deeplearning.ai

Object Detection

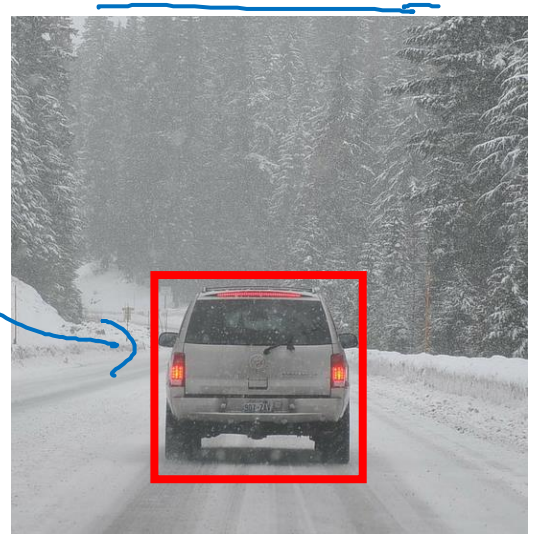
Object localization

What are localization and detection?

Image classification

"Car"

Classification with localization



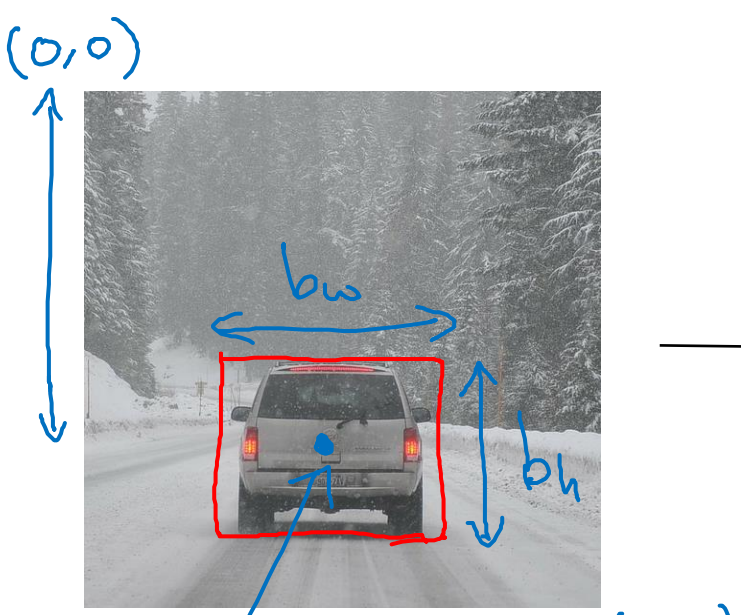
"Car"

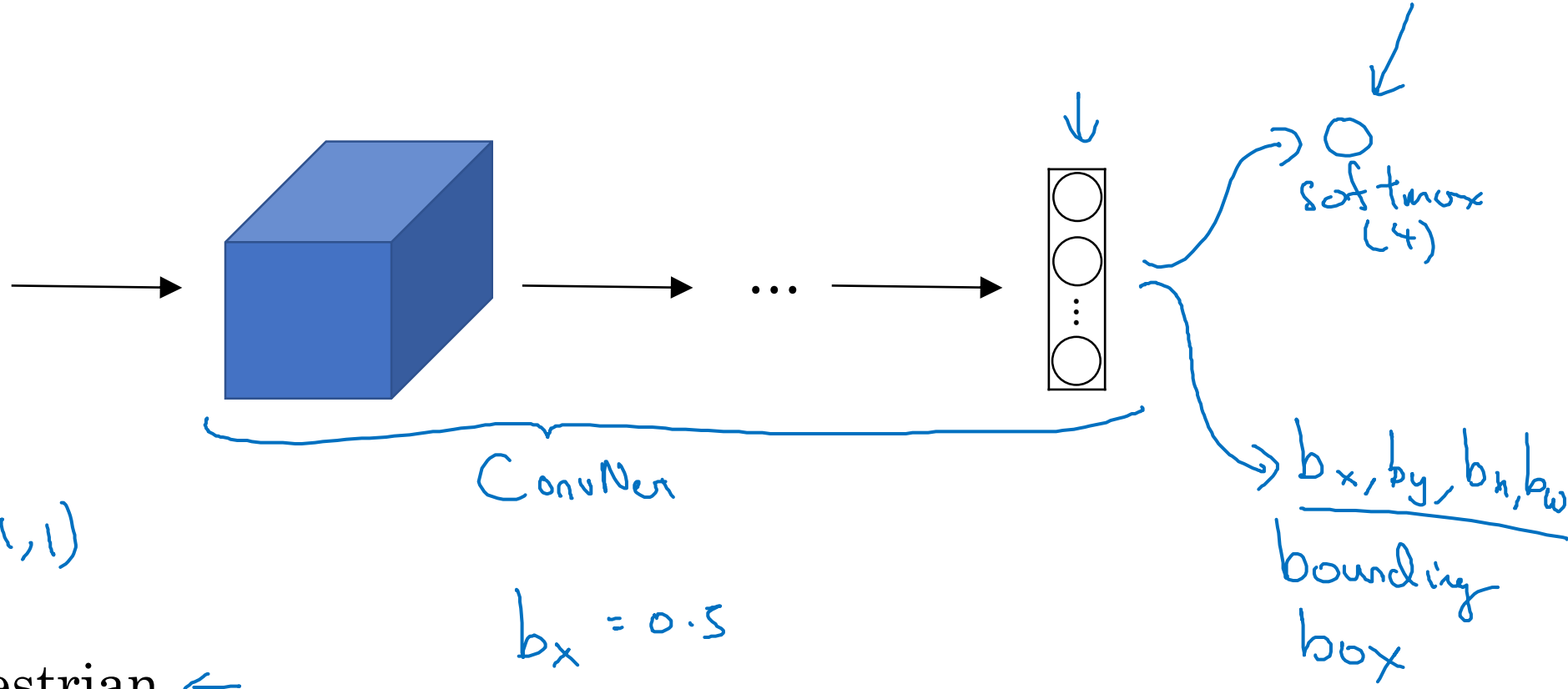
1 object

Detection

multiple objects

Classification with localization





1 - pedestrian ←
2 - car ←
3 - motorcycle ←
4 - background

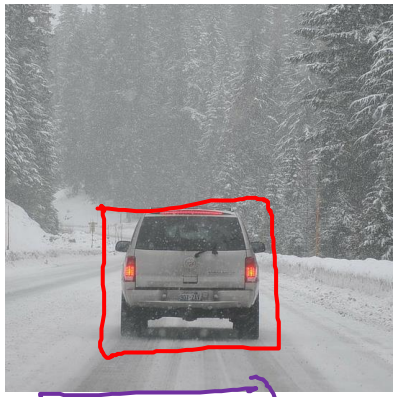
$b_x = 0.5$
$b_y = 0.7$
$b_h = 0.3$
$b_w = 0.4$

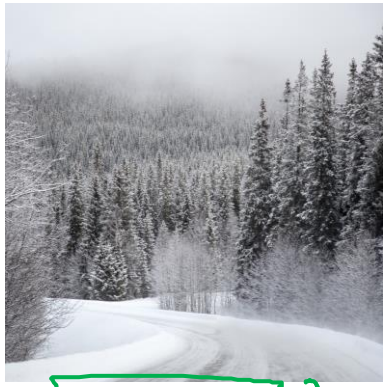
Andrew Ng

Defining the target label y

1 - pedestrian
2 - car
3 - motorcycle
4 - background

Need to output b_x, b_y, b_h, b_w, class label (1-4)

$x =$ 



$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

p_c: is there an object?

In practice, you could use improbably use a log likelihood loss for the c1, c2, c3 to the softmax, output one of those elements, usually you can use squared error or something like squared error for the bounding box coordinates. and then for Pc, you could use something like the logistic regression loss although even if you use squared error will probably work ok.

(x, y)

$$\mathcal{L}(\hat{y}, y) = \begin{cases} (\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2 + \cdots + (\hat{y}_8 - y_8)^2 & \text{if } y_1 = 1 \\ (\hat{y}_1 - y_1)^2 & \text{if } y_1 = 0 \end{cases}$$

p_c

"don't care"

Andrew Ng