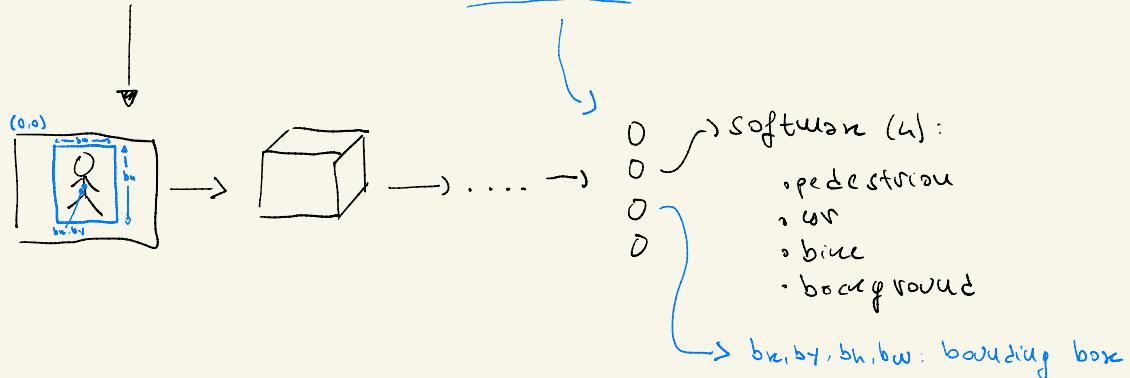



CONV NETS - OBJECT DETECTION

Image Classification + Image Classification with Localization + Detection



→ Target label y :

There will be prob. 1 - pedestrian
2 - car
3 - bike
actually 4 - background

+ b_x, b_y, b_w, b_h

$$\rightarrow \gamma =$$

$$L(\hat{y}, y) = \begin{cases} P_c=1 & = (\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2 + \dots + (\hat{y}_8 - y_1)^2 \\ \text{or} \\ y_1=1 & \\ P_c=0 & = (\hat{y}_1 - y_1)^2 \\ \text{or} \\ y_1=0 & \end{cases}$$

$$\gamma = \begin{bmatrix} P_c \\ b_x \\ b_y \\ b_w \\ b_h \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}$$

1 or 0: is there an obj?
is it a pedestrian?
a car? a bike?

[LANDMARK DETECTION]

→ Define coordinates x, y for the landmark of interest:

l_{lx}, l_{ly} = left corner of an eye

r_{rx}, r_{ly} = right = = = =

m_{mx}, m_{ly} = center of mouth

→ Build a training set, such as: (x, y) where

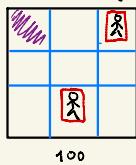
$$y = \begin{bmatrix} 1 \\ l_{lx} \\ l_{ly} \\ r_{rx} \end{bmatrix}$$

is the sparse

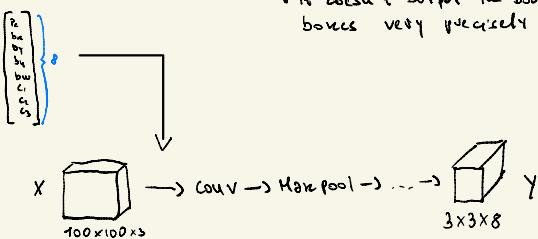
→ Build a convert, such as the output has the same size

Object Detection

- Sliding windows detection: too slow and computational costly
- Sliding windows detection, convolutionally implemented: moves all the prediction at the same time, making it much faster.
- Bounding box prediction: yolo algorithm:



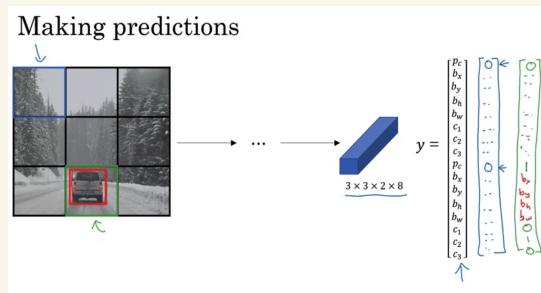
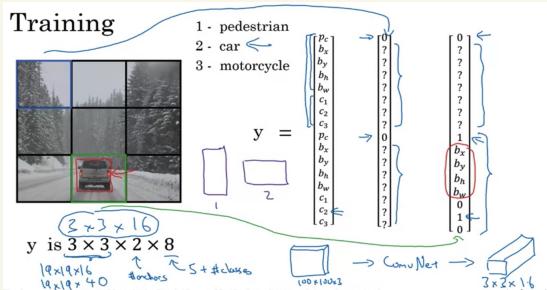
- Divide the img in n grids
 - For each grid the training labels are: $y = \begin{bmatrix} p_x \\ b_x \\ b_y \\ b_w \\ b_h \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$
 - Target output volume: \rightarrow input-output match
1 object per grid cell \rightarrow limitation
- $3 \times 3 \times 8$
- \rightarrow it outputs precise bounding box
(1 object per grid cell \rightarrow limitation)
- \rightarrow Convolutional implementation



Bounding box encoding: b_x, b_y, b_w, b_h are specific to the grid in question

- Intersection over Union (IoU): to evaluate obj localization. It's a measure of the overlap between 2 bounding boxes
- Non-Max Suppression: allows you to suppress bounding boxes that have a lower probability associated
- Anchor boxes: multiple objects in 1 grid

YOLO Algorithm:



+ outputting the non-max suppressed outputs