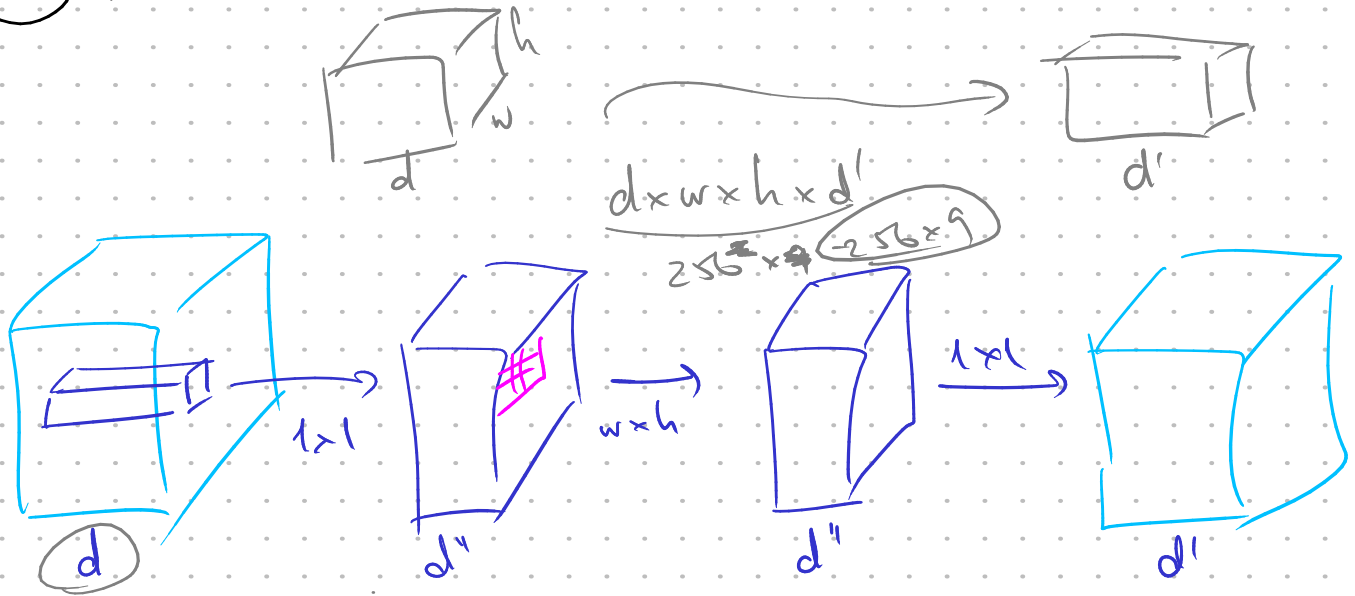


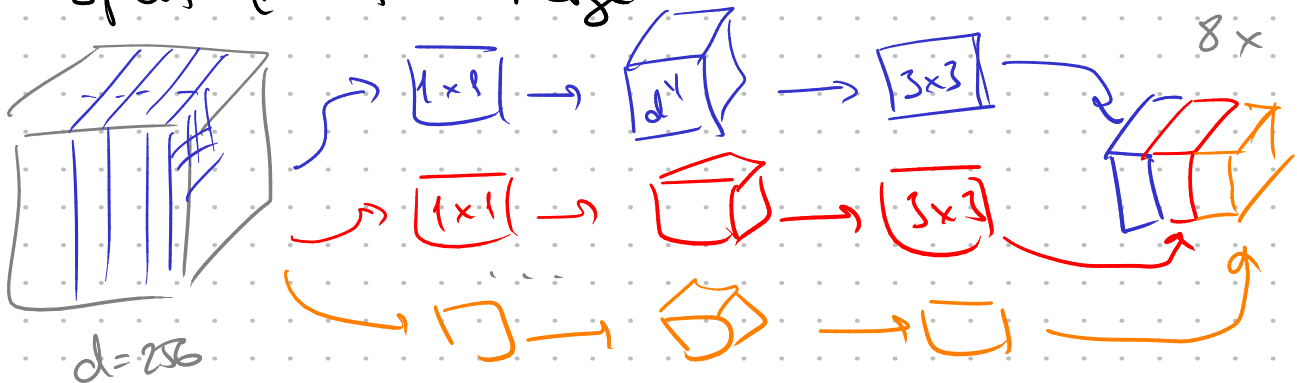
⑤ Bottleneck



$$\begin{matrix}
 d \times 1 \times 1 \times d' & + & d'' \times w \times h \times d'' & + & d'' \times 1 \times 1 \times d' \\
 \boxed{256} & & \boxed{16} & & 16 & & 16 & & 256
 \end{matrix}$$

$$256 \times (16 + 9 + 16) = 256 \times 41$$

⑥ Split-transform-merge $d''=32$



$$8 \times (256 \times 1 \times 1 \times 32 + 32 \times 3 \times 3 \times 32)$$

$$= 256 (8 \times 32 + 8 \times 4 \times 9) = 256 \times 32 \times 17$$

depthwise
convolution

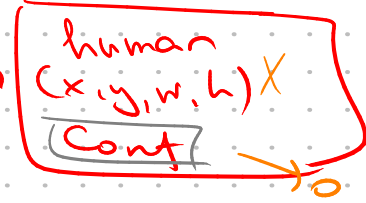
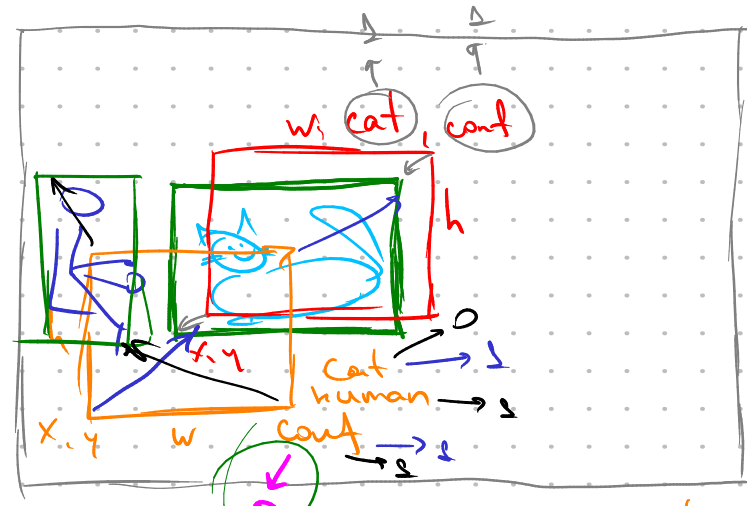
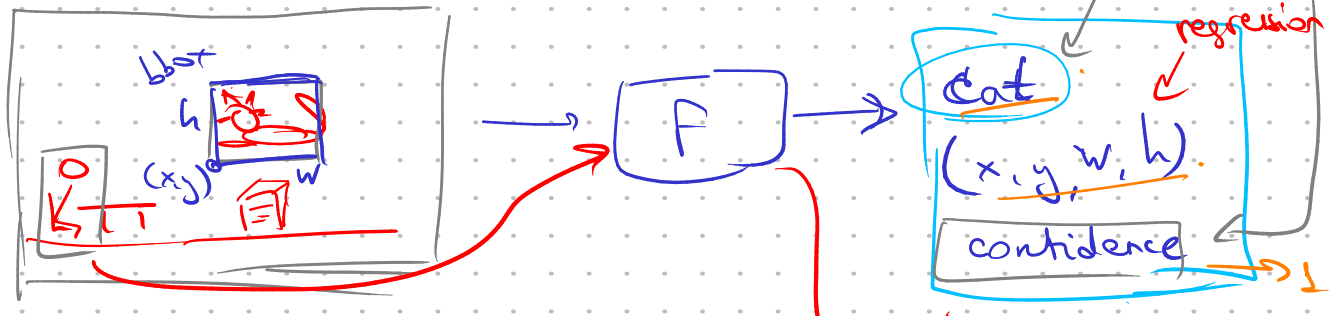
Object detection



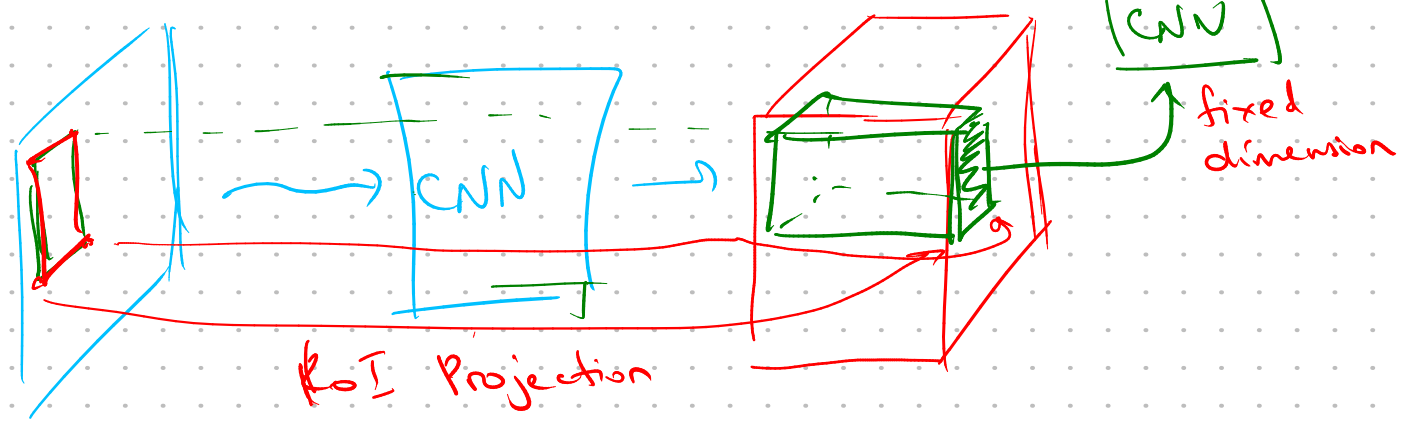
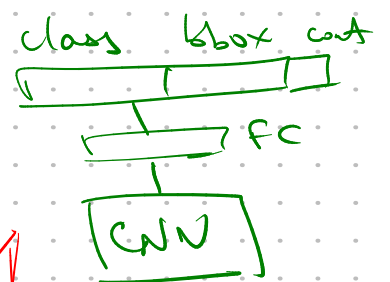
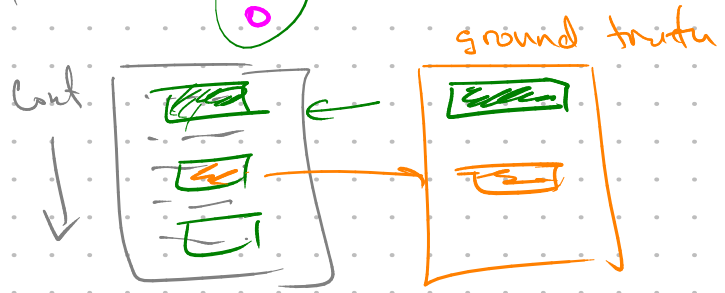
Object localization



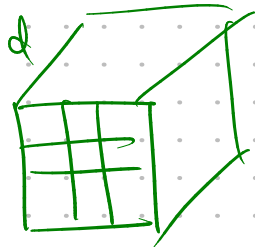
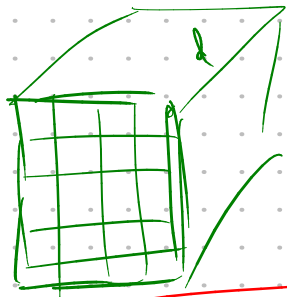
Classification



$$IoU = \text{Jaccard} = \frac{|A \cap B|}{|A \cup B|}$$



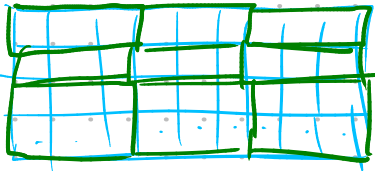
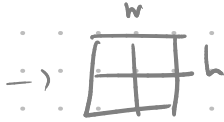
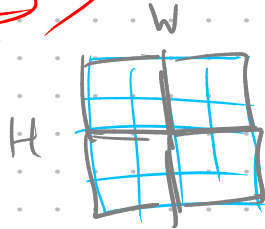
ROI Projection
Region of interest



CNN



RoI pooling



max-pool

