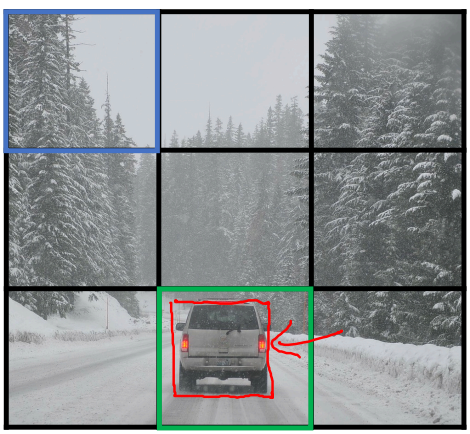
deeplearning.ai

Object Detection

Putting it together: YOLO algorithm

Training



1 - pedestrian
2 - car
3 - motorcycle

y =

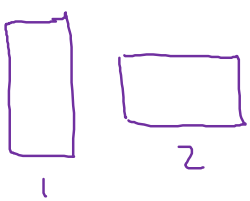
p_c b_x b_y b_h b_w c_1 c_2 c_3 p_c b_x b_y b_h b_w c_1 c_2 c_3

0 ? ? ? ? ? ? ? 0 ? ? ? ? ? ? ?

Anchor 1 Anchor 2

Cell 1

0 ? ? ? ? ? ? ? 1 b_x b_y b_h b_w 0 1 0



3 x 3 x 16

y is 3×3×2×8

(In Pratice) 19x19x16
19x19x40

#anchors

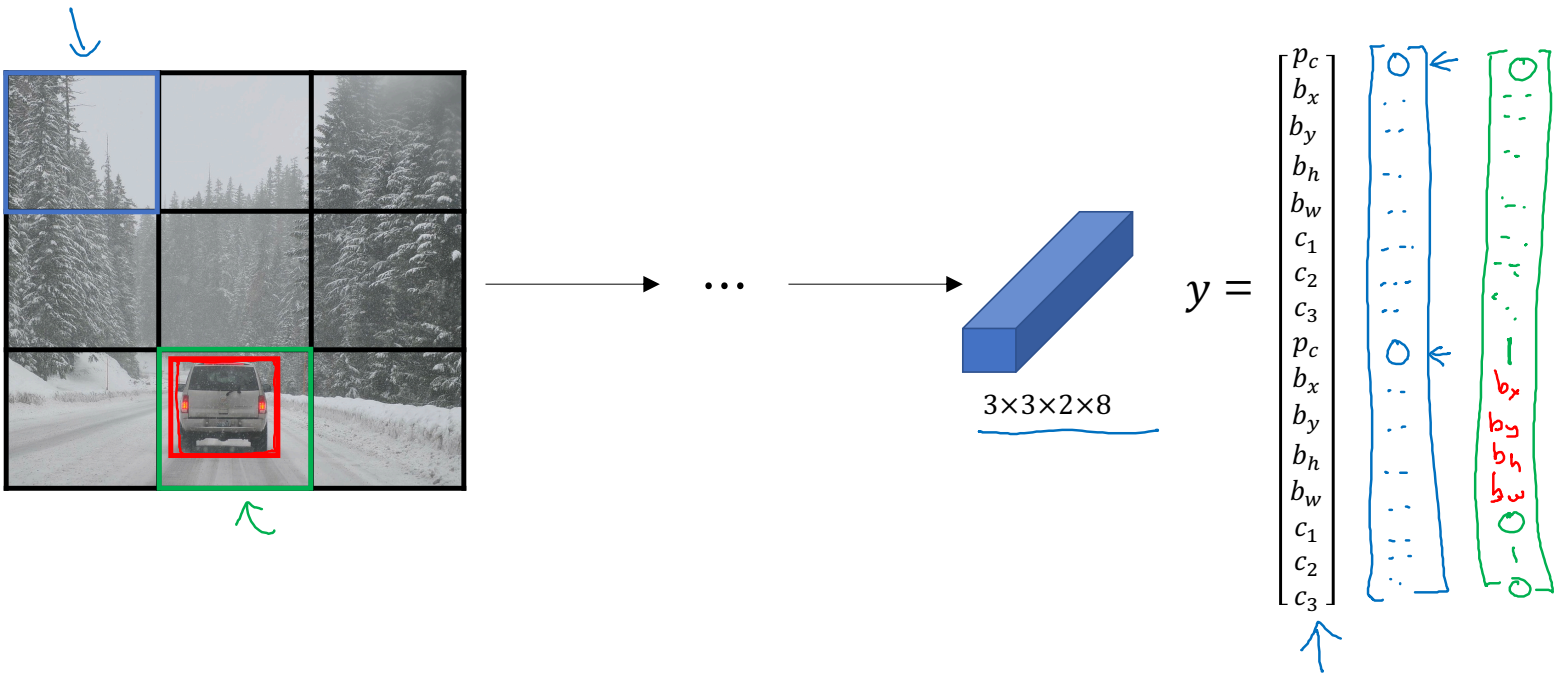
5 + #classes

100x100x3 → ConvNet → 3x3x16

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

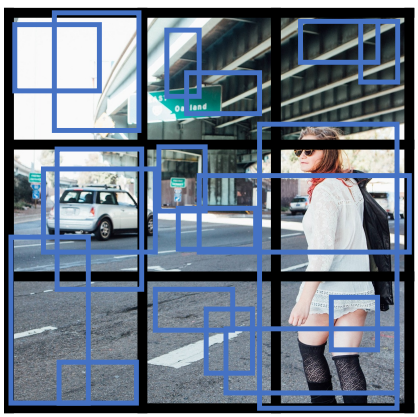
Andrew Ng

Making predictions



Andrew Ng

Outputting the non-max supressed outputs



- For each grid call, get 2 predicted bounding boxes.
- Get rid of low probability predictions.
- For each class (pedestrian, car, motorcycle) use non-max suppression to generate final predictions.

Andrew Ng