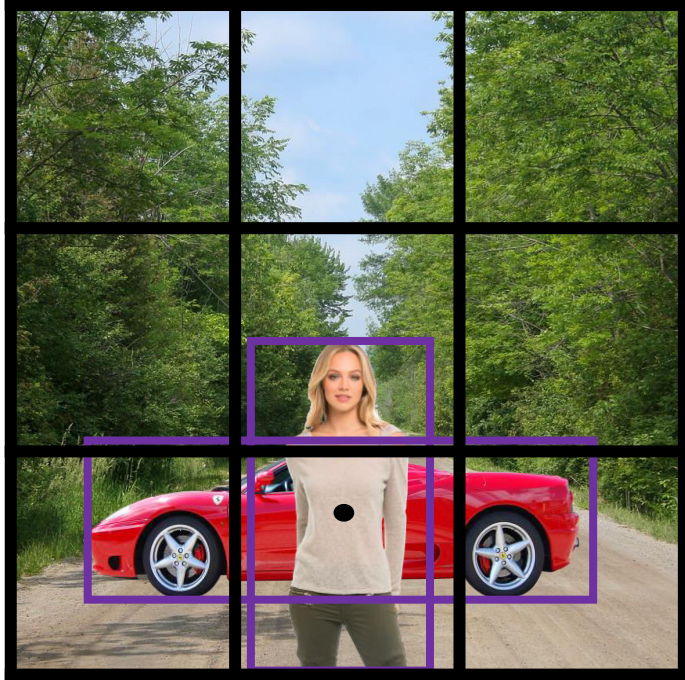Object Detection
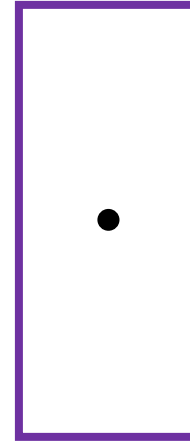
Anchor boxes

deeplearning.ai

Overlapping objects:

$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} \Bigg\} \text{pedestrian, car, motorcycle}$$

Anchor box 1:

Anchor box 2:
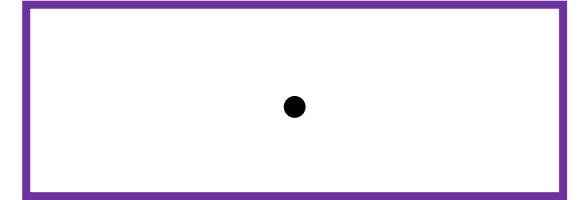
$$Y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \\ p_c \\ b_x \\ b_y \\ b_H \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} \begin{array}{l} \Big\} \text{Anchor box 1} \\ \\ \Big\} \text{Anchor box 2} \end{array}$$

[Redmon et al., 2015, You Only Look Once: Unified real-time object detection]

Andrew Ng

Anchor box algorithm

Previously:

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

∴ y dim is $3 \times 3 \times 8$
grid   8 elements in y

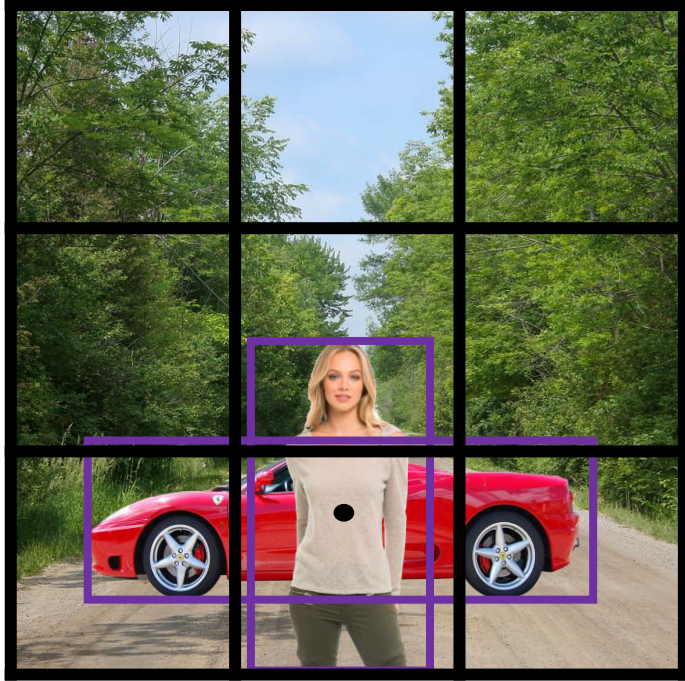$(P_c, b_x, b_y, b_H, b_W, c_1, c_2, c_3)$

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.

Anchor box 1
Anchor Box 2
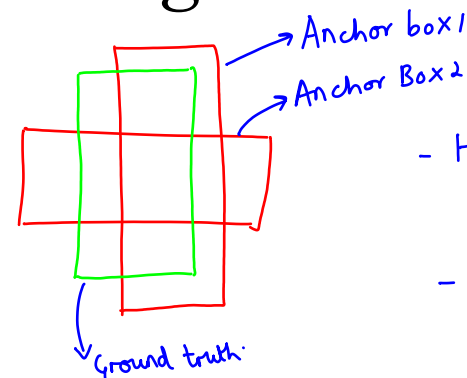
- Highest IoU will be of Anchor Box 1 wrt ground truth
- There can be more than 2 Anchor Boxes

Ground truth.

Andrew Ng

Anchor box example

Anchor box 1:

→ pedestrian

Anchor box 2:

→ car

→ Helps specialized output
tall skinny → for person
flat, long → car

Drawbacks

① If we had say 3 objects
in a grid cell & only 2
Anchor boxes?
This wouldn't work

② What if Both Anchor boxes
were horizontal shaped?
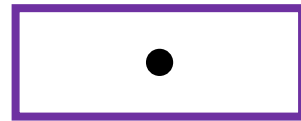You would
need to design
your Algo such that the
bigger IoU Anchor box
is chosen

Ø How to choose
Anchor Box shape?
① By hand
② Kmeans - find
shapes that fit
your dataset well.

$$
y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \\ p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}
\begin{bmatrix} 1 \\ b_x \\ b_y \\ b_H \\ b_w \\ 1 \\ 0 \\ 0 \\ 1 \\ b_x \\ b_y \\ b_H \\ b_w \\ 0 \\ 1 \\ 0 \end{bmatrix}
$$

If there was
No human

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

} Anchor Box 1

} Anchor Box 2

Andrew Ng