



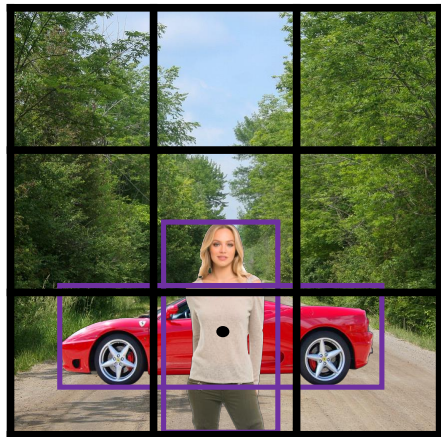
deeplearning.ai

Object Detection

Anchor boxes

detect multiple objects in a grid cells.

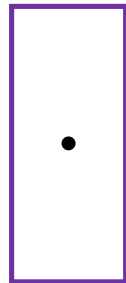
Overlapping objects:



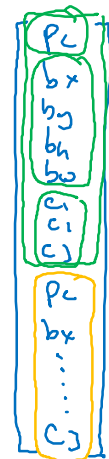
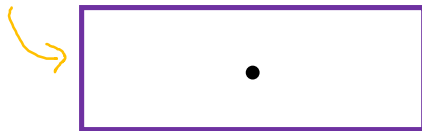
$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Annotations: A green arrow points to p_c , a blue arrow points to b_y , and a blue bracket groups c_1, c_2, c_3 .

Anchor box 1:



Anchor box 2:



Anchor box 1

Anchor box 2

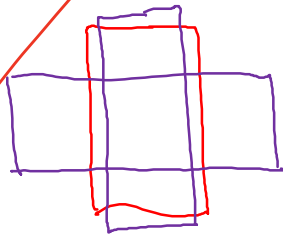
$y =$

Anchor box algorithm

Previously:

Each object in training image is assigned to grid cell that contains that object's midpoint.

Output y :
 $3 \times 3 \times 8$



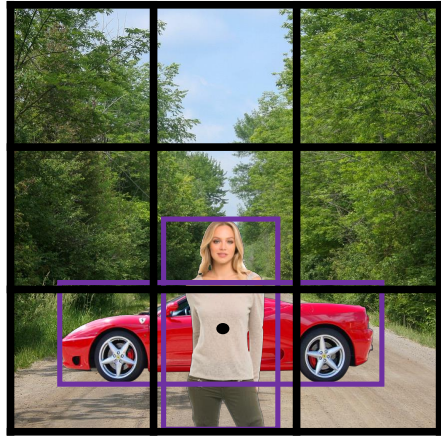
With two anchor boxes:

Each object in training image is assigned to grid cell that contains object's midpoint and anchor box for the grid cell with highest IoU.

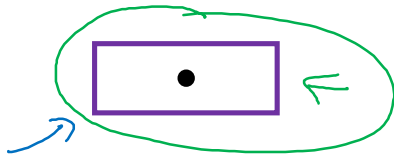
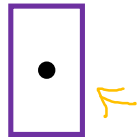
(grid cell, anchor box)

Output y :
 $3 \times 3 \times 16$
 $3 \times 3 \times 2 \times 8$

Anchor box example



Anchor box 1: Anchor box 2:



$y =$

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

Handwritten annotations for the first vector y (orange text):

- b_x (next to b_x)
- b_y (next to b_y)
- b_h (next to b_h)
- b_w (next to b_w)
- 1 (next to c_1)
- 0 (next to c_2)
- 0 (next to c_3)
- 1 (next to p_c)
- b_x (next to b_x)
- b_y (next to b_y)
- b_h (next to b_h)
- b_w (next to b_w)
- 0 (next to c_1)
- 0 (next to c_2)
- 0 (next to c_3)

Handwritten annotations for the second vector y (green text):

- b_x (next to b_x)
- b_y (next to b_y)
- b_h (next to b_h)
- b_w (next to b_w)
- 0 (next to c_1)
- 0 (next to c_2)
- 0 (next to c_3)

Handwritten notes:

- car only? (at the top left)
- anchor box 1 (bracketed next to the first vector)
- anchor box 2 (bracketed next to the second vector)