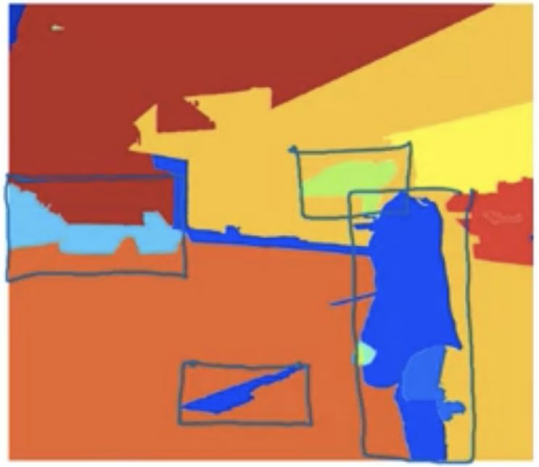
Tema 17 : Localización de Objetos y Landmarks

Region proposal: R-CNN



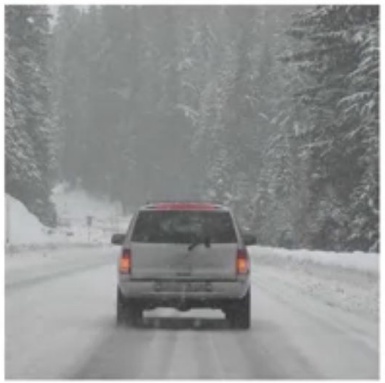
Segmentation algorithm

~2,000

[Girshik et. al, 2013, Rich feature hierarchies for accurate object detection and semantic segmentation]

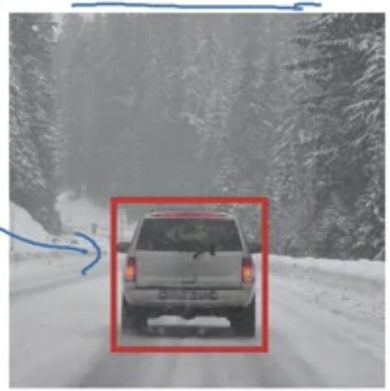
What are localization and detection?

Image classification



"Car"

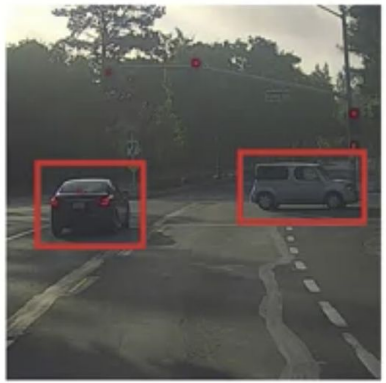
Classification with localization



"Car"

1 object

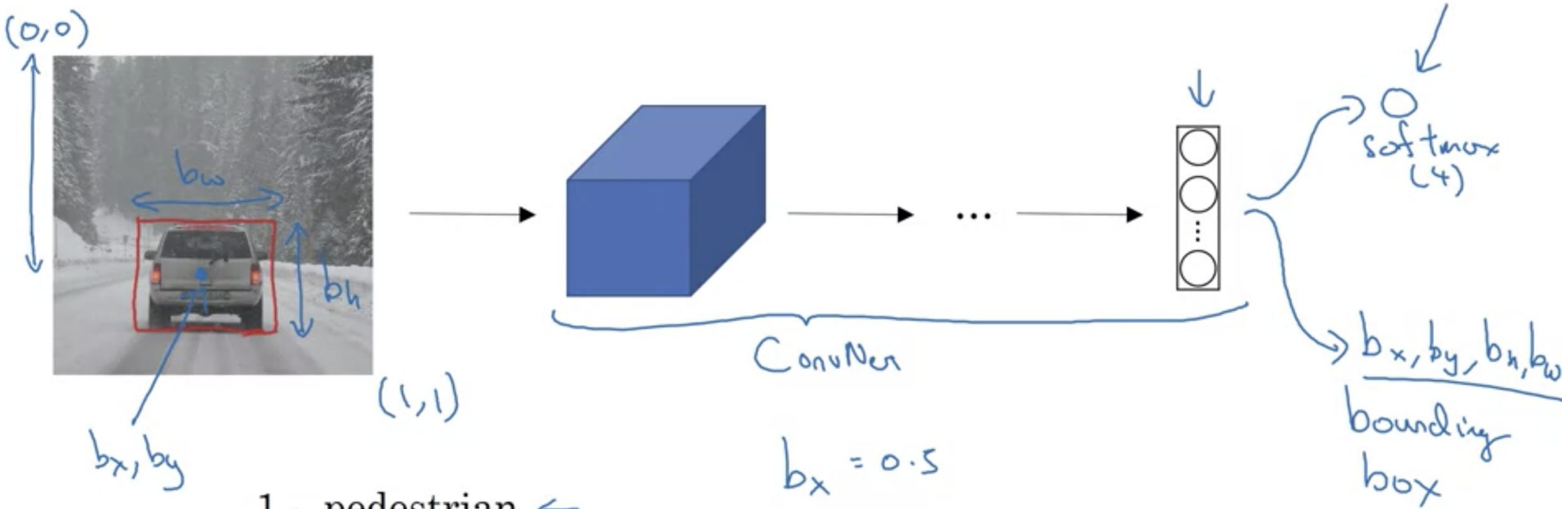
Detection



multiple objects

Andrew Ng

Classification with localization



Andrew Ng

Defining the target label y

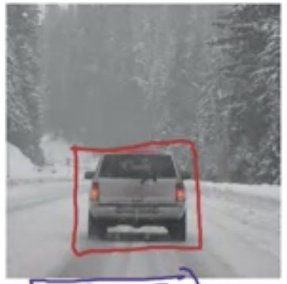
1 - pedestrian
2 - car
3 - motorcycle
4 - background

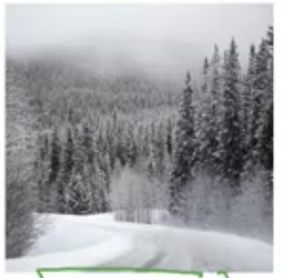
Need to output b_x, b_y, b_h, b_w, class label (1-4)

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

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

p_c: is there an object?

$x =$ 



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

(x, y)

$$\begin{bmatrix} 0 \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix}$$

p_c

← "don't care"

Andrew Ng

Landmark detection

b_x, b_y, b_h, b_w

$l_{1x}, l_{1y}, l_{2x}, l_{2y}, l_{3x}, l_{3y}, l_{4x}, l_{4y}, \ldots, l_{64}, l_{64y}$

X, y

$l_{1x}, l_{1y}, \ldots, l_{32x}, l_{32y}$

Andrew Ng