deeplearning.ai

Object Detection

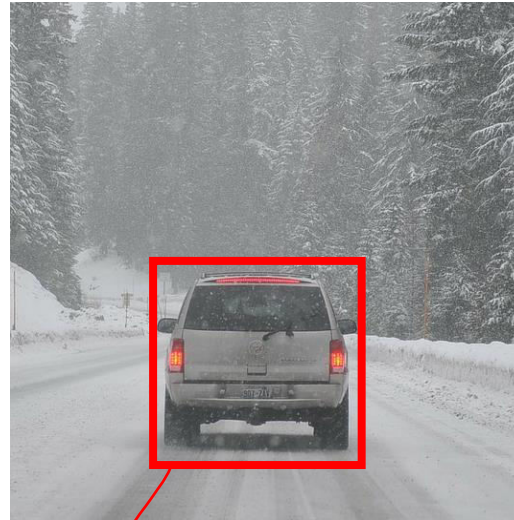
Object localization

What are localization and detection?

Image classification

Classification with localization

Detection

Localization = Draw a Bounding box around the car

- Usually 1 object in classification & localization
- In Detection, there are often more than 1 objects that may belong to more than 1 class
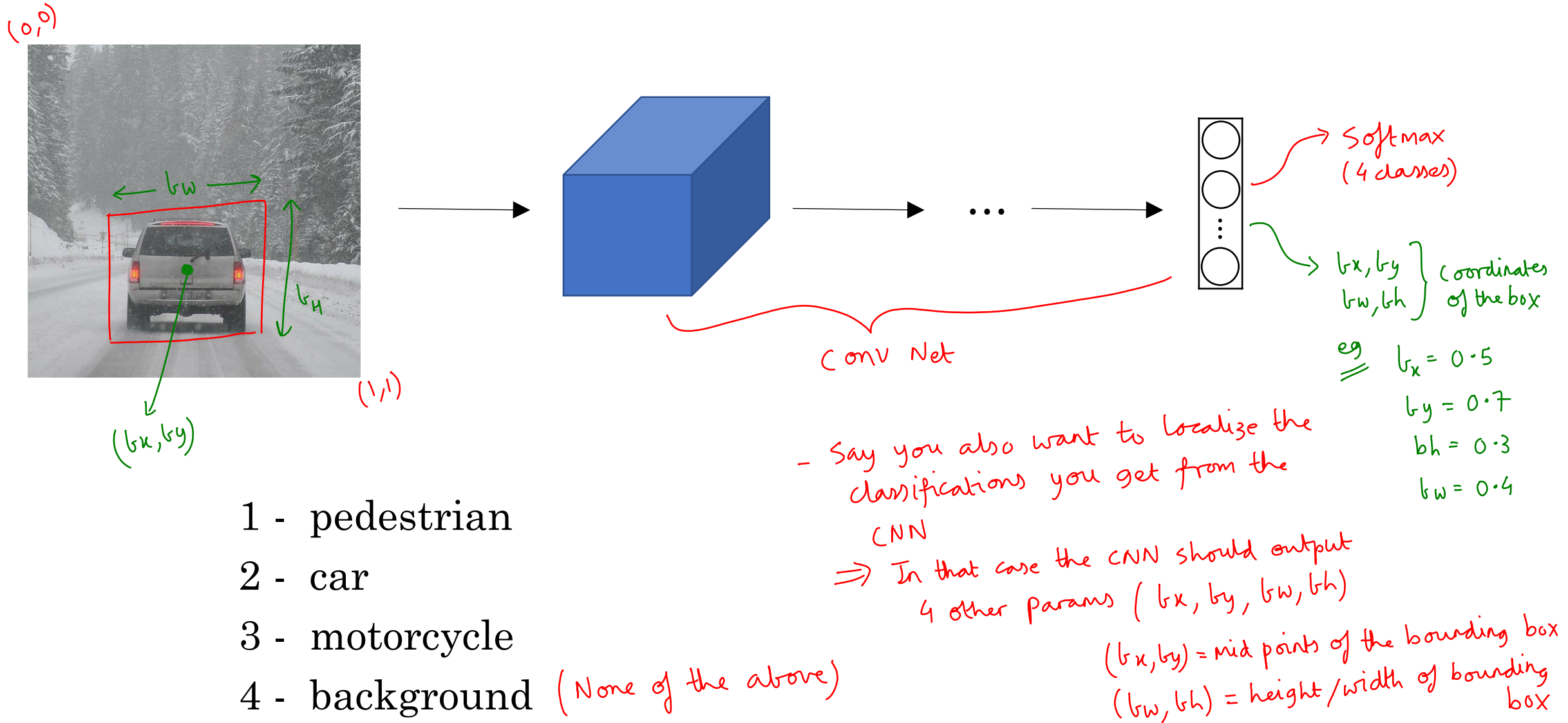
Detect & localize

- Multiple objects
- May have multiple classes

Classification with localization

(0,0)

b_w

b_h

(b_x, b_y)

(1,1)

1 - pedestrian
2 - car
3 - motorcycle
4 - background (None of the above)

Conv Net

Softmax
(4 classes)

b_x, b_y
b_w, b_h } coordinates of the box

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

- Say you also want to localize the classifications you get from the CNN

⇒ In that case the CNN should output 4 other params (b_x, b_y, b_w, b_h)

(b_x, b_y) = mid points of the bounding box
(b_w, b_h) = height / width of bounding box
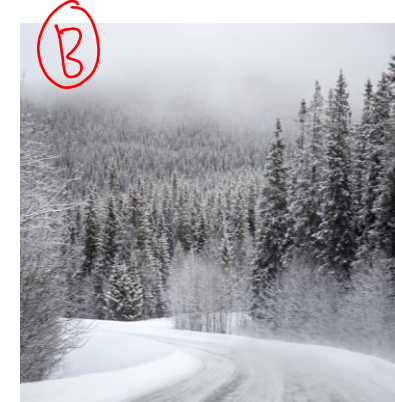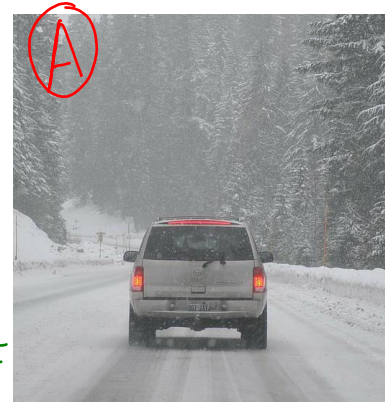
Andrew Ng

Defining the target label y

Classes

1 - pedestrian

2 - car

3 - motorcycle

4 - background

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

Probability of

$$y = \begin{bmatrix} P_c \\ b_x \\ b_y \\ b_h \\ b_w \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}$$

P_c — is there an object

b_x, b_y — mid point

b_h, b_w — height/width

C_1, C_2, C_3 — classes

P_c = is there an object

i.e, it is 1 if class = 1, 2, 3

$P_c = 0$ if class = 4

A   B

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

→ object does exist

Remember

$y_1 \rightarrow P_c$

$y_2, y_3, y_4, y_5 \rightarrow b$

$y_6, y_7, y_8 \rightarrow C$

- Remember, we assume only 1 object per Image

example

$X = \text{Ⓐ}$

then $y = \begin{bmatrix} 1 \\ b_x \\ b_y \\ b_w \\ b_H \\ 0 \\ 1 \\ 0 \end{bmatrix}$

$X = \text{Ⓑ}$

then $y = \begin{bmatrix} 0 \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix}$? = Dont care value

b

C

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

→ object doesnt exist

Generally → least squares loss is NOT used for classification
for P_c, we use logistic loss, for b_x, b_y, b_w, b_h (Squared error) & for C_1, C_2, C_3 (Softmax/log likelihood)