deeplearning.ai

# Object Detection

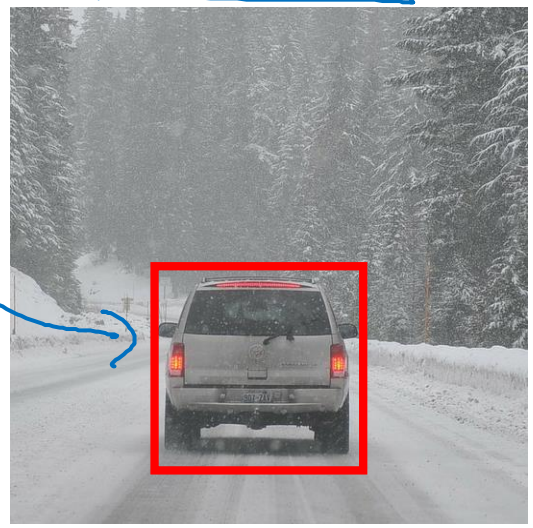
## Object localization

# What are localization and detection?

## Image classification

"Car"

## Classification with localization



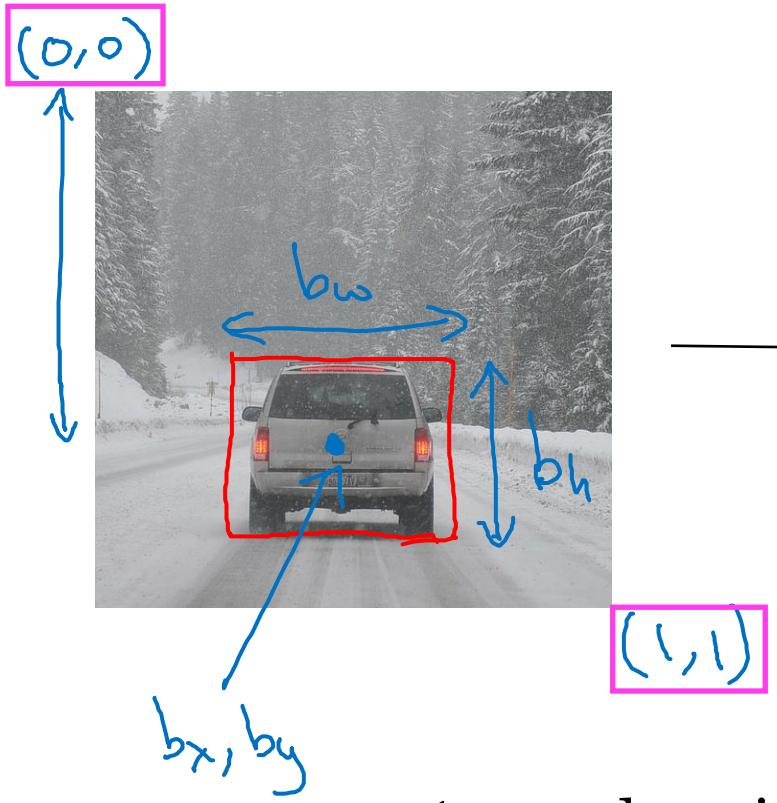
"Car"

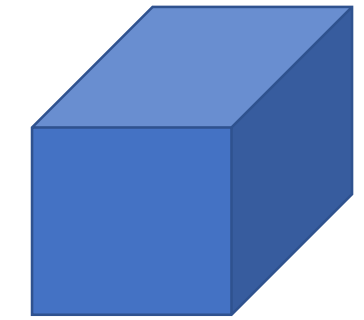
1 object

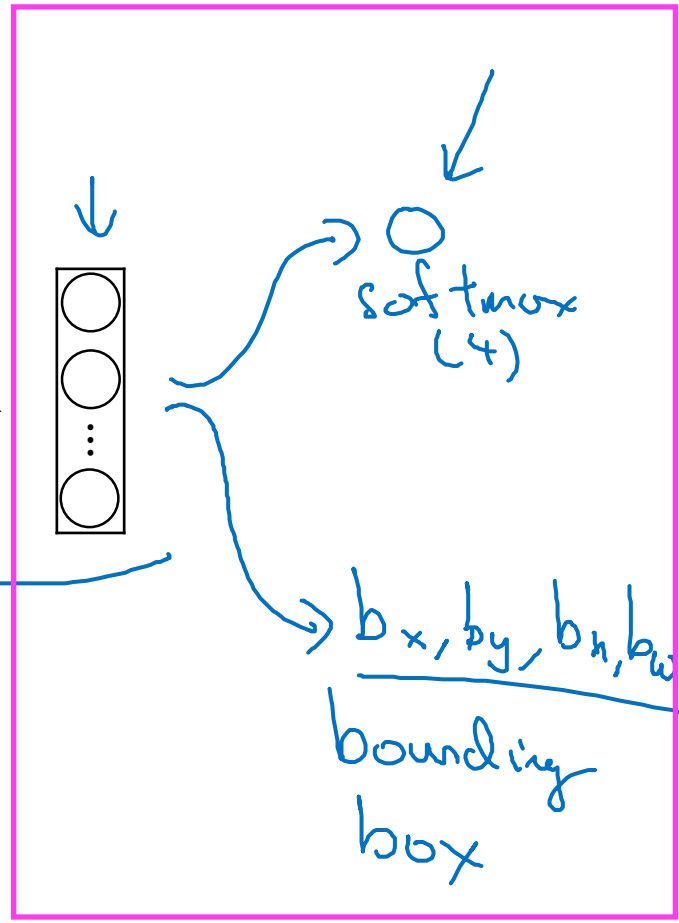
## Detection

multiple objects

# Classification with localization






ConvNet

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

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

# Defining the target label y

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

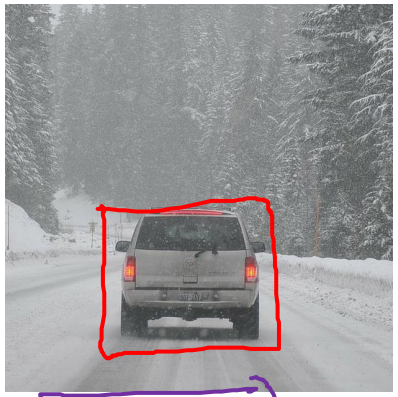
Need to output $b_x, b_y, b_h, b_w$, class label (1-4)

$\mathcal{L}(\hat{y}, y) =$

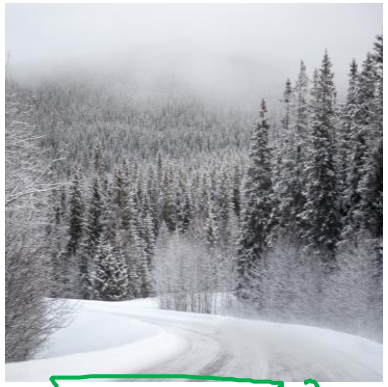
$(\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2 + \cdots + (\hat{y}_8 - y_8)^2$ if $y_1 = 1$

$(\hat{y}_1 - y_1)^2$ if $y_1 = 0$

$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$ $p_c$: is there an object?

$x =$ 

$(x, y)$ $\begin{bmatrix} 1 \\ b_x \\ b_y \\ b_h \\ b_w \\ 0 \\ 1 \\ 0 \end{bmatrix}$



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

Andrew Ng