



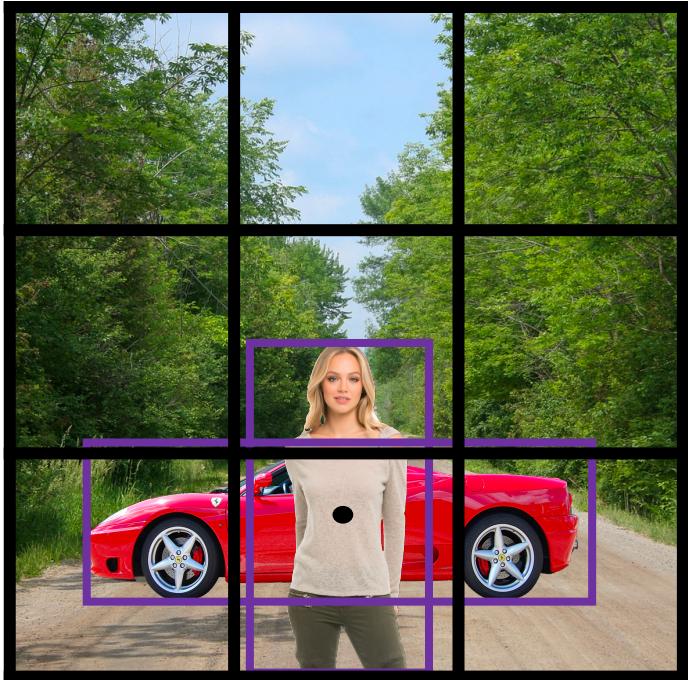
deeplearning.ai

Object Detection

Anchor boxes

여태까지 배운 object detection의 문제점 중 하나는 한 grid cell은 오직 한 개의 object만 detect 할 수 있다는 점이다.

Overlapping objects:



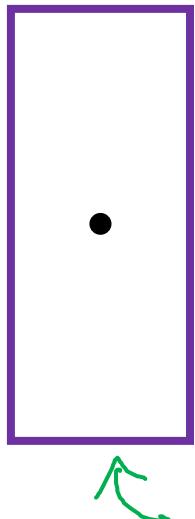
사람과 차의 midpoint가
거의 같은 곳에 위치하며
같은 grid cell에 들어간다.

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

y 의 output이 이런 vector라면
2개의 detection을 산출하지
못할 것이다.

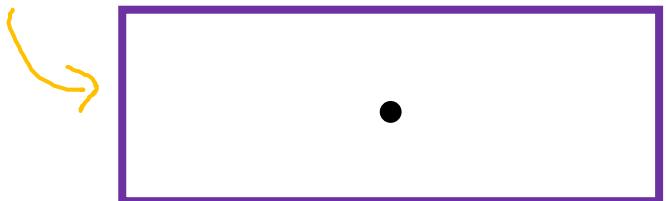
2개의 앵커박스를 predefine하고,
2개의 앵커박스로 2 개의 prediction을 associate할 수 있다.

Anchor box 1:

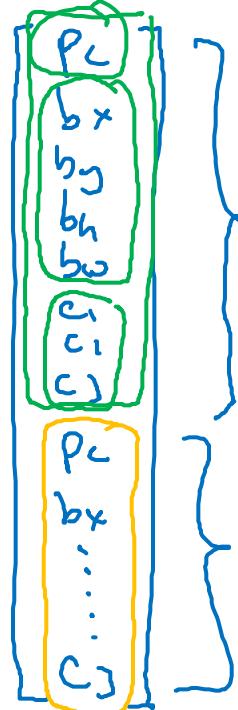


사람 모양이 앵커박스 2보다 1에 더
가깝기 때문에 상단 8개 숫자를
pedestrian 정보를 encoding하기
위해 사용할 수 있다.

Anchor box 2:



$$y =$$



Anchor box 1
8 outputs associated with
anchor box 1

Anchor box 2

Anchor box algorithm

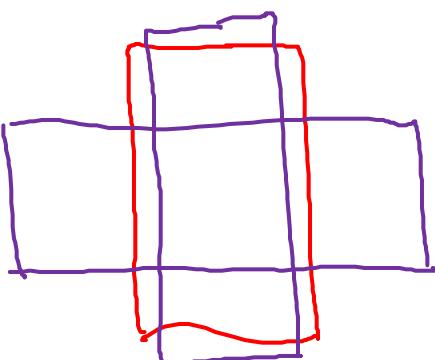
Previously:



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

Output y :

$$\underline{3 \times 3} \times \underline{8}$$



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.

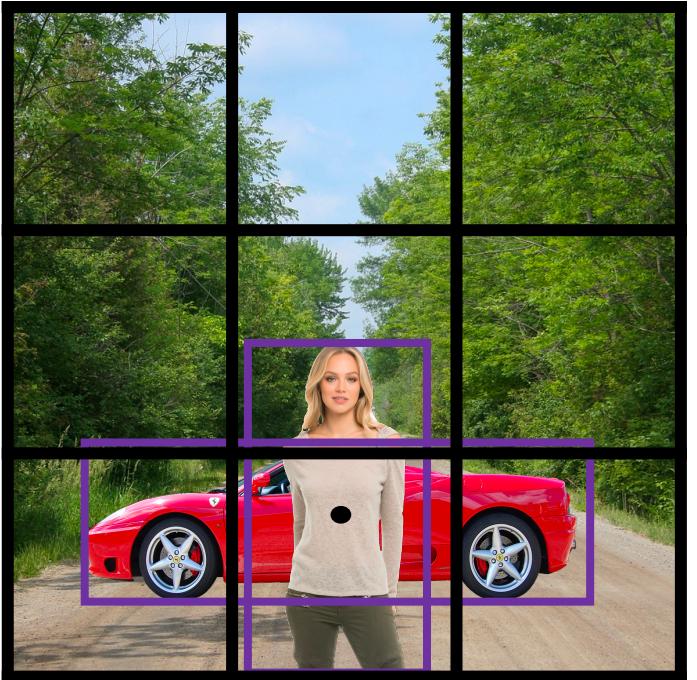
각 오브젝트는 grid cell과 anchor box 페어
(grid cell, anchor box)
에 assigned 된다.

Output y :

$$3 \times 3 \times \underline{16}$$

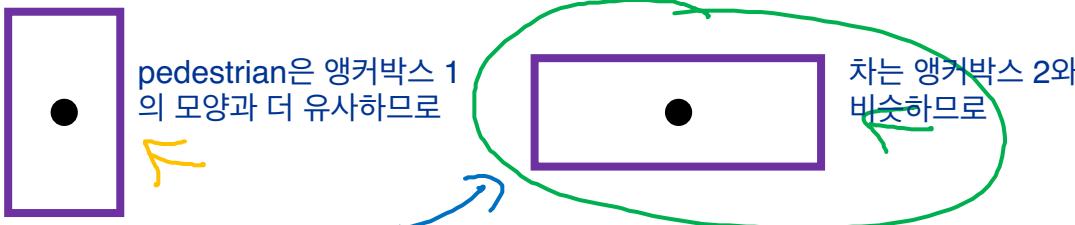
$$3 \times 3 \times \underline{3} \times \underline{8}$$

Anchor box example



요 grid cell에 대한 y를 기술해보자

Anchor box 1: Anchor box 2:



이 방법론이 잘 동작하지 못하는 두 가지 경우가 있다:

1. 두 개의 앵커를 갖고 있는데 3개의 오브젝트가 같은 grid cell에 있는 경우
2. 같은 grid cell에 연관된 두 개의 오브젝트들이 같은 앵커 박스를 가질 경우

$y =$

