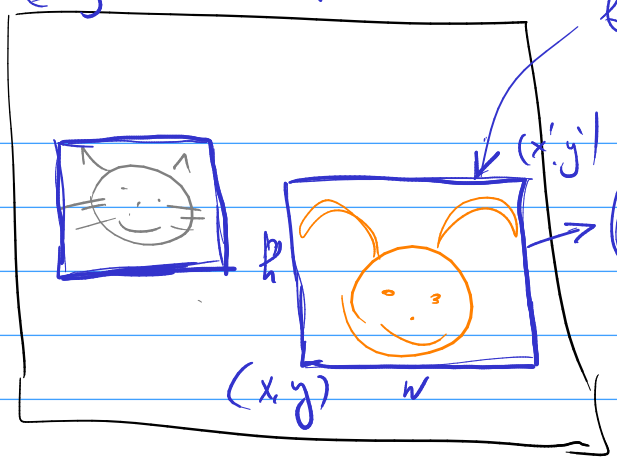
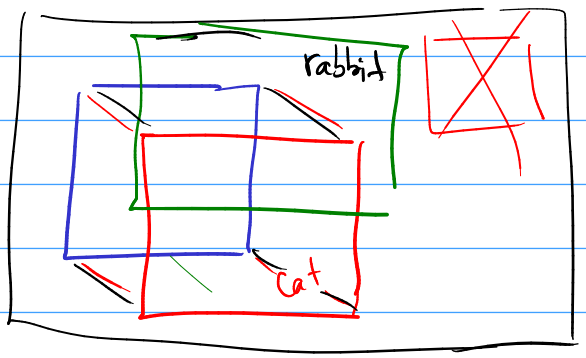
Object localization

Object detection



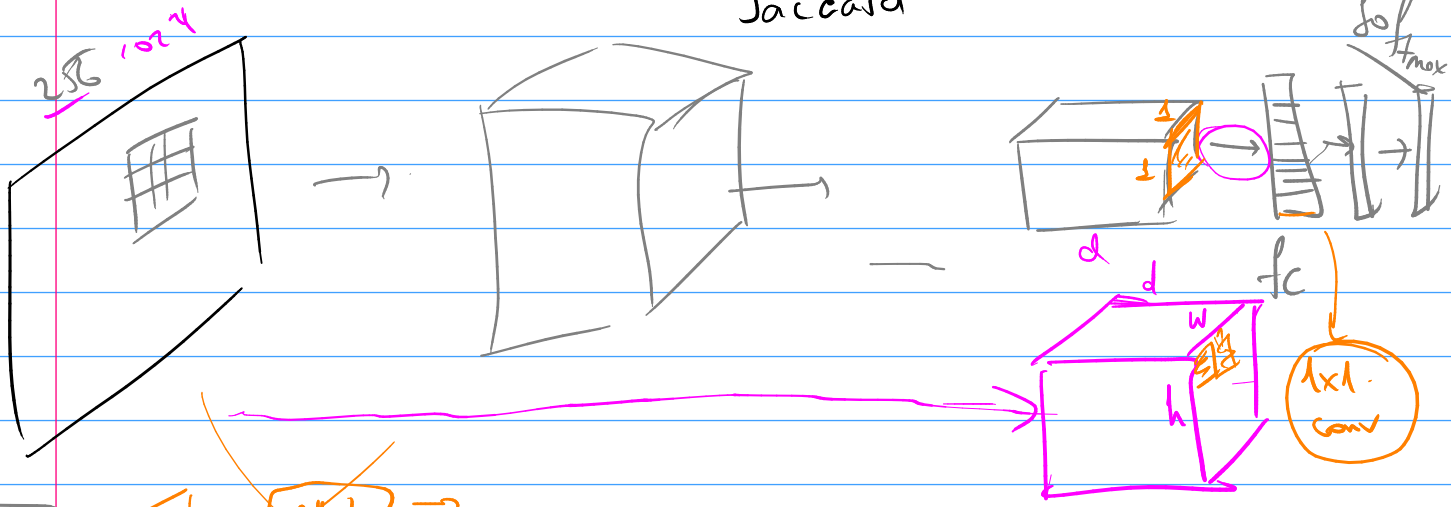


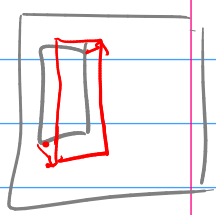
$(p, \bar{c}, x, y, w, h)$

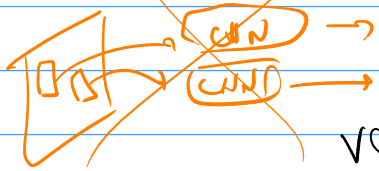
$(\bar{c}, x, y, w, h)$

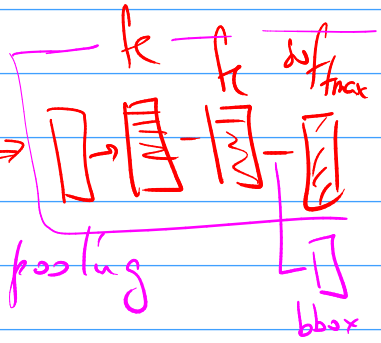
$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

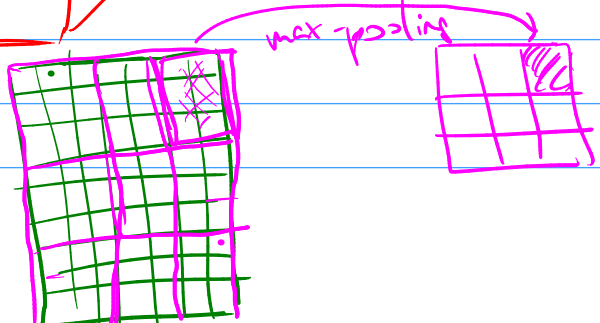
Jaccard

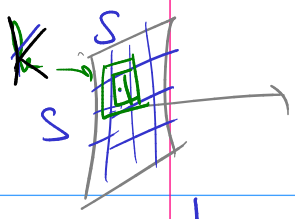












$S \times S \times (C + 5k)$

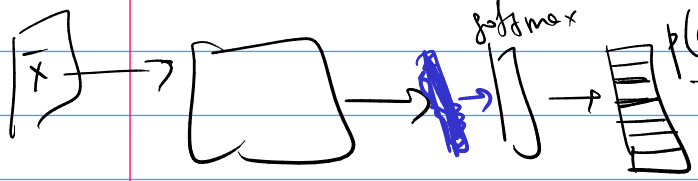
regression bbox

$$L = \sum_{s=1}^{S^2} \sum_{k \geq 1}^{k} \left(\lambda_1 \left[\frac{\text{npob}}{\text{bbox}} \right] \cdot \left((\hat{x} - x)^2 + (\hat{y} - y)^2 + \right. \right.$$

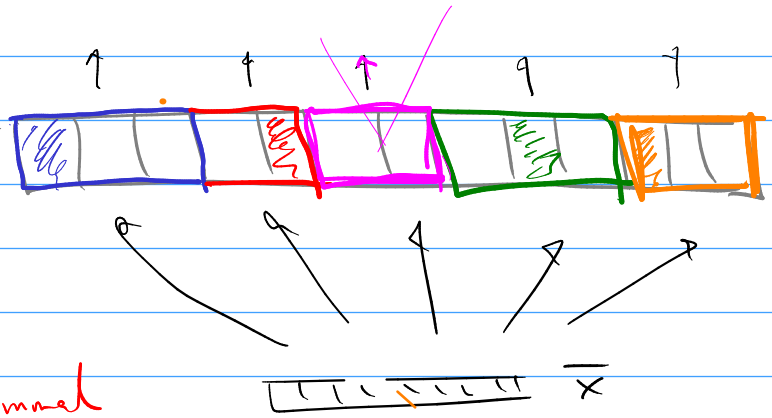
$$\left. + \left(\sqrt{\hat{w}} - \sqrt{w} \right)^2 + \left(\sqrt{\hat{h}} - \sqrt{h} \right)^2 \right) +$$

$$+ \lambda_2 \left[\text{npob\_bbox} \right] \cdot L_{class}(\hat{c}, c)$$

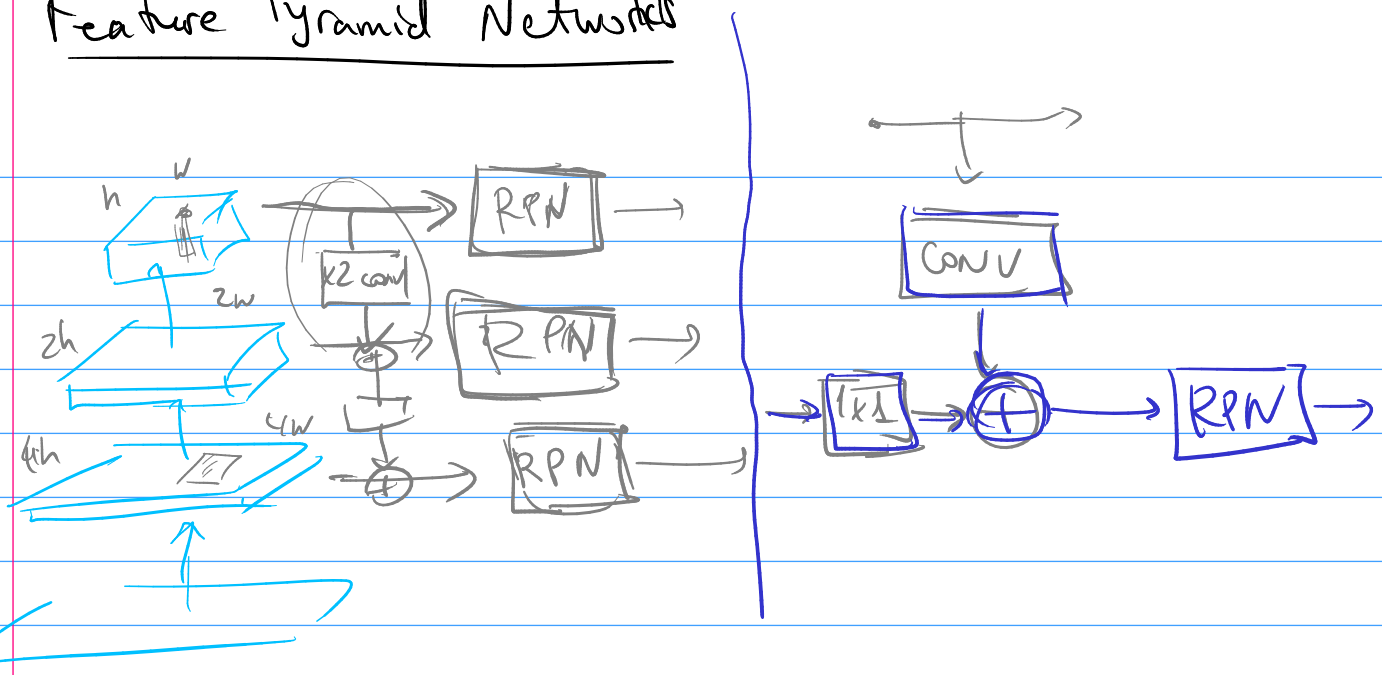
$$\left. + \lambda_3 \left[\frac{\text{npob}}{\text{bbox}} \right] \cdot (p - 1)^2 + \lambda_4 \left[\frac{\text{temp}}{\text{bbox}} \right] \cdot (p - 0)^2 \right)$$



Hierarchical softmax



Feature Pyramid Networks



p - уверенность

$y = 1$, да
$= 0$, нет

$$L(p, y) = -y \log p - (1-y) \log(1-p)$$

$p = 0.8,\ y = 1 \rightarrow L = -\log \frac{4}{5} = 0.32$

10^5

$p = 0.2,\ y = 1 \rightarrow L = -\log \frac{1}{5} = 2.32$

10