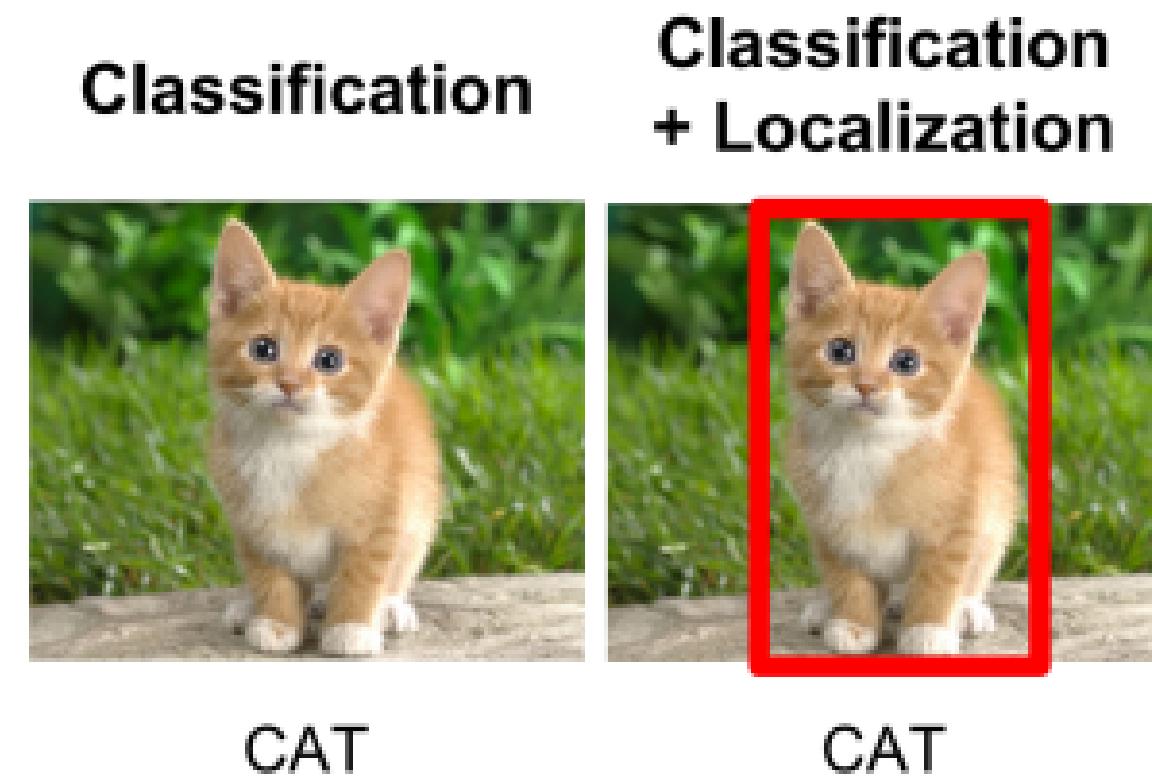
Bounding box





Disadvantage : Only one object can be detected

:(
:(



Note : This Architecture is used for
Object Localization

Old Approach: Sliding Window



Old Approach: Sliding Window



Classify this region!



Neural Network classifier

Mountain

Old Approach: Sliding Window



Classify this region!



Neural Network classifier

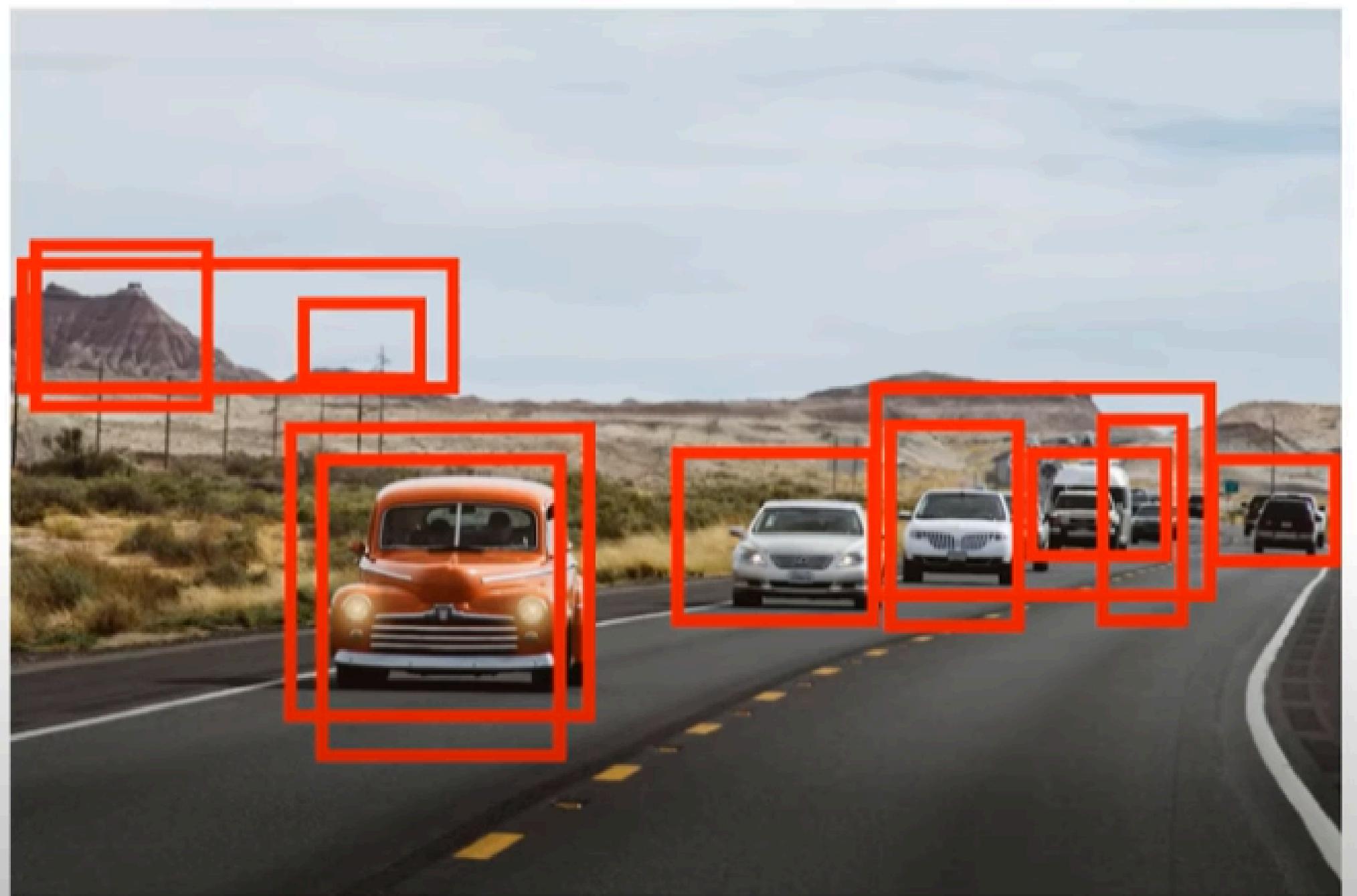
Car

RCNN

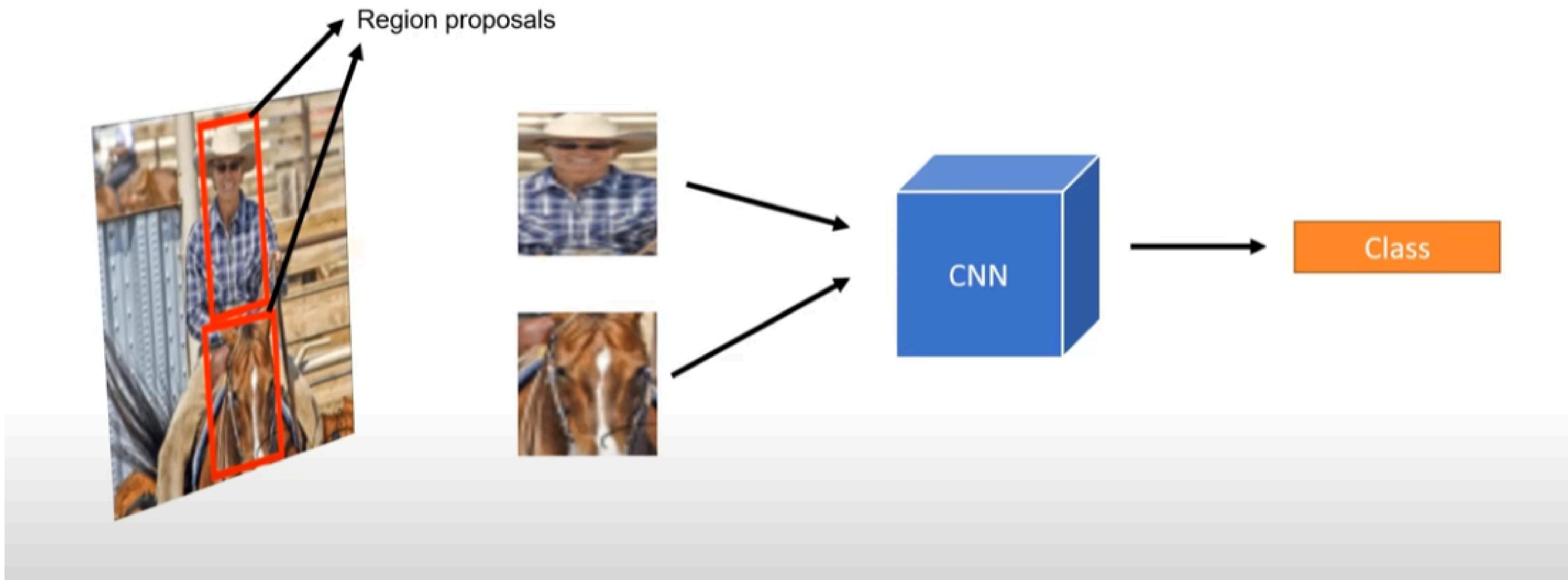
It uses an external algorithm to propose some regions.

Selective Search

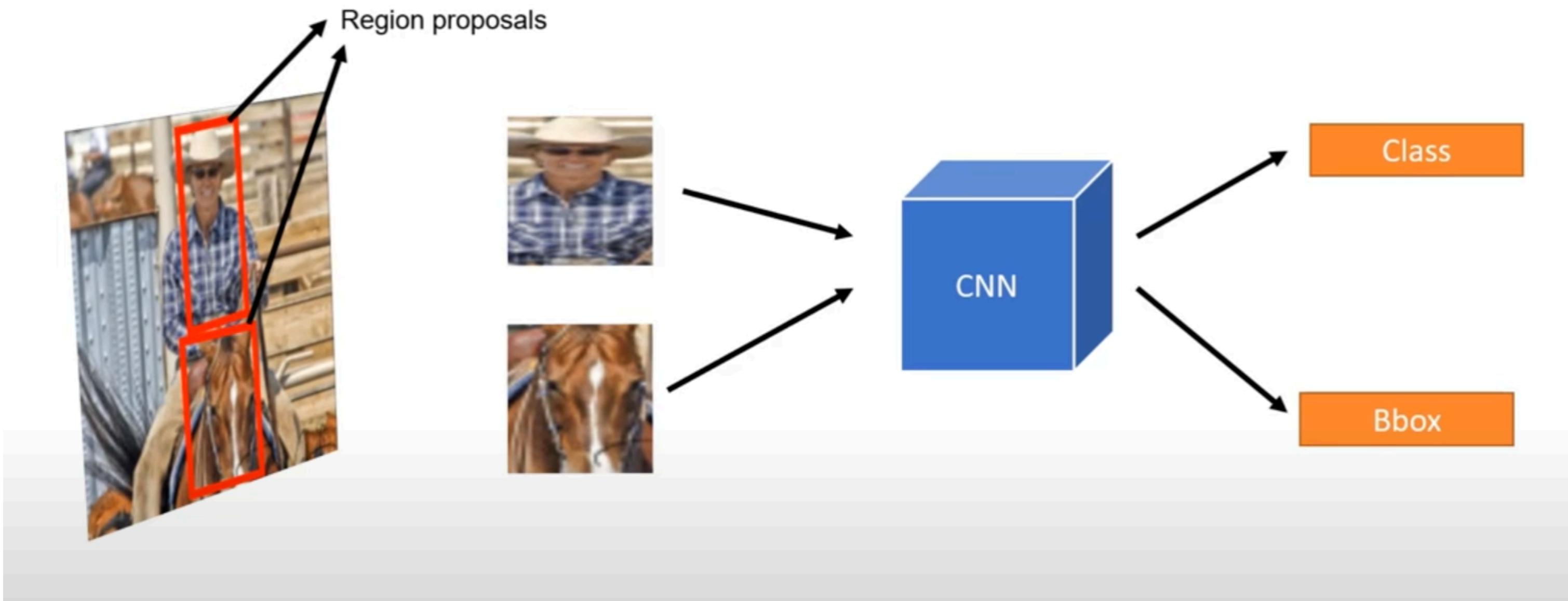
It gives us like 2000 region proposals within one or two seconds



RCNN



RCNN



RCNN

Region proposal: (p_x, p_y, p_h, p_w)



Bbox

Transform: (t_x, t_y, t_h, t_w)

Output: (b_x, b_y, b_h, b_w)



RCNN

Region proposal: (p_x, p_y, p_h, p_w)



Bbox

Transform: (t_x, t_y, t_h, t_w)

Output: (b_x, b_y, b_h, b_w)



Translation:

$$b_x = p_x + p_w t_w$$

(Horizontal translation)

$$b_y = p_y + p_h t_h$$

(Vertical translation)

Log-space scale transform:

$$b_w = p_w \exp(t_w)$$

(Horizontal scale)

$$b_h = p_h \exp(t_h)$$

(Vertical scale)

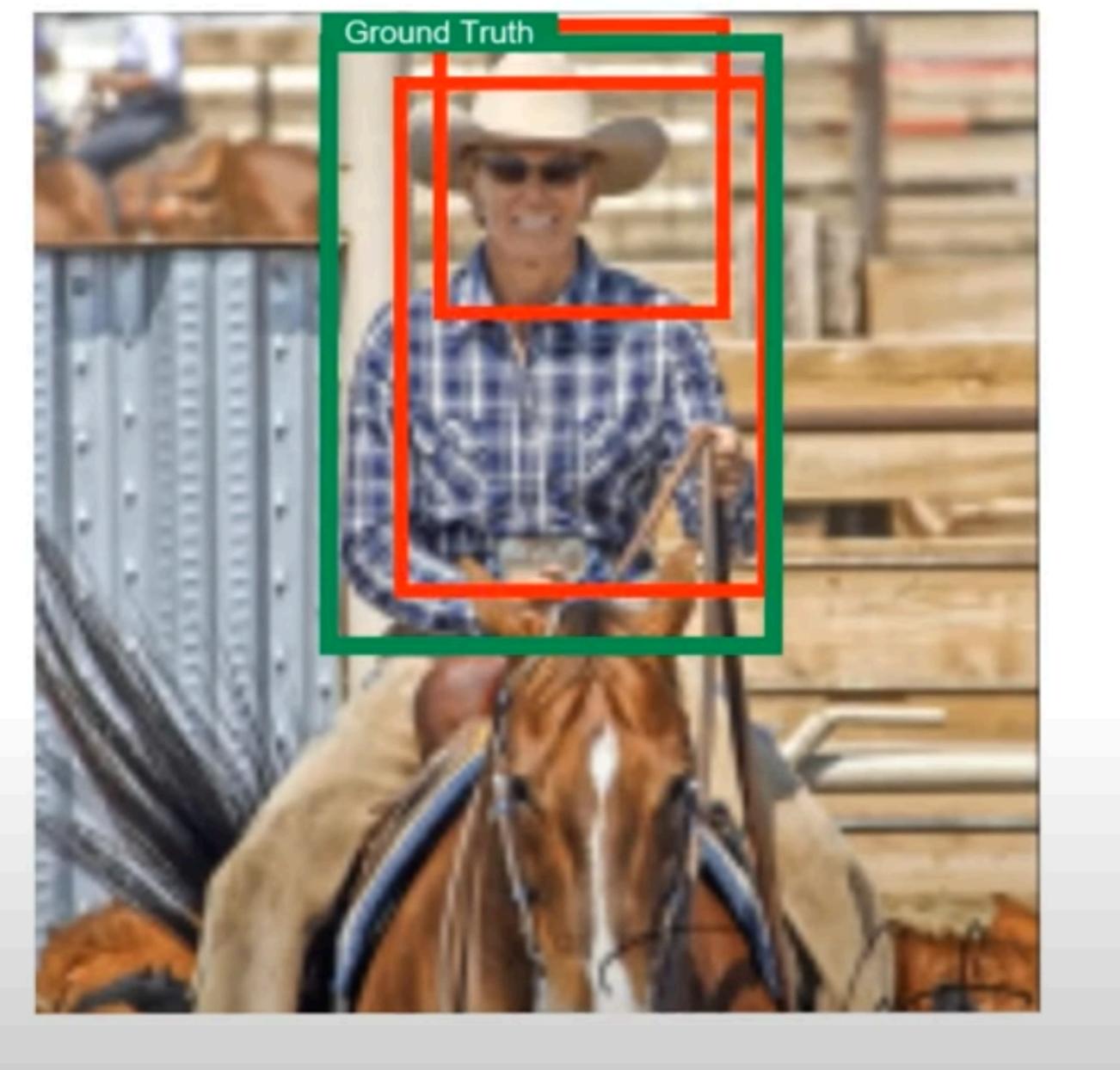
RCNN

What if we output two boxes both pointing to the same object?



Non-Maximum suppression

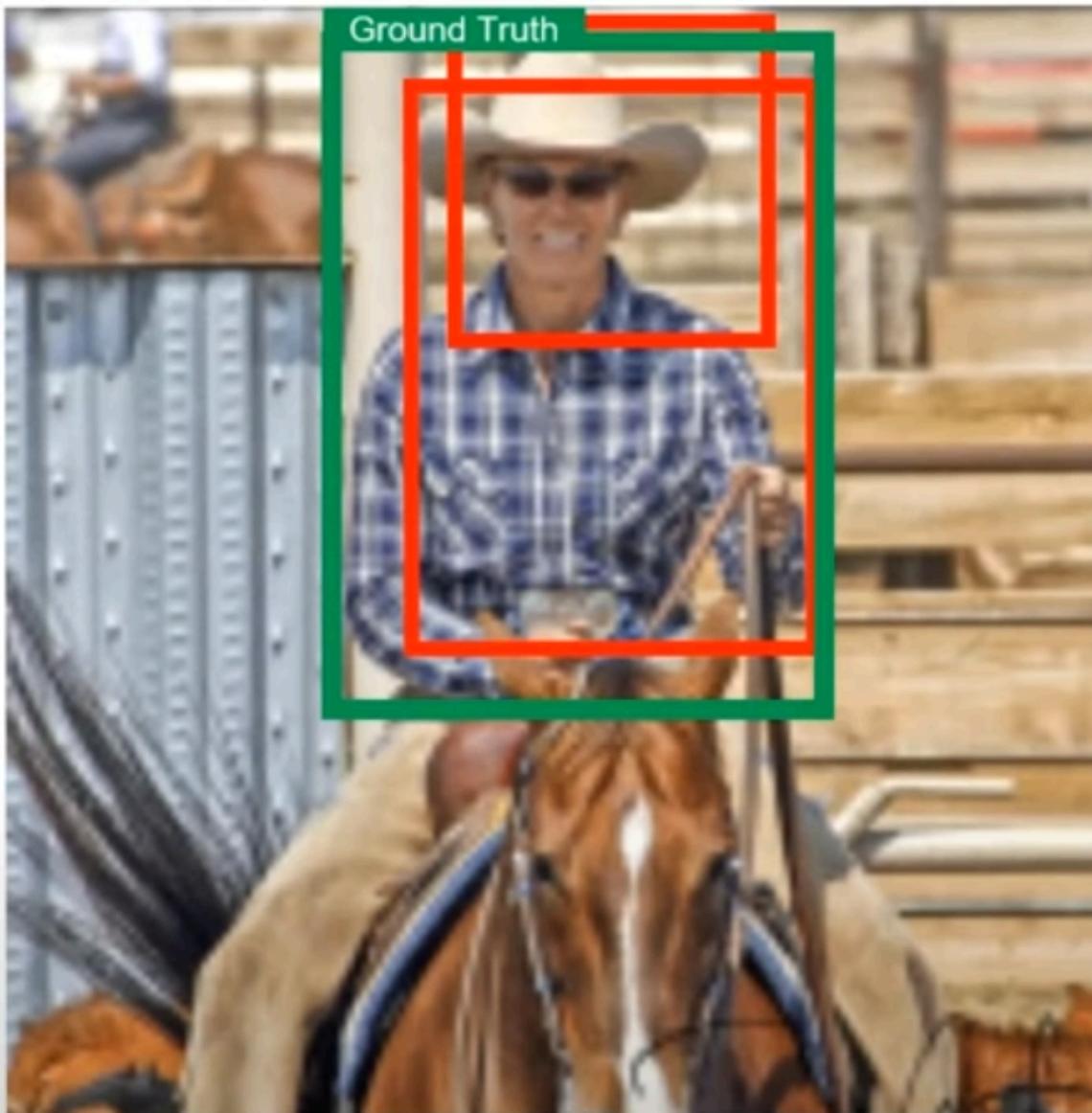
Non-Maximum suppression



Non-Maximum suppression

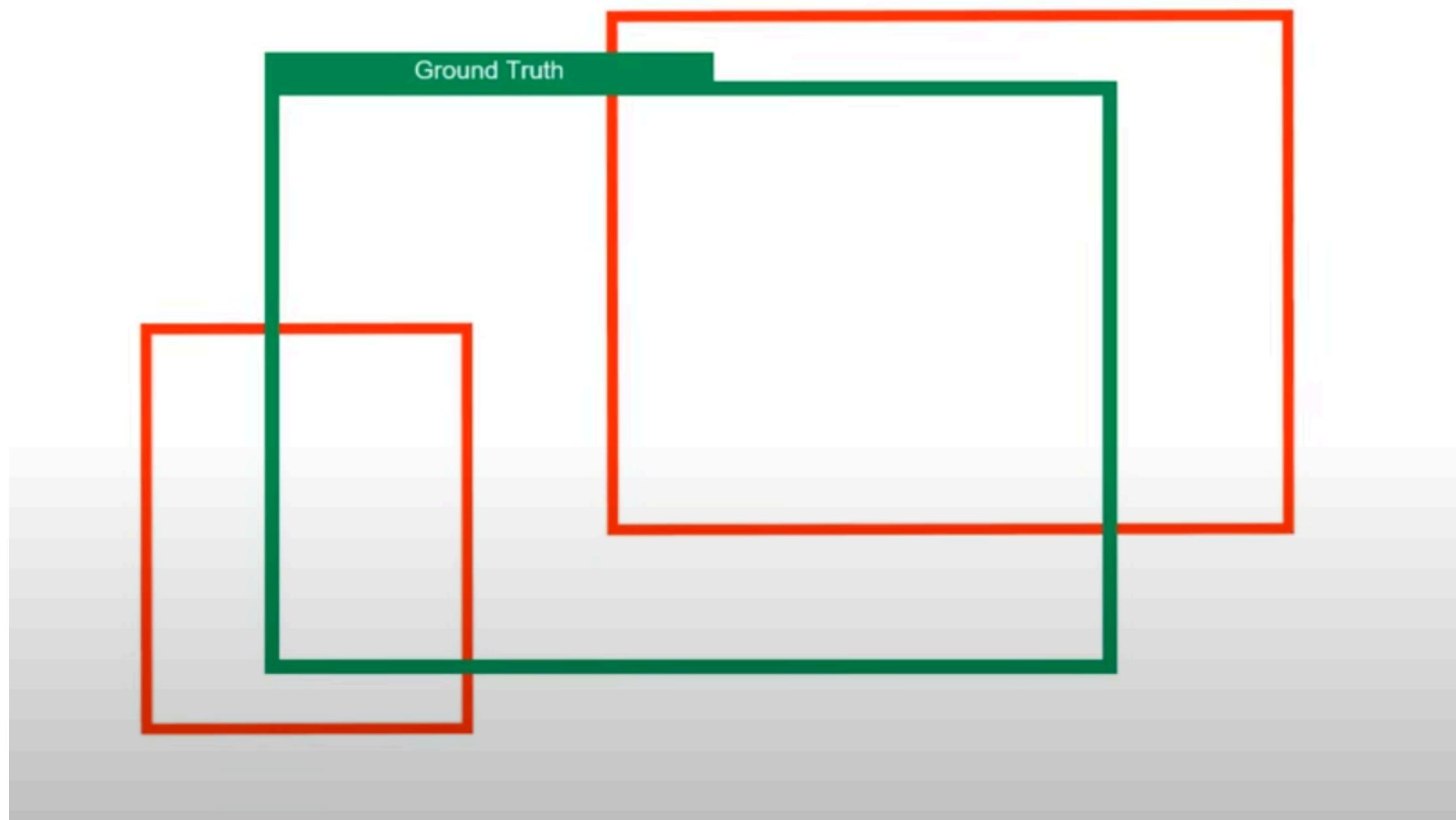
Which one to choose?

Intuitively: The one that is closer to the ground truth

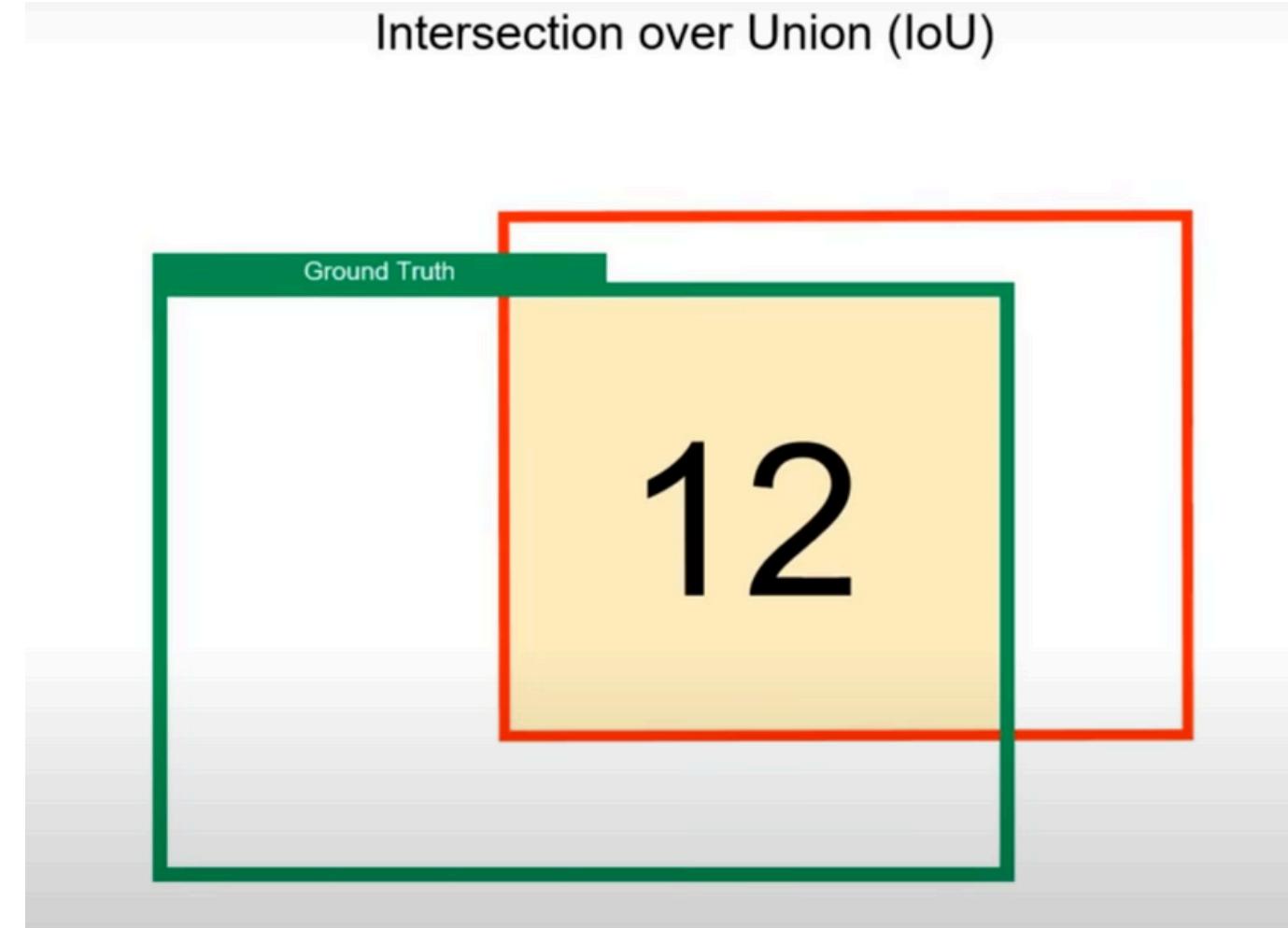


Non-Maximum suppression

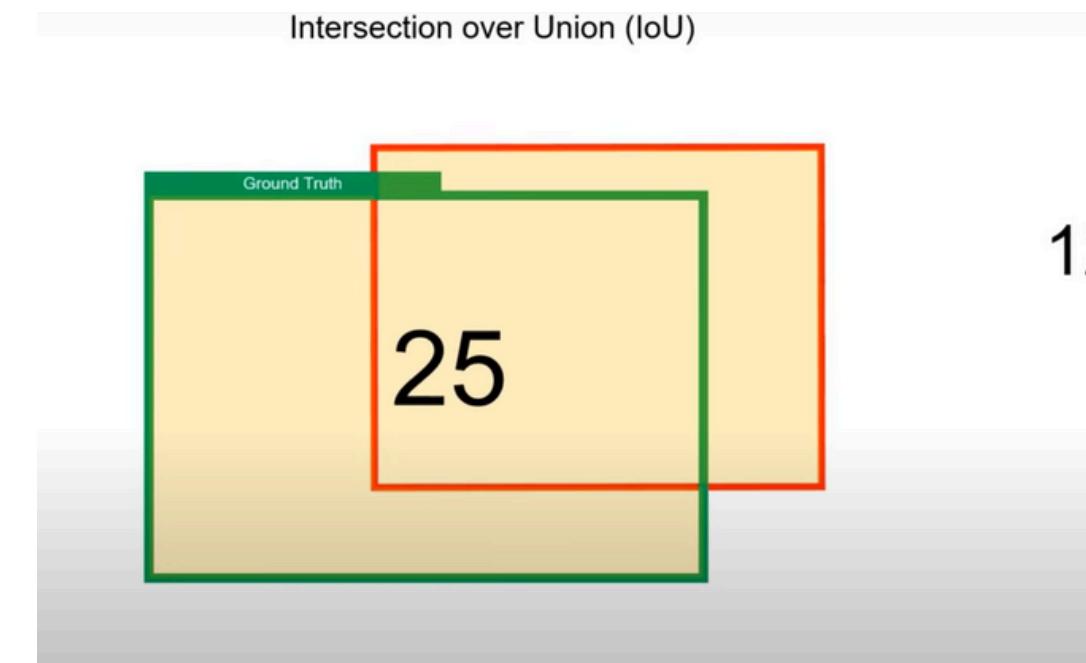
Intersection over Union (IoU)



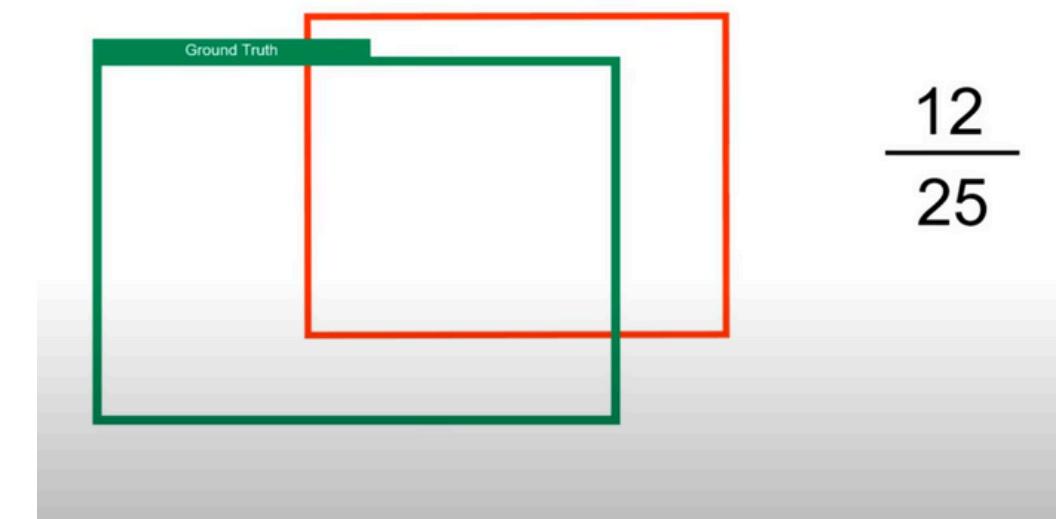
Non-Maximum suppression



Intersection

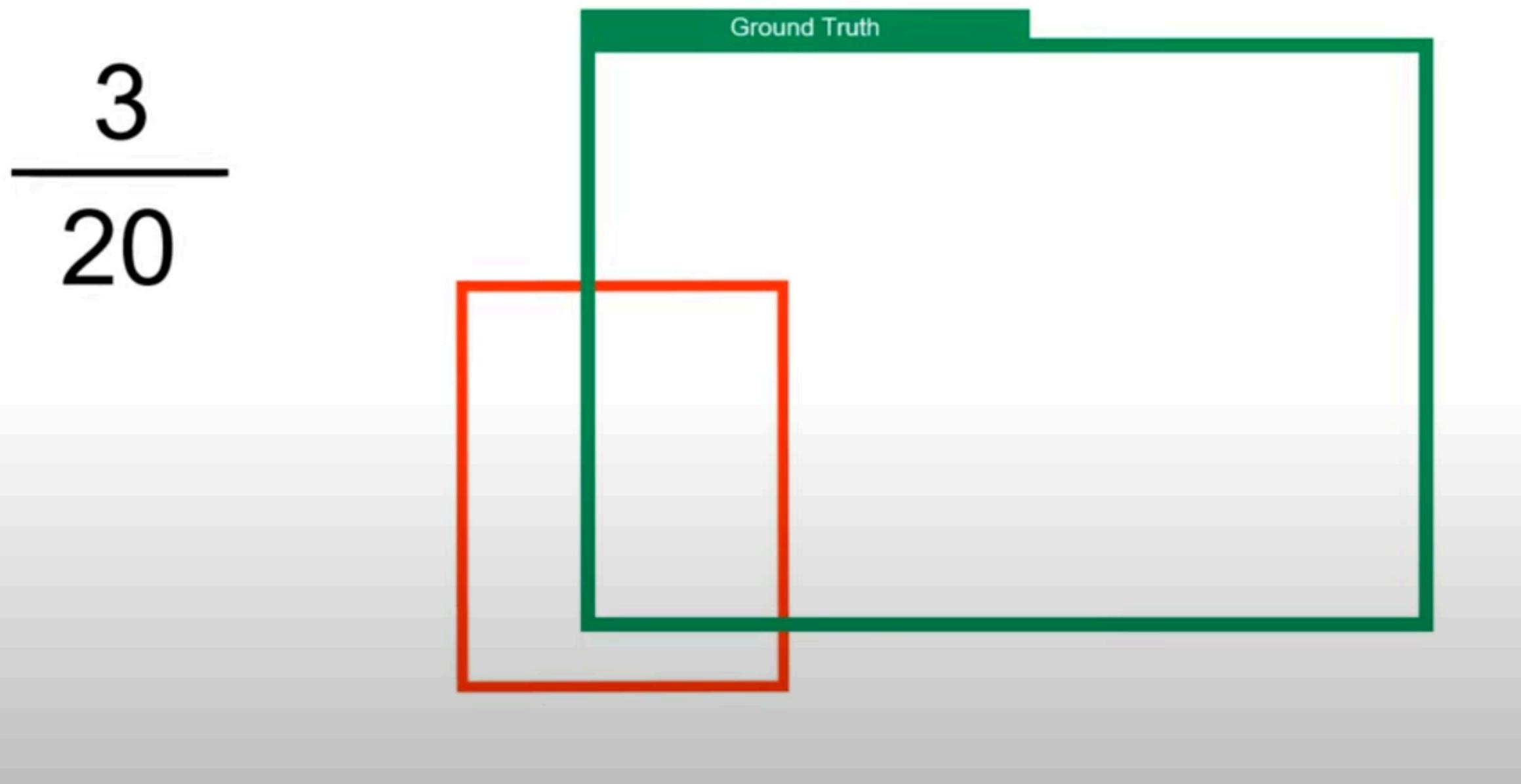


Union

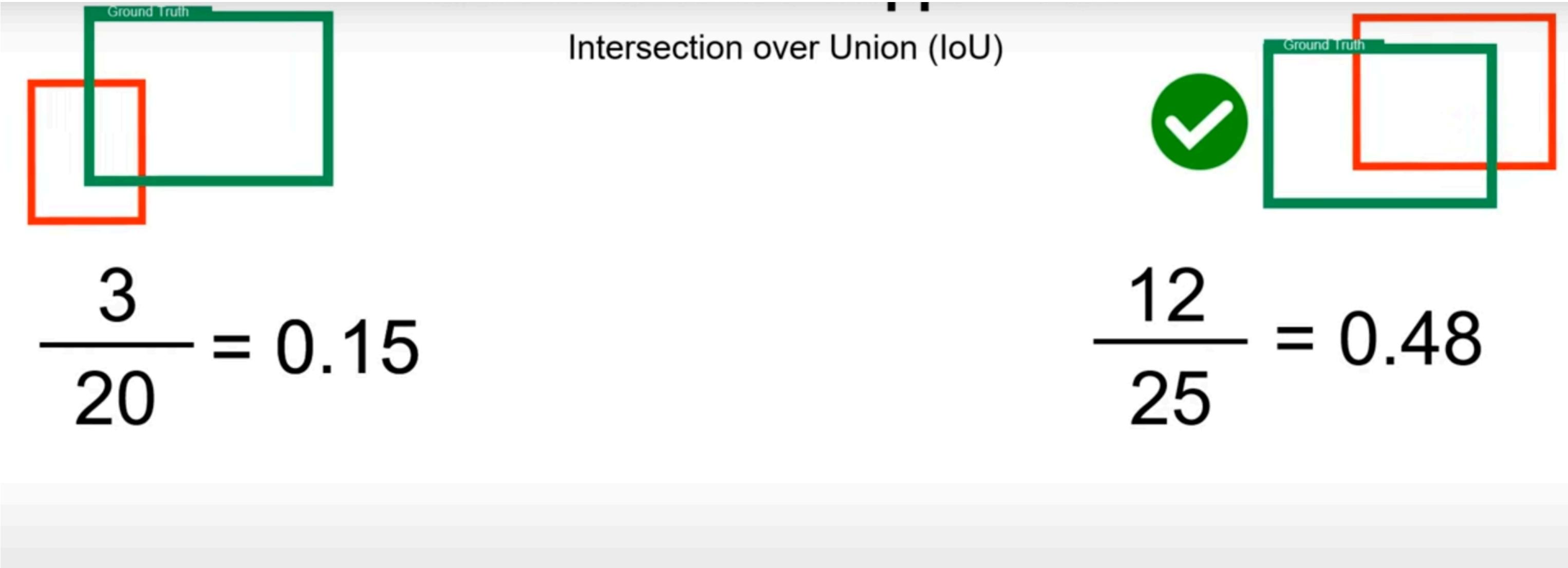


$$\frac{12}{25}$$

Non-Maximum suppression

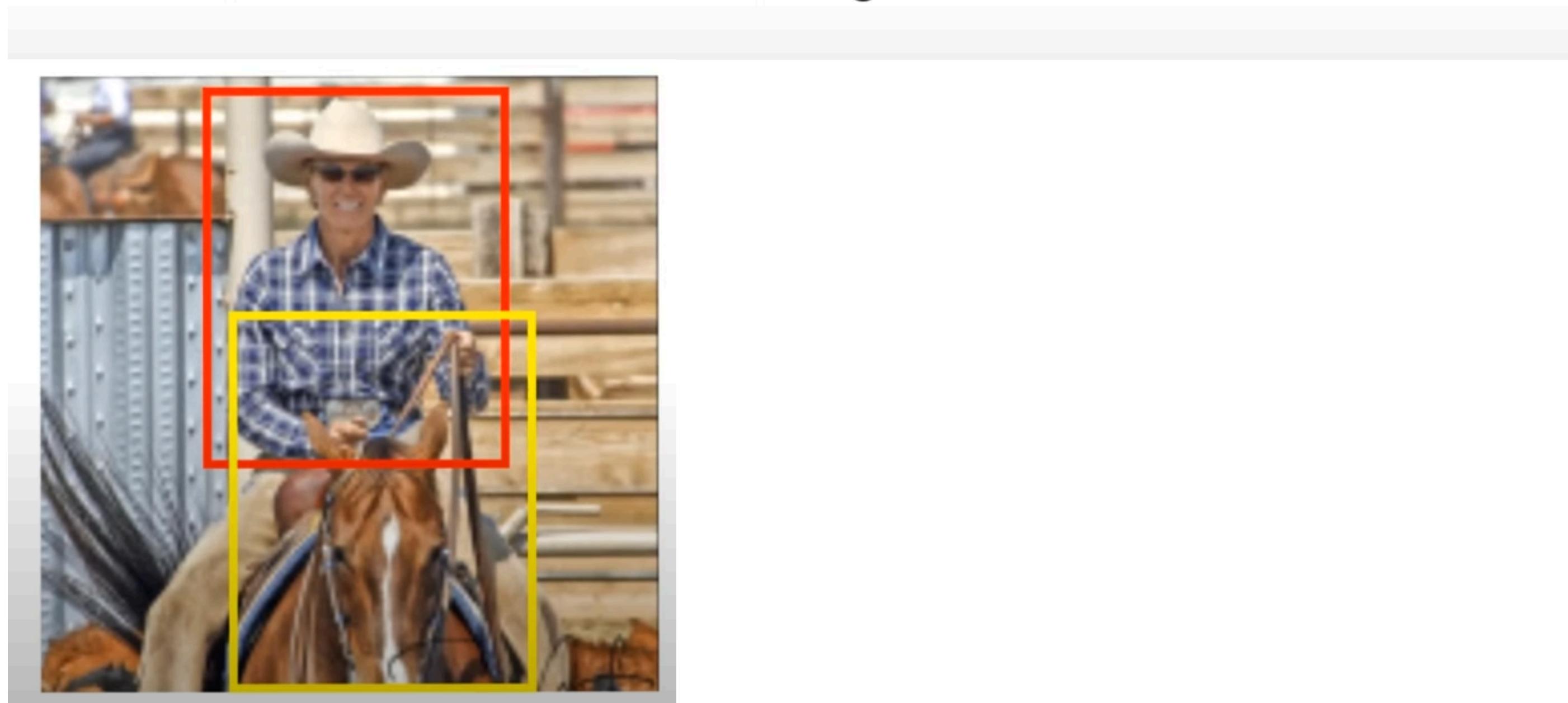


Non-Maximum suppression



Non-Maximum suppression

Note: We use non-max suppression when the object is the same for all of the bounding boxes

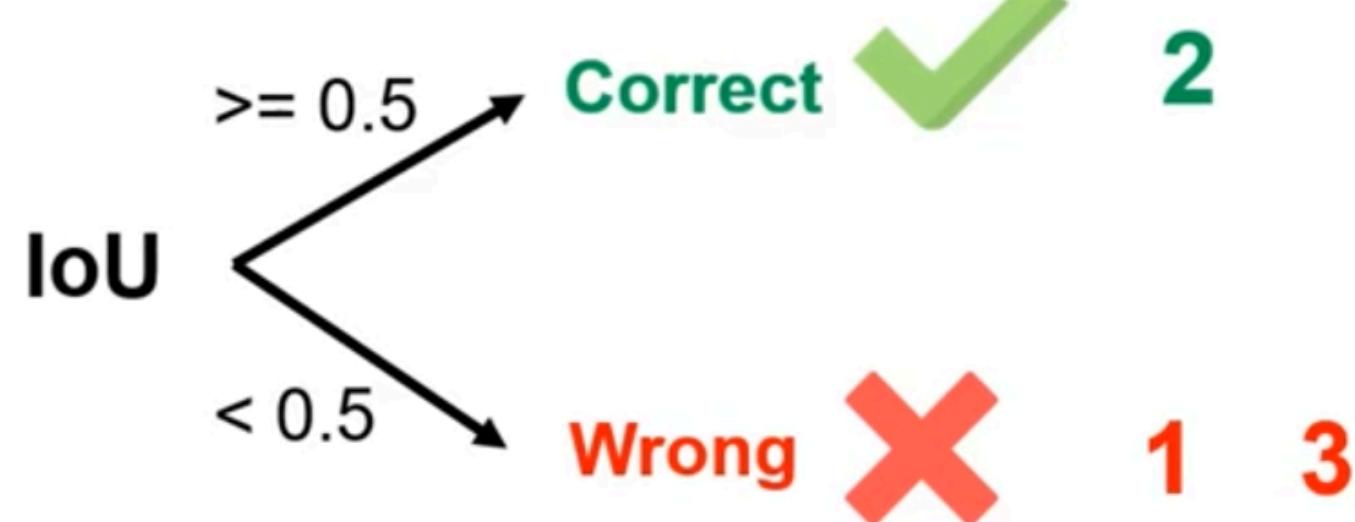


How to Evaluate our Model?

Mean Average Precision(mAP)

Mean Average Precision(mAP)

Which predicted bounding boxes are correct?

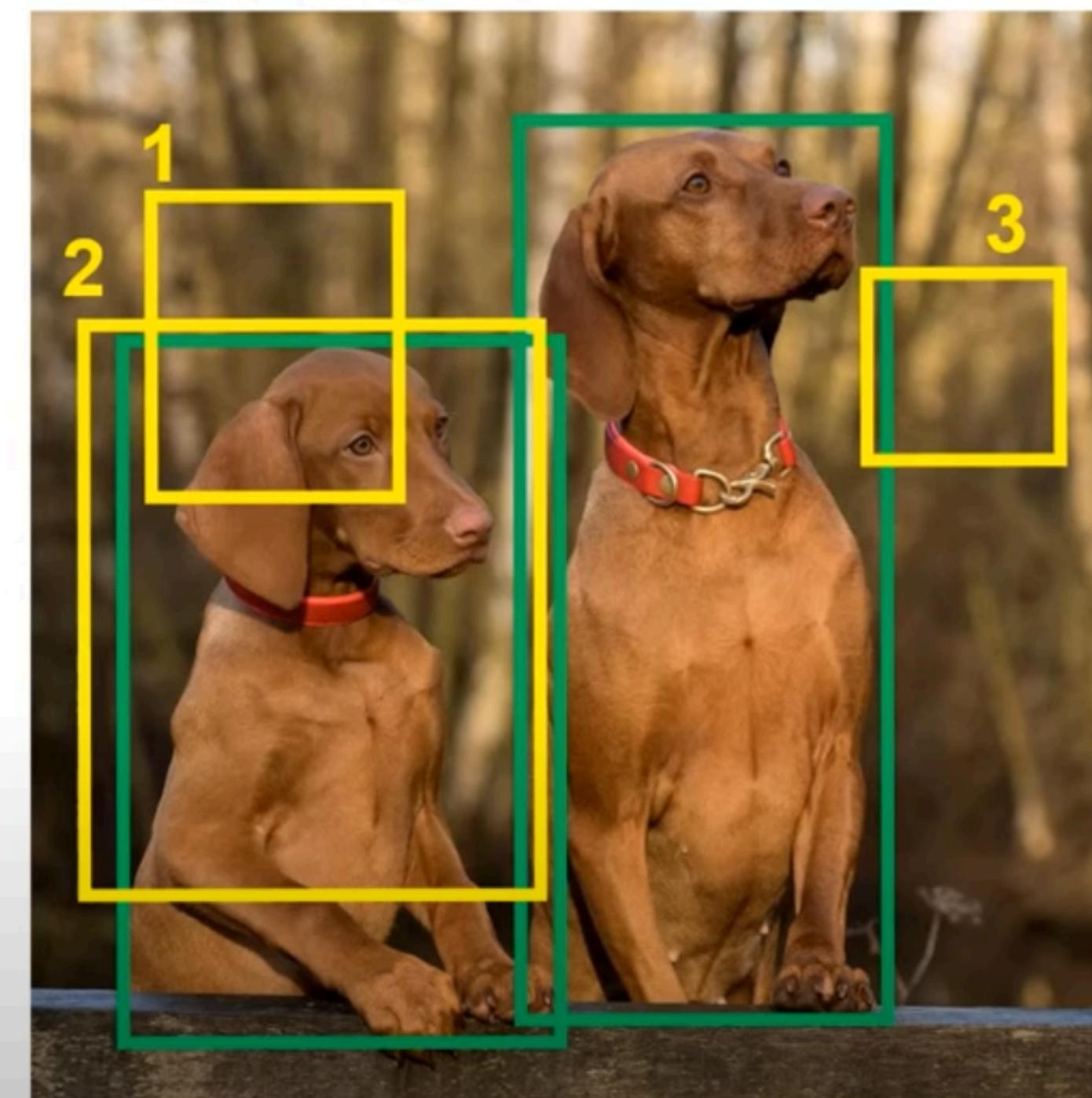


Precision

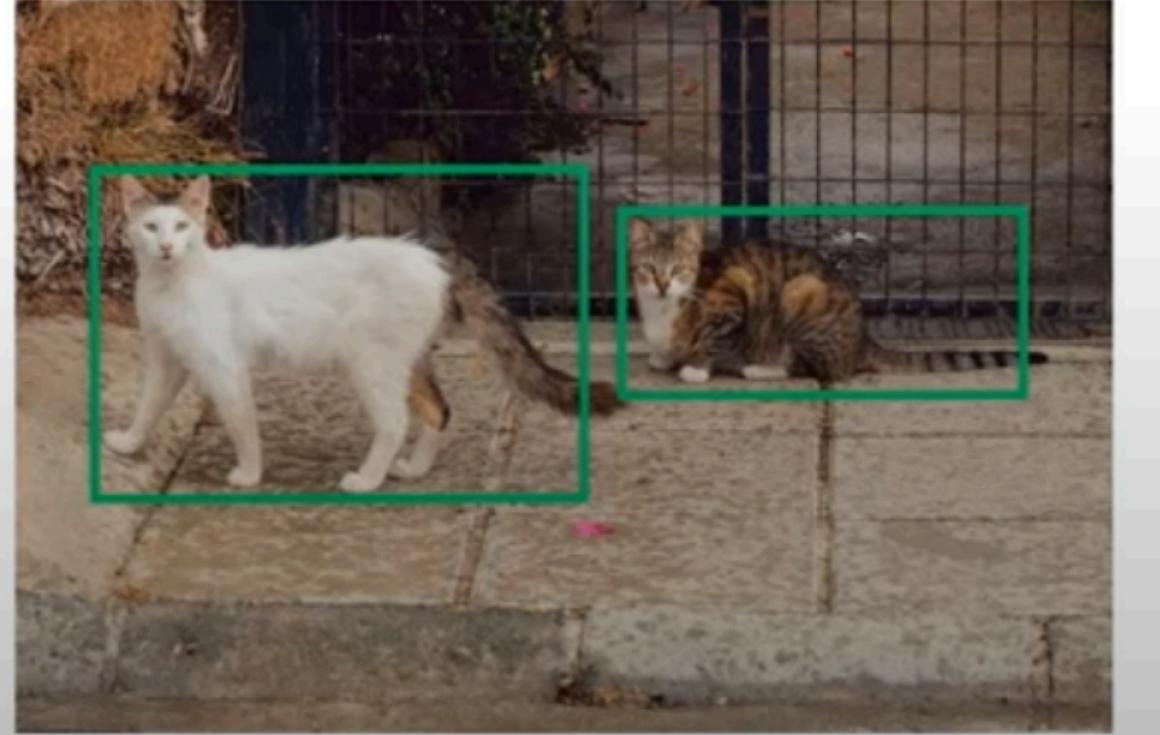
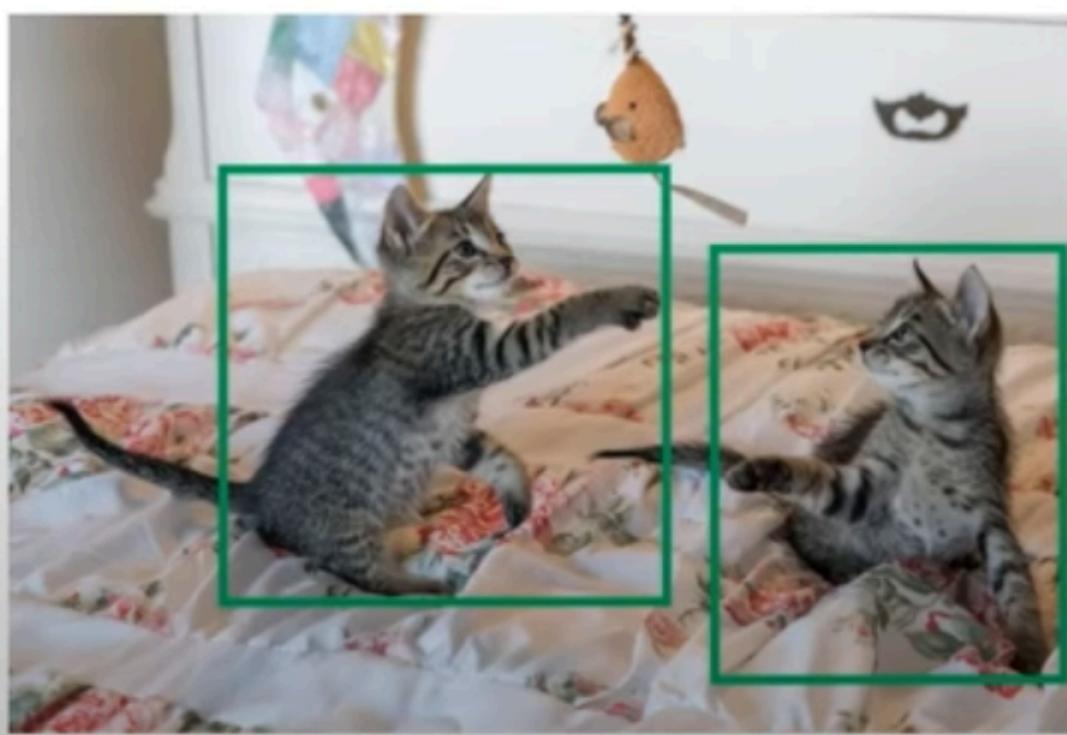
$$\frac{1}{3}$$

Recall

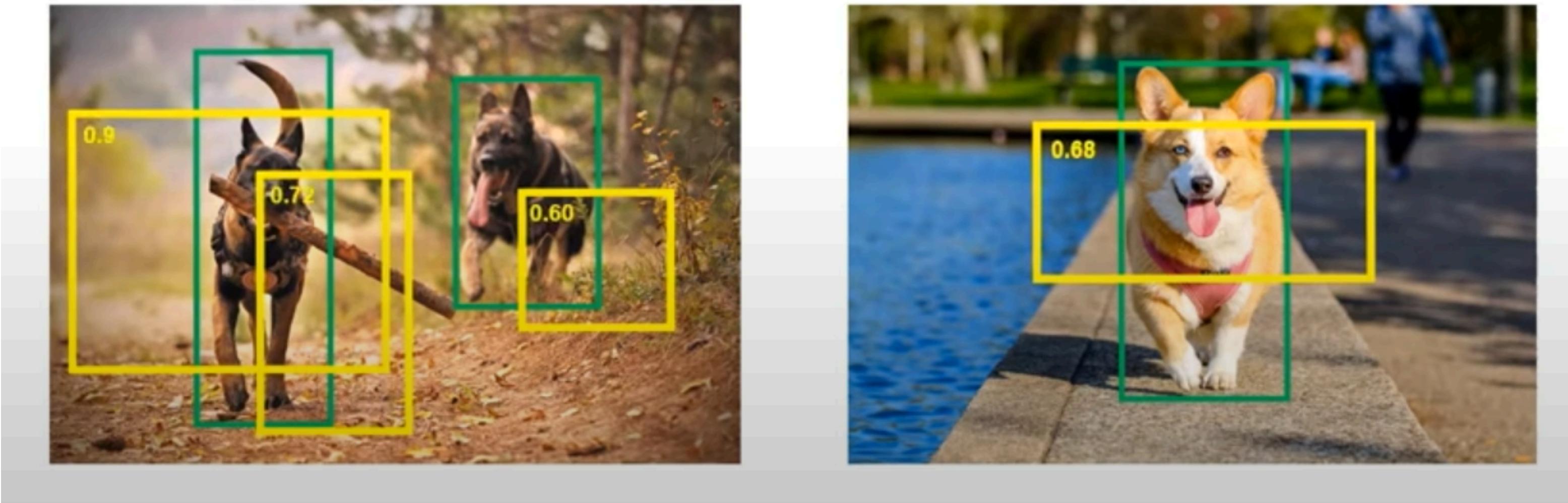
$$\frac{1}{2}$$



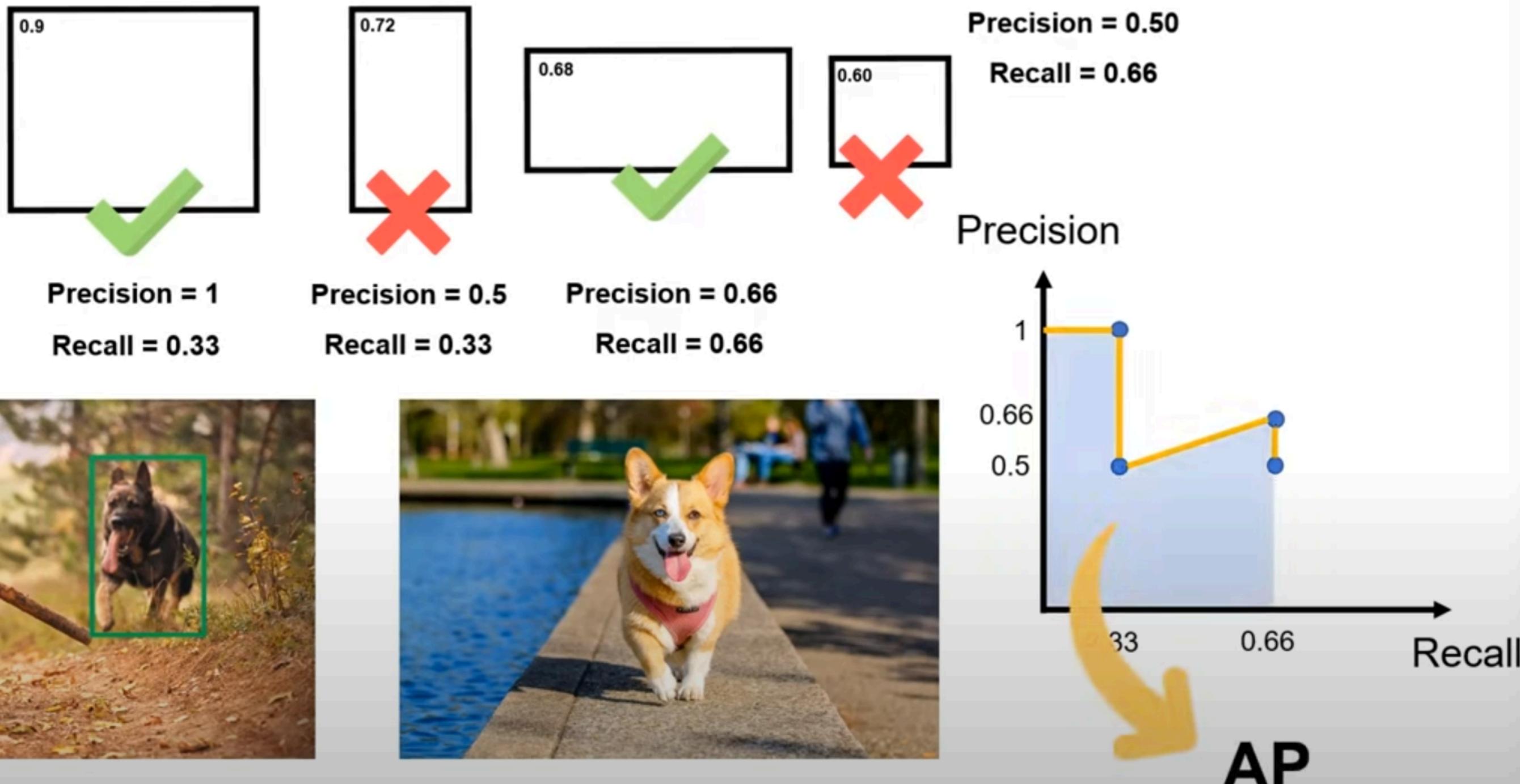
Mean Average Precision(mAP)



Mean Average Precision(mAP)



Mean Average Precision(mAP)



Mean Average Precision(mAP)



AP = 0.58

mAP = 0.60

AP = 0.62

Mean Average Precision(mAP)

mAP@0.5 = 0.60

mAP@0.55 = 0.57

mAP@0.60 = 0.53

...

mAP@0.95 = 0.23



mAP@0.5:0.95:0.05 = 0.55