Object Detection

Object localization

deeplearning.ai
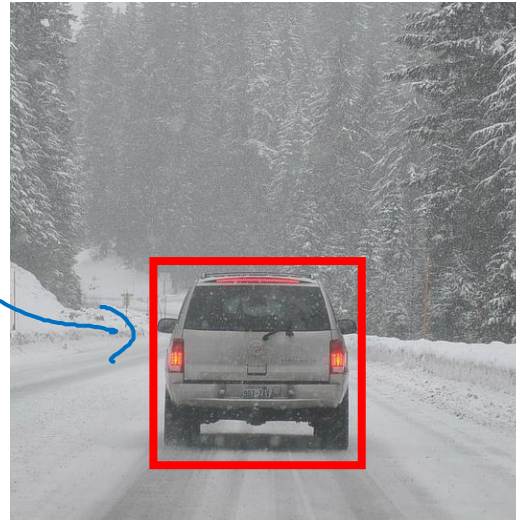
What are localization and detection?

Image classification

Classification with localization
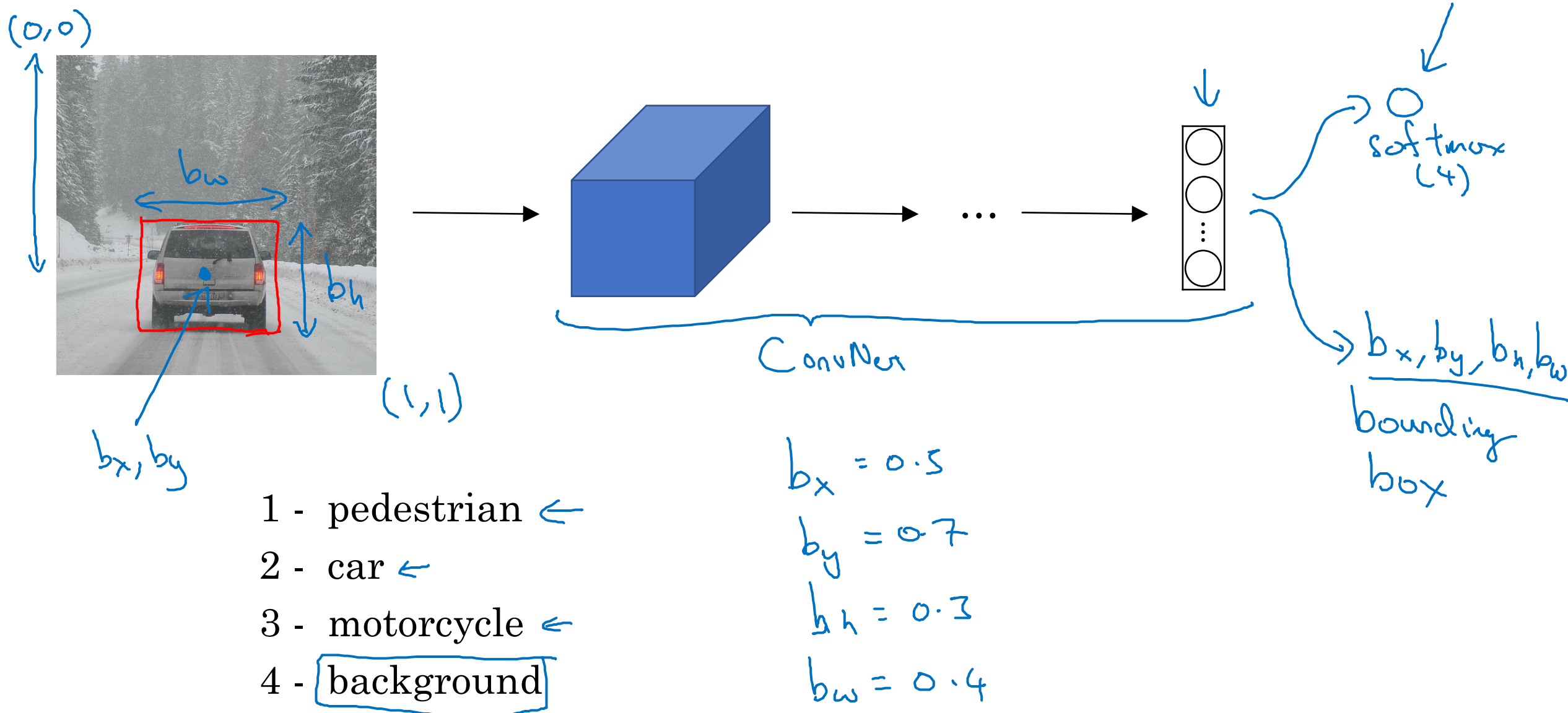
Detection

"Car"

"Car"

1 object

multiple objects
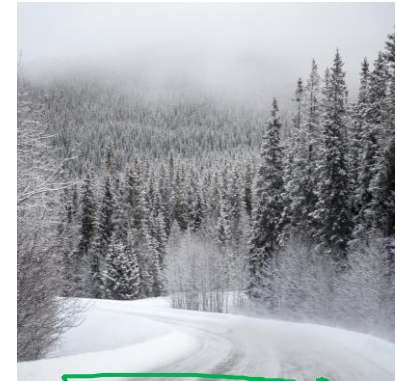
Andrew Ng

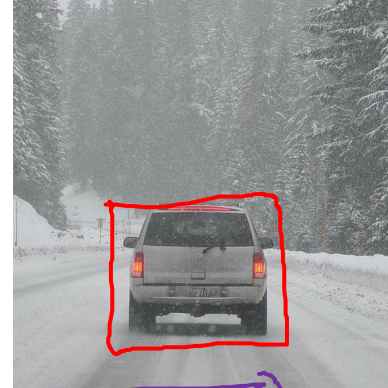# Classification with localization

$(0,0)$

b_w

b_h

$(1,1)$

b_x, b_y

1 - pedestrian ←

2 - car ←

3 - motorcycle ←

4 - background

ConvNet

softmax (4)

b_x, b_y, b_h, b_w

bounding box

$b_x = 0.5$

$b_y = 0.7$

$b_h = 0.3$

$b_w = 0.4$

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)

$x =$

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

$$y = \begin{bmatrix} P_c \\ \hline b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$ is there any object?

(x, y)

$$\begin{bmatrix} 1 \\ b_x \\ b_y \\ b_h \\ b_w \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix}$$ ← "don't care" P_c

Andrew Ng