

Image Classification

Process of assigning labels is called **annotation**



only Image class by CNN

Object Detection

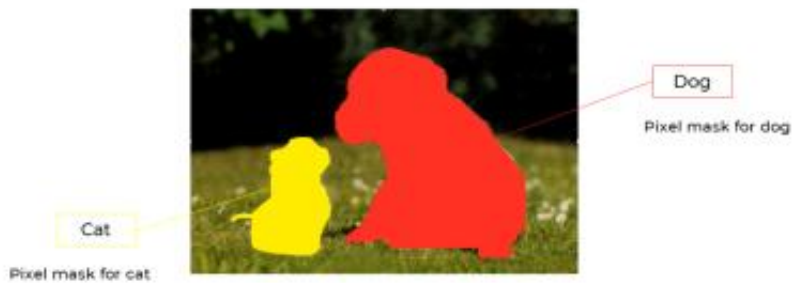


$$\begin{bmatrix} 1 \\ 100 \\ 20 \\ 170 \\ 200 \end{bmatrix} = \text{Regression}$$

Image + Bounding boxes

(x,y)

Image Segmentation



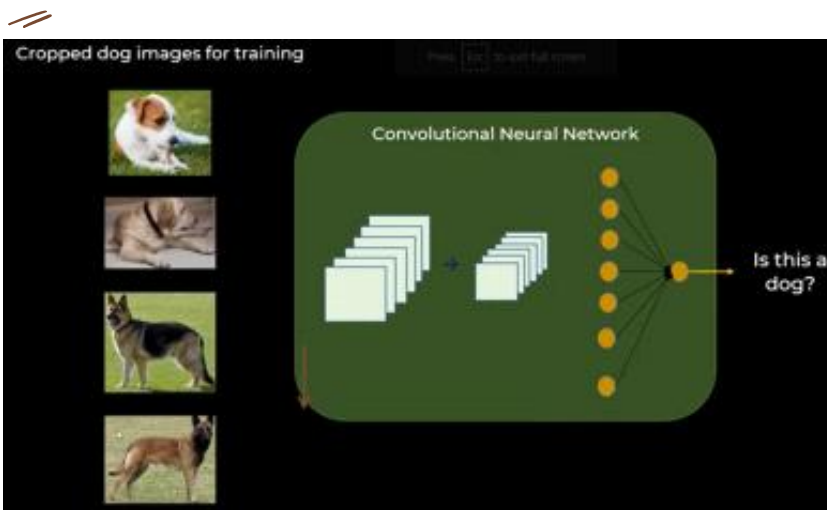
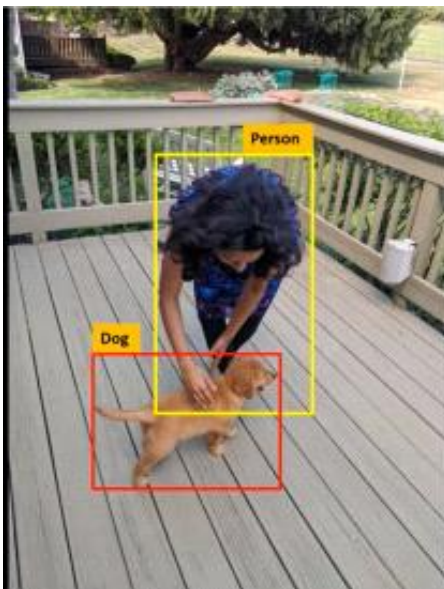
14 million hand annotated images

Annotated bounding boxes for at least 1 million images

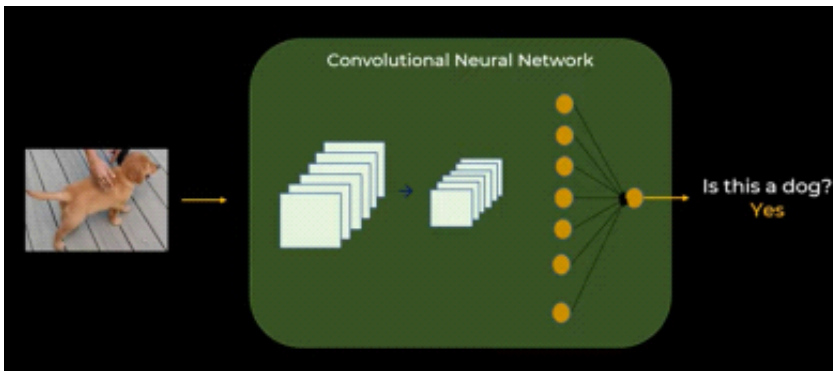
Annotated using crowdsourcing



Sliding Window Object Detection Technique -

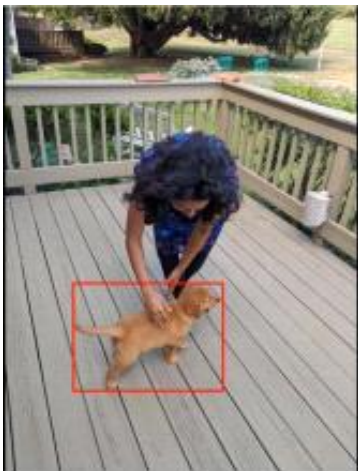


Test data below =>



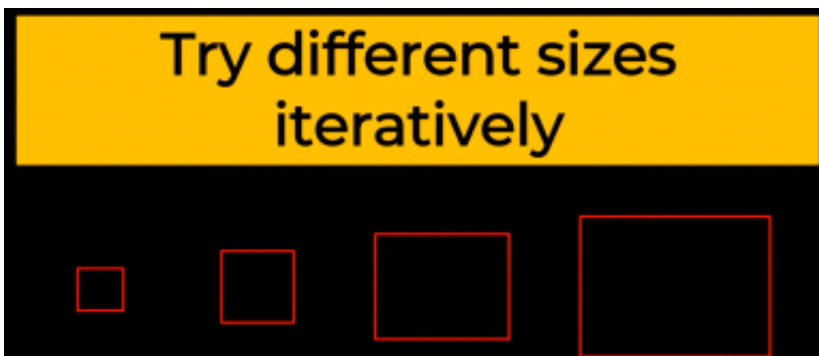
Window is passed to CNN and asked whether
it's a dog or not?? — No dog

1) Move the window to right and keep on
moving the window in the CNN



finally we are able to locate the dog

* But the question remains — what should be the
window size?? — Trial and Error




⇒ Computationally Expensive



YOLO-

Image Classification

Is this a dog or a person?




Neural Network Output

Dog = 1
Person = 0

Object Localization


Where exactly is the dog in this image?





Neural Network Output

Dog = 1
Person = 0
+
Bounding Box

Object Localization



P_c	1
B_x	50
B_y	70
B_w	60
B_h	70
C_1	1
C_2	0

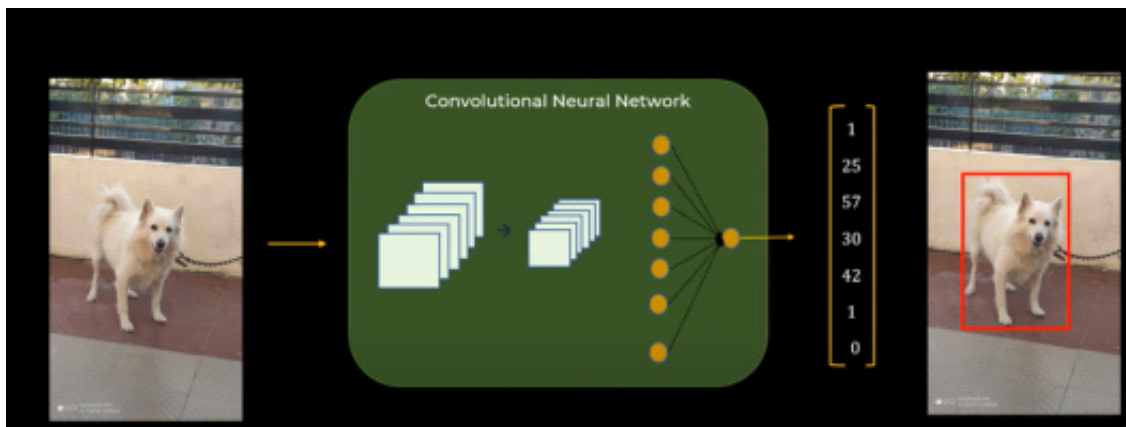
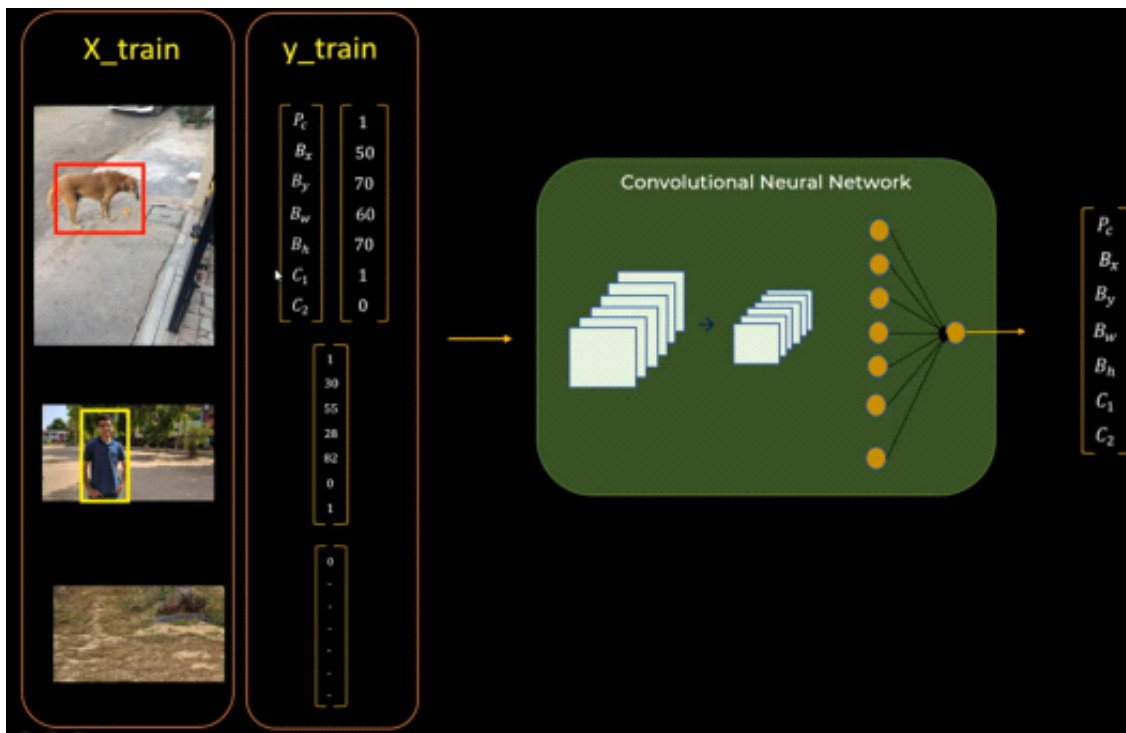
1
30
28
28
82
0
1

0
-
-
-
-
-
-

⇒ No object in the Image

$C_1 = \text{Dog class}$
 $C_2 = \text{Person Class}$

⇒ vector of size 7

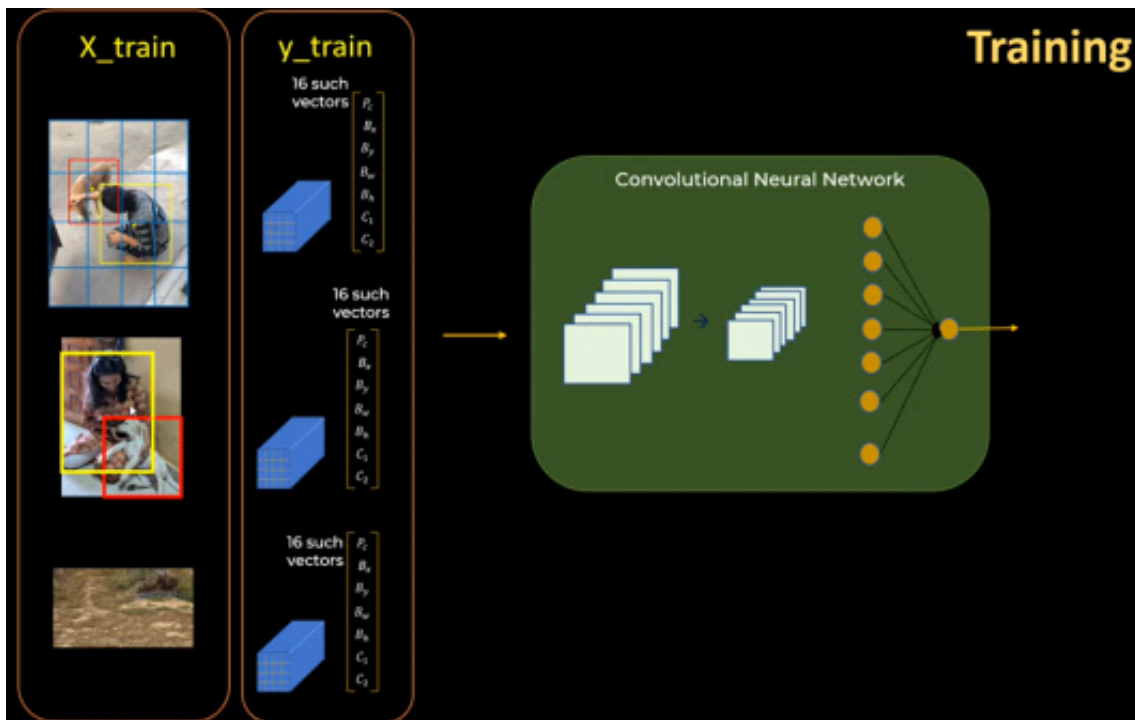
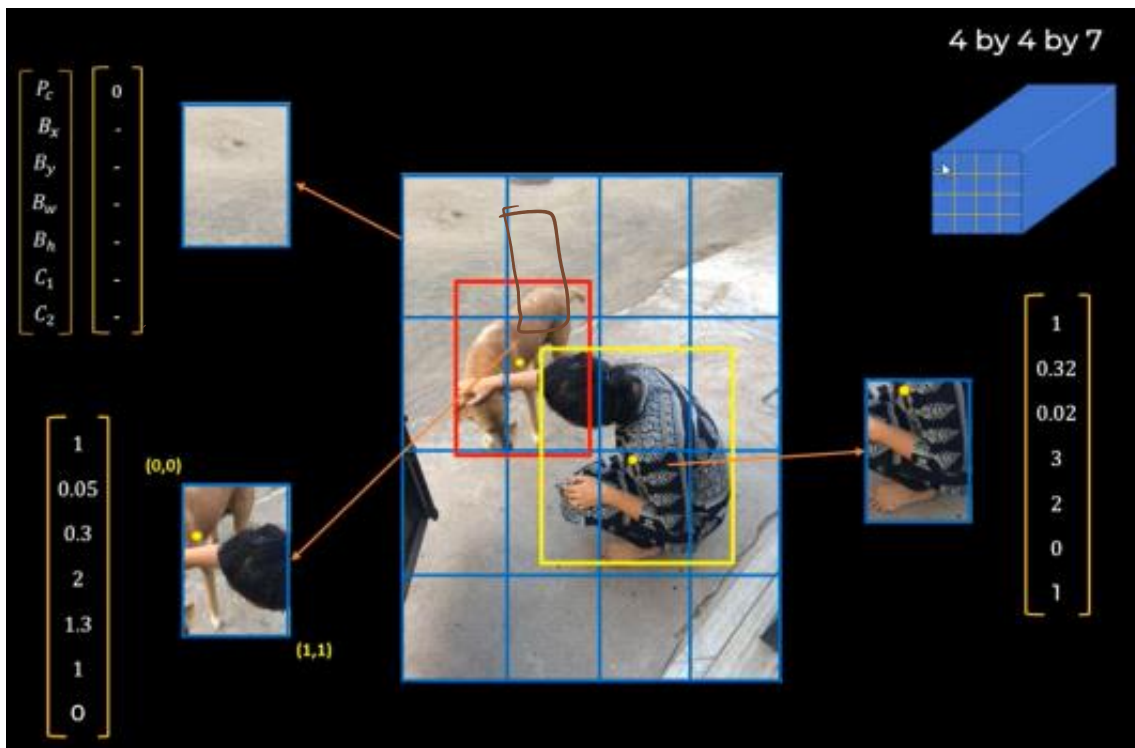


Output a new Image + Bounding boxes

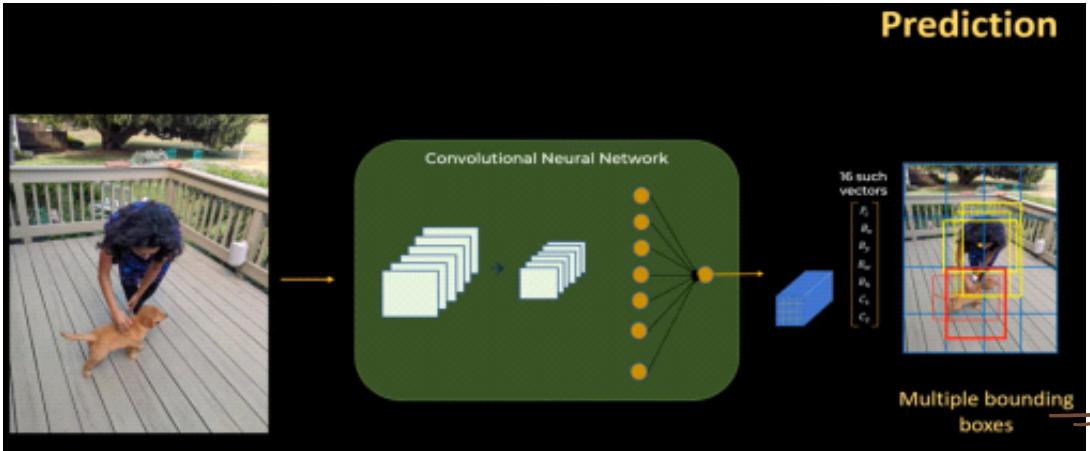
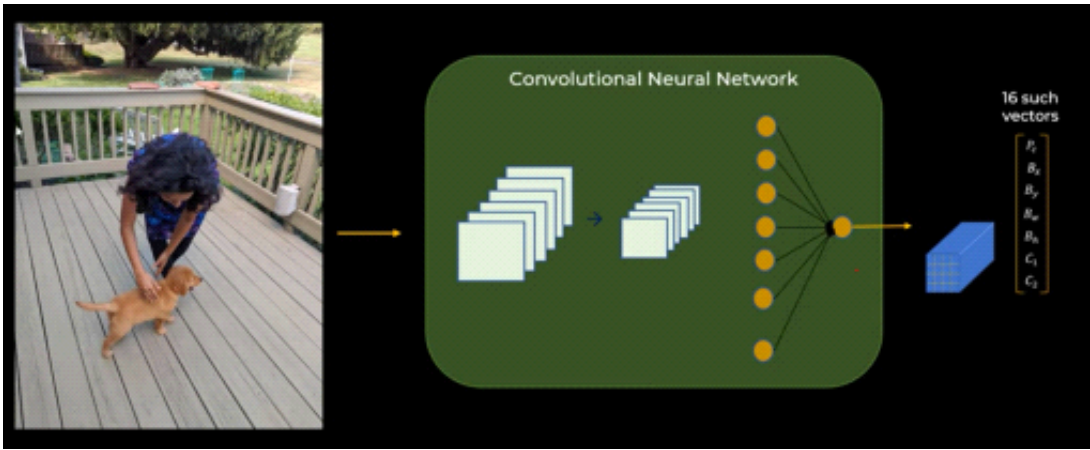


⇒ Multiple objects - What to do??
then hard to determine the size of output

Multiple Objects in one Image -

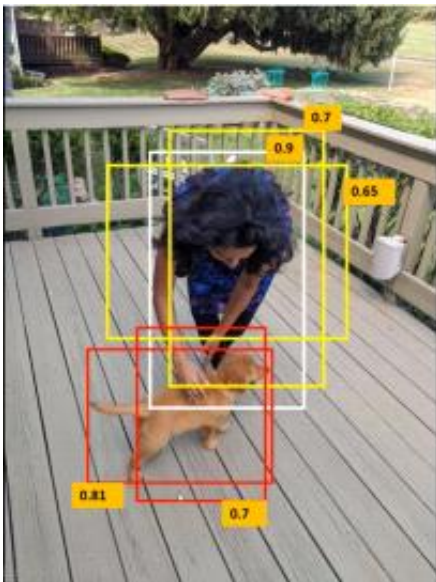


⇒ Each image will have a grid - standard for all images

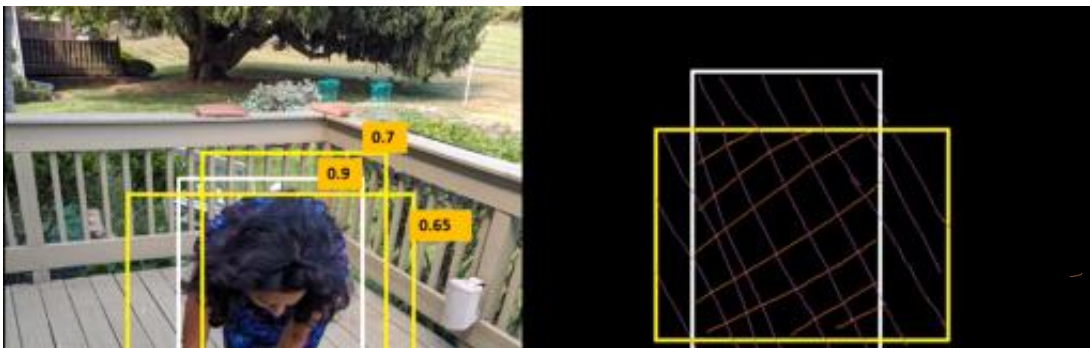


⇒ In one forward pass you can get all the outputs =

⇒ What if this happens??



⇒ Can take max of Prob. - If there is another person then how will it be next?



⇒ finding overlapping area

Intersection over union = $\text{Intersect area} / \text{union area}$

Intersection over union : IOU

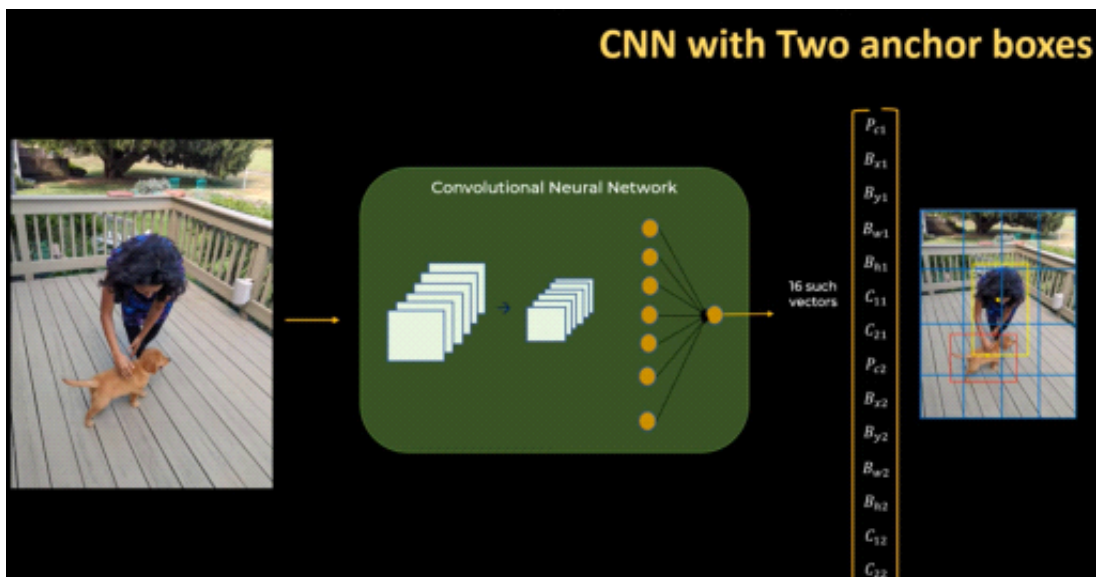
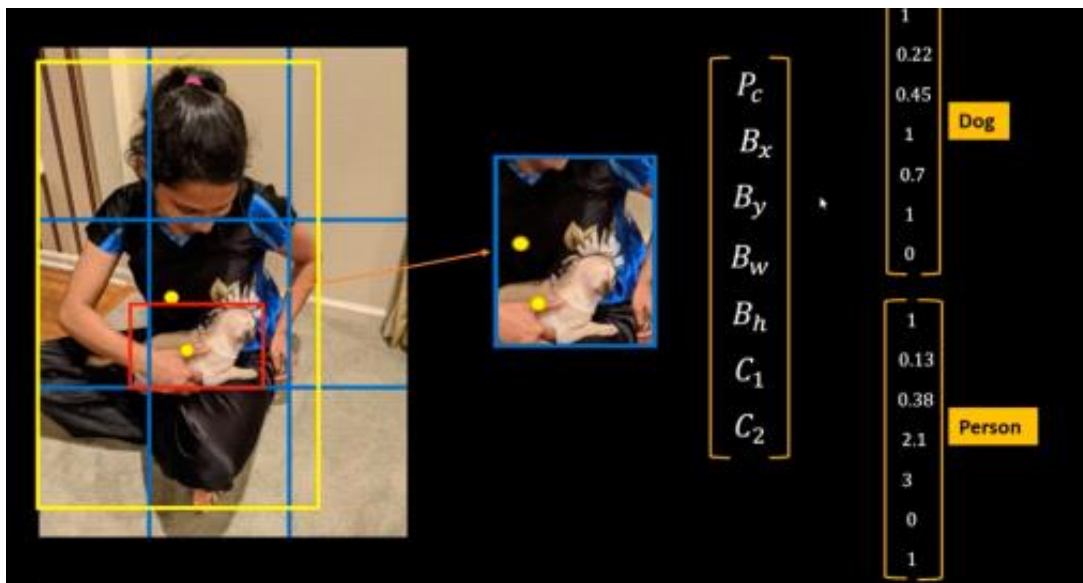
⇒ finding overlapping area

Non max suppression

What if one grid cell has center of two objects?

P_c

1
0.22
0.45



<https://jonathan-hui.medium.com/ssd-object-detection-single-shot-multibox-detector-for-real-time-processing-9bd8deac0e06>

<https://www.mihaileric.com/posts/object-detection-with-rcnn/>

https://mlwhiz.com/blog/2019/12/05/od/?utm_campaign=demystifying-object-detection-and-instance-segmentation-for-data-scientists&utm_medium=social_link&utm_source=missingletter