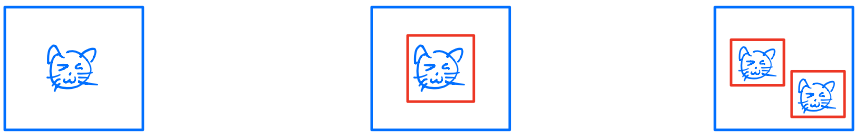
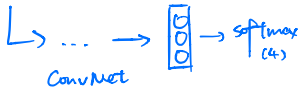
Object Localization

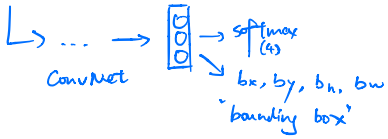
Image Classification

Classification w/ Localization

Detection







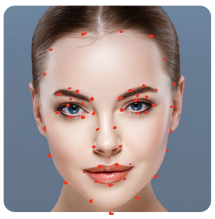
Define the target label y

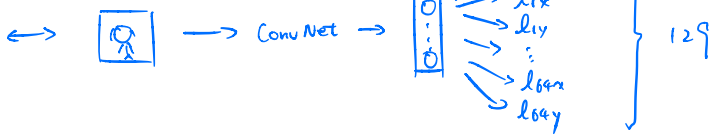
$$\Rightarrow y = \begin{bmatrix} P_c \\ b_x \\ b_y \\ b_h \\ b_w \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}$$ is there any object?

e.g. $\begin{bmatrix} 1 \\ 0.4 \\ 0.4 \\ 0.2 \\ 0.2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ ? \\ ? \\ \vdots \\ ? \end{bmatrix}$

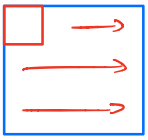
$$\Rightarrow L(\hat{y}, y) = \begin{cases} (\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2 + \cdots + (\hat{y}_8 - y_8)^2 & \text{if } y_1 = 1 \\ (\hat{y}_1 - y_1)^2 & \text{if } y_1 = 0 \end{cases}$$

Landmark Detection

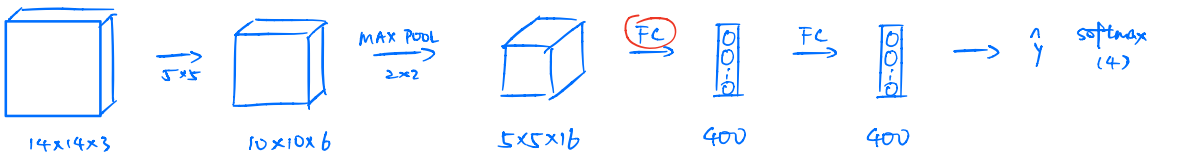


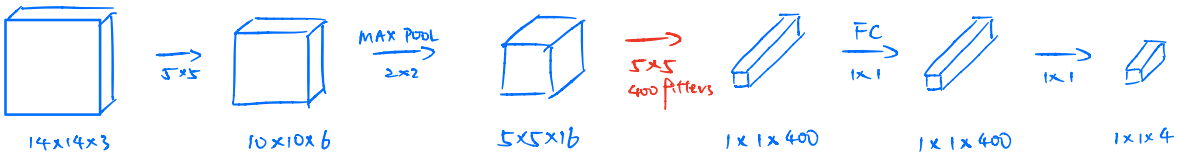


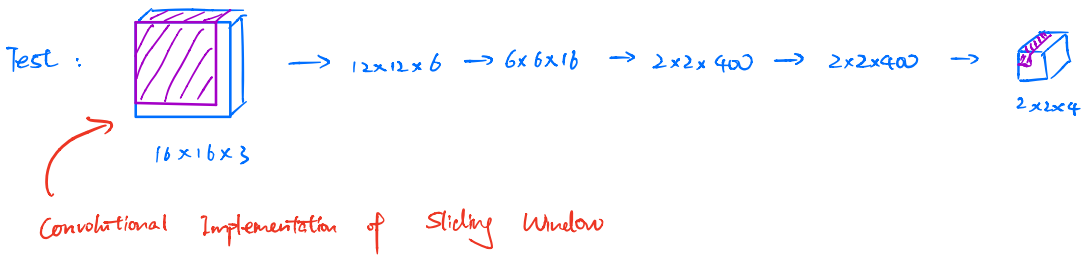
Object Detection — Sliding Window



Computation Cost



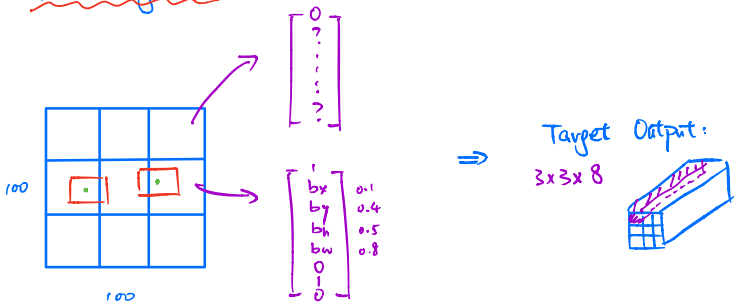




Convolutional Implementation of Sliding Window

Bounding Box Prediction

YOLO Algorithm



⇒ Intersection over Union (IoU)

$$IoU = \frac{\text{Size of } \square}{\text{Size of } \square}$$

"Correct" if IoU ≥ 0.5, 0.6

Non – Max Suppression

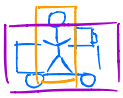
Multiple Detection on the same object

Pick the box w/ the largest P_c

Suppress others w/ high IoU ≥ 0.5

Anchor Boxes

Overlapping objects:



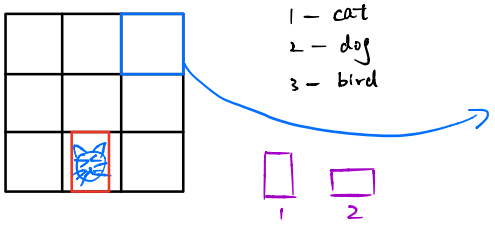
$$y = \begin{bmatrix} P_c \\ b_x \\ b_y \\ b_h \\ b_w \\ C_1 \\ C_2 \\ C_3 \end{bmatrix} \Rightarrow$$

Anchor Box 1

Anchor Box 2

$$y = \begin{bmatrix} P_c \\ b_x \\ \vdots \\ C_3 \\ P_c \\ b_x \\ \vdots \\ C_3 \end{bmatrix}$$ Anchor Box 1 / Anchor Box 2

YOLO Algorithm



y is $3 \times 3 \times (2 \times 8)$ ← $5 +$ # classes

↑ # anchors

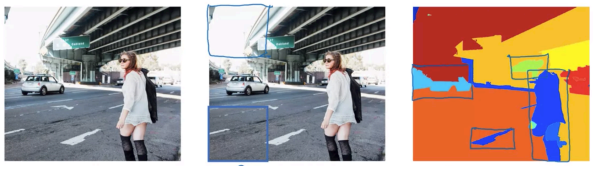
in practice: $19 \times 19 \times (2 \times 8)$

$19 \times 19 \times (5 \times 8)$

$3 \times 3 \times 16$

Region Proposal

R-CNN



Segmentation Algorithms

~ 2,000 blocks

R-CNN: Propose regions. Classify proposed regions one at a time. Output label + bounding box

Fast R-CNN: Propose regions. Use convolutional implementation of sliding windows to classify all the proposed regions.

Faster R-CNN: Use convolutional network to propose regions.