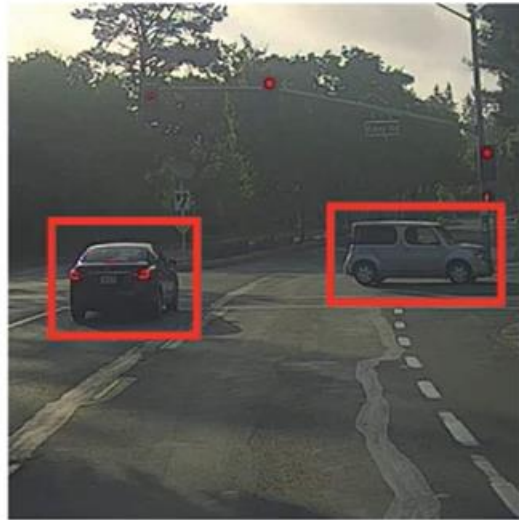


Lecture07

YOLO



Sliding window detection



Training set:

x

y



1



1



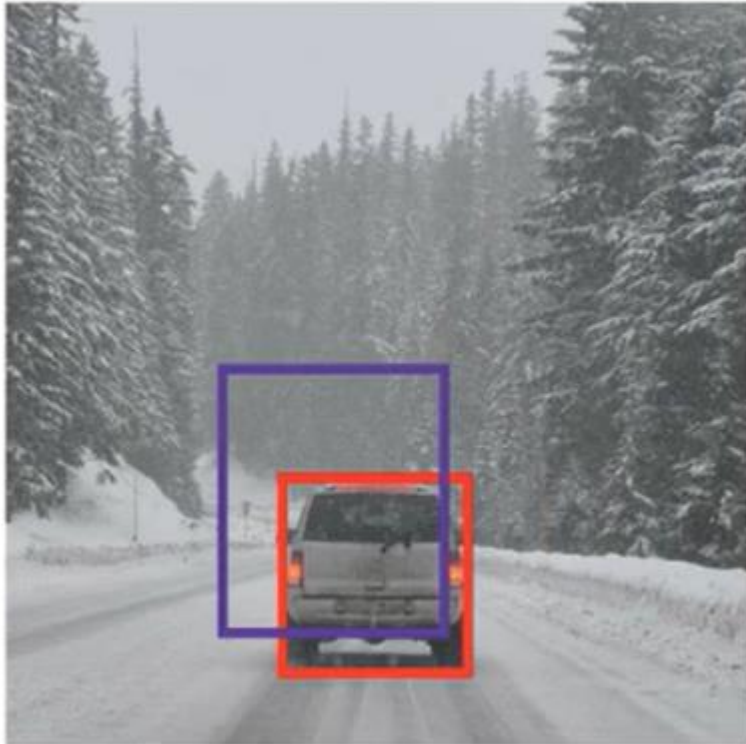
1



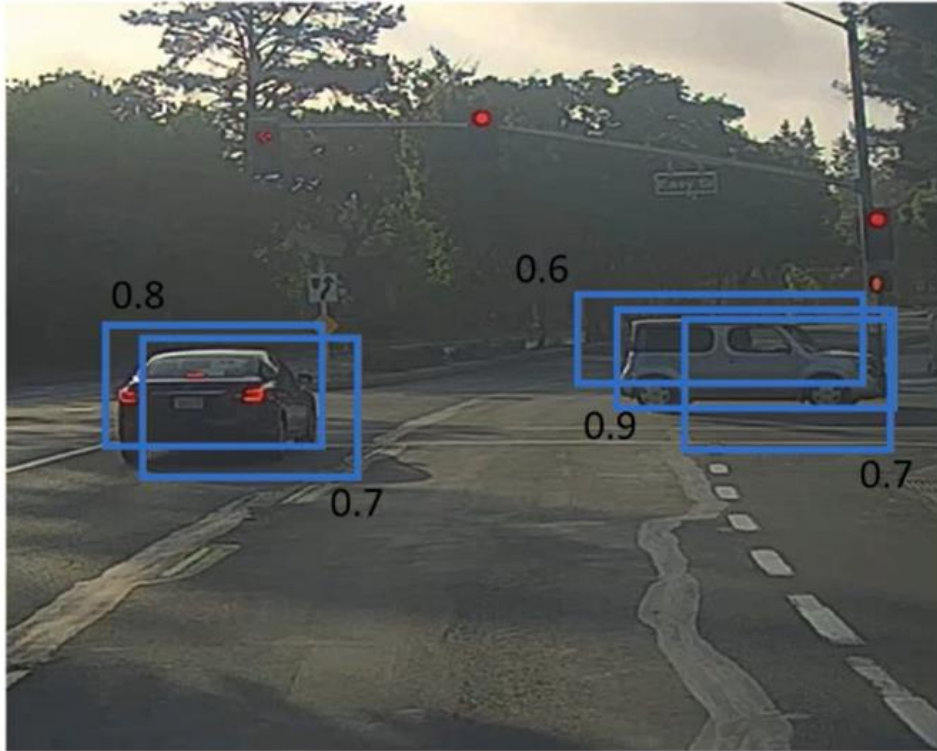
0



Intersection over union (IoU)



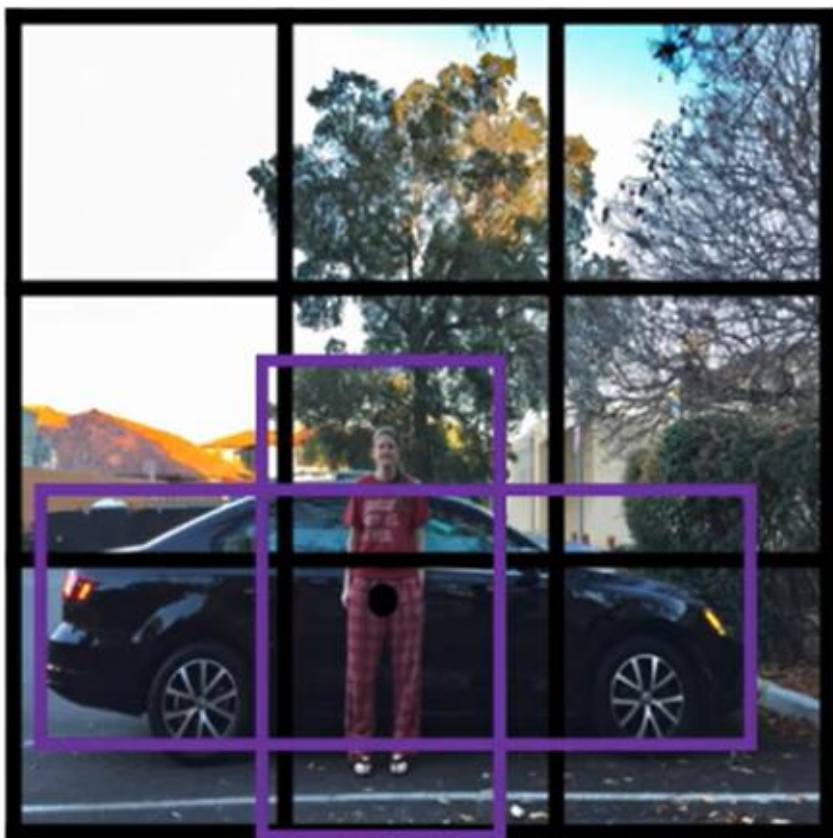
Non-max supression



Objetos traslapados



Objetos traslapados



Anchor box 1:



Anchor box 2:



$$y = \begin{bmatrix} P_c^{(1)} \\ b_x^{(1)} \\ b_y^{(1)} \\ b_h^{(1)} \\ b_w^{(1)} \\ c_1^{(1)} \\ c_2^{(1)} \\ \vdots \\ c_j^{(1)} \\ P_c^{(2)} \\ b_x^{(2)} \\ b_y^{(2)} \\ b_h^{(2)} \\ b_w^{(2)} \\ c_1^{(2)} \\ c_2^{(2)} \\ \vdots \\ c_j^{(2)} \end{bmatrix}$$



Anchor box (caja ancla)

Algoritmo de caja ancla

Anteriormente: Cada objeto en la imagen de entrenamiento se asigna a la celda de la cuadrícula que contiene el punto medio de ese objeto.

Con dos cajas ancla: Cada objeto en la imagen de entrenamiento se asigna a la celda de la cuadrícula que contiene el punto medio del objeto y a la caja ancla para la celda de la cuadrícula con el mayor IoU (Intersección sobre Unión).



YOLO



y is $3 \times 3 \times 2 \times 8$

- 1 - pedestrian
- 2 - car
- 3 - motorcycle

$y =$

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

$$\begin{aligned} & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\ & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\ & + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\ & + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \\ & + \sum_{i=0}^{S^2} \mathbf{1}_{ij}^{\text{obj}} \sum_{c \in \text{Classes}} (p_i(c) - \hat{p}_i(c))^2 \end{aligned}$$



Non-max suppressed outputs



Non-max suppression para cada clase

