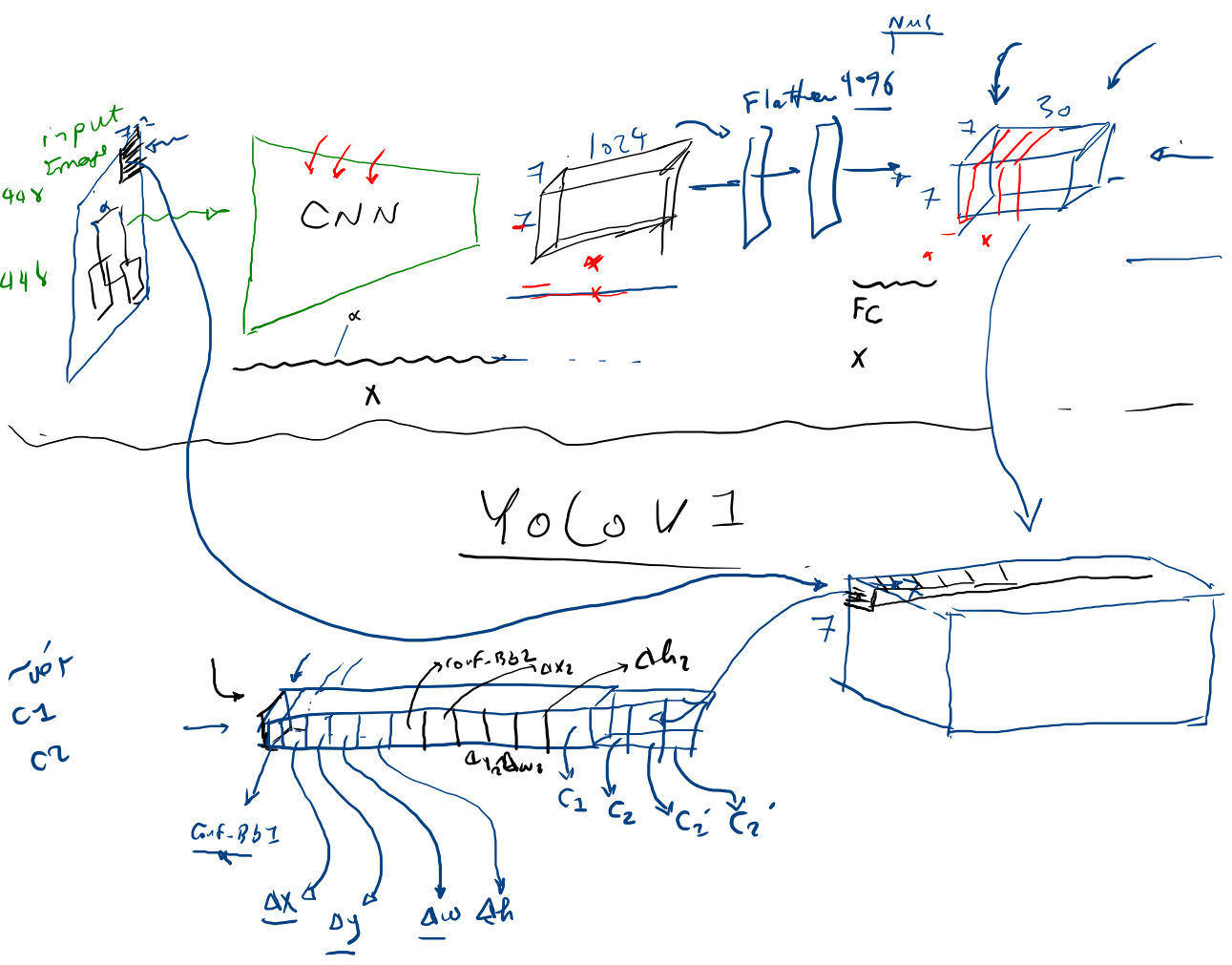
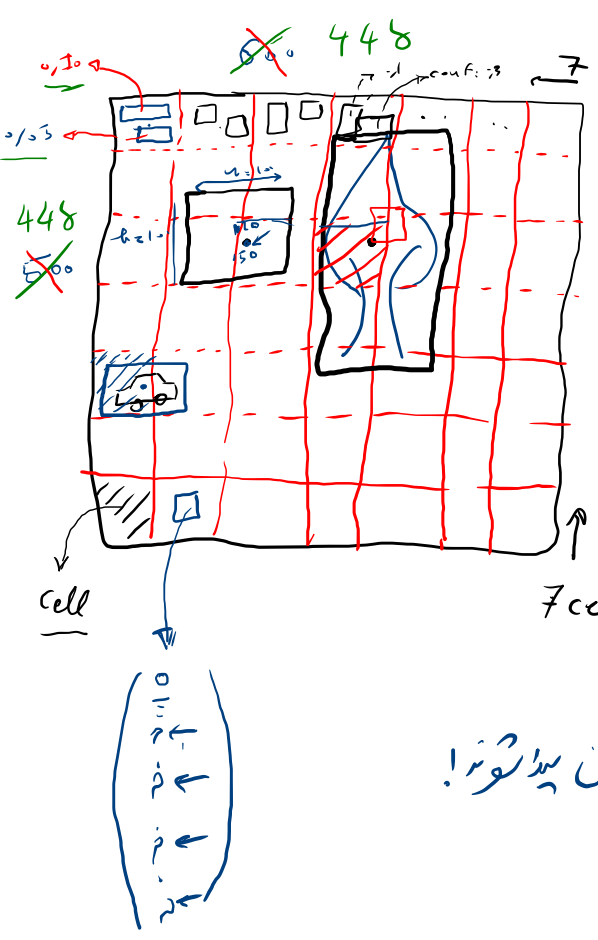
Yolo v1



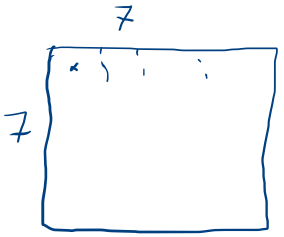
$7 \times 7 \times 30 = 1470$

α



sliding window

Feature map



classes = 2

$S = 7$

تعداد باکس‌ها $B = 2$

تعداد کل: Number of Bbs $= 7 \times 7 \times 2 \times (5 + 7)$

Bb info: $(x, y, w, h, p), (C_1, C_2, \cdots, C_{20})$

۱. تصویر ورودی Resize شود! (448×448)

۲. در نهایت با object جعبه Bounding Box مشخص می‌شود!

جهت سریع‌تر شدن محاسبات، اعداد مختصات به [0,1] مپ شود!

$$\Delta X = \frac{X - X_a}{w_{cell}}$$

(X: center; X_a → X_left top_cell)

$$\Delta y = \frac{y - y_a}{h_{cell}}$$

(y: center; y_a → y_left top_cell)

$$\Delta w = \frac{w}{w_{img}}$$

(w → w_Bb)

$$\Delta h = \frac{h}{h_{img}}$$

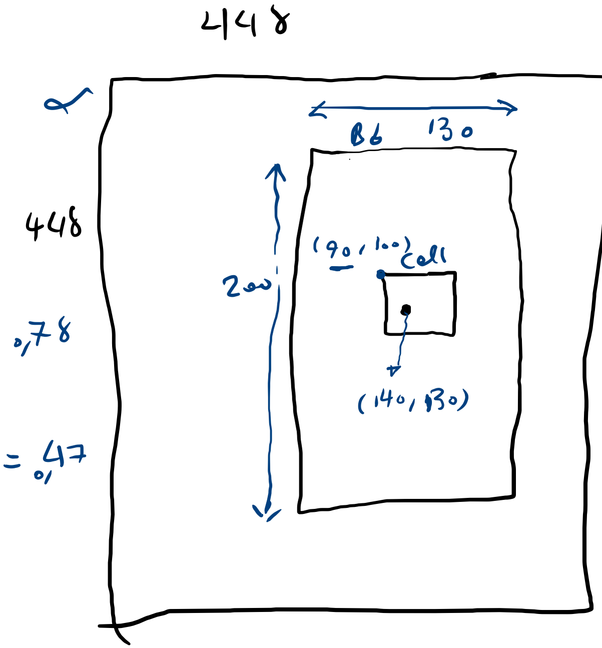
(h → h_Bb)

$$\Delta X = \frac{140 - 90}{\frac{448}{7}} = \frac{50}{64} = 0{,}78$$

$$\Delta y = \frac{130 - 100}{64} = \frac{30}{64} = 0{,}47$$

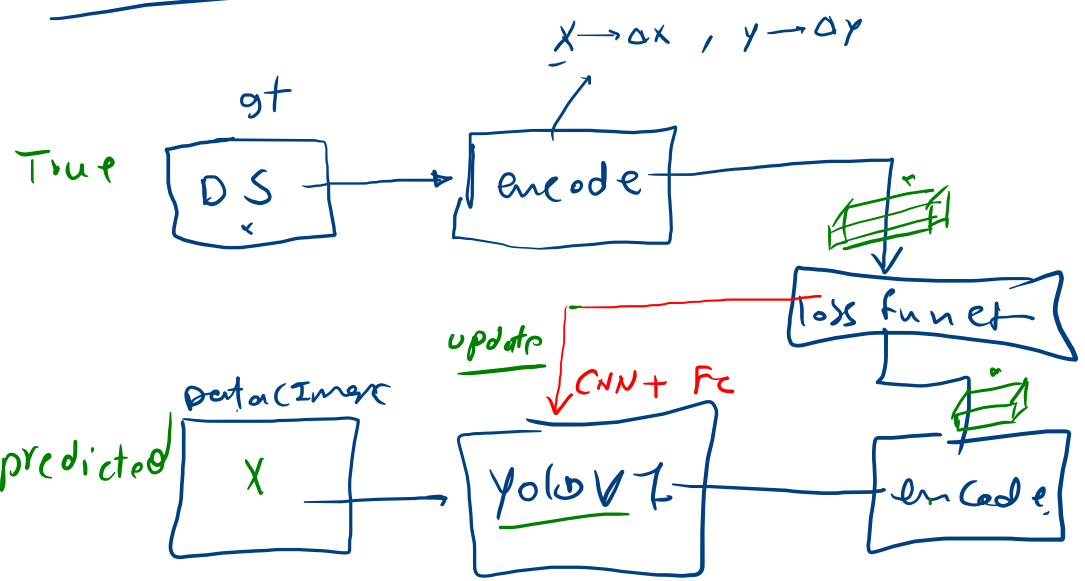
$$\Delta w = \frac{130}{448} = 0{,}29$$

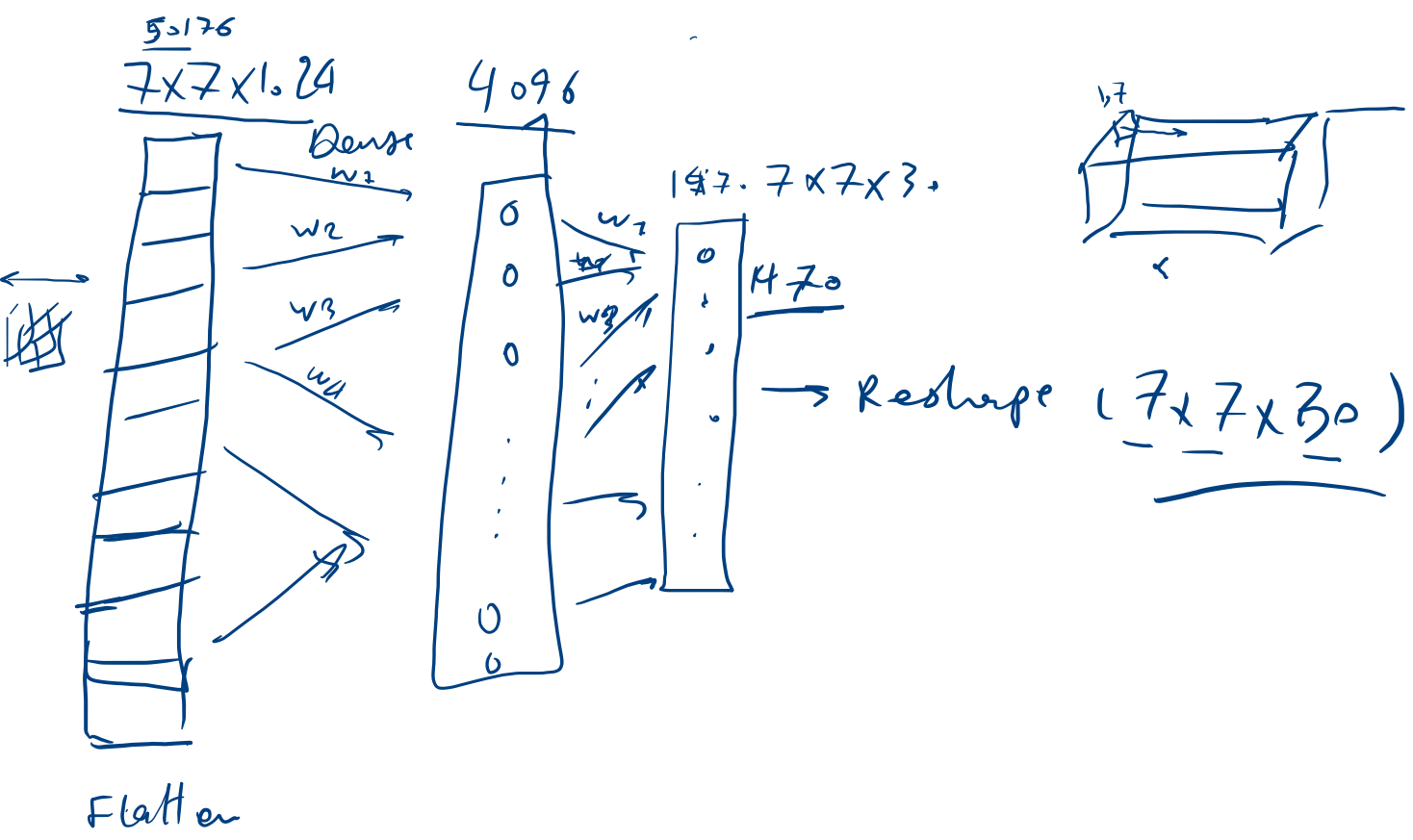
$$\Delta h = \frac{200}{448} = 0{,}45$$



	SSD	Yolo *
mAP	✓	X
FPS	X	✓
	✓	X

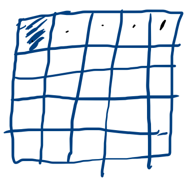
Yolo Train

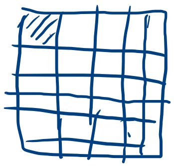

x → ax , y → ay
gt
True
DS
encode
loss funct
update
CNN + Fc
DataCImage
predicted
X
YoloV7
encode


5x176
7x7x1024
Dense
W1
W2
W3
W4
4096
0
0
0
0
0
W1
W2
W3
W4
147. 7x7x30
0
1470
1,7
x
→ Reshape (7x7x30)
Flatten

loss - yolo v1

gt



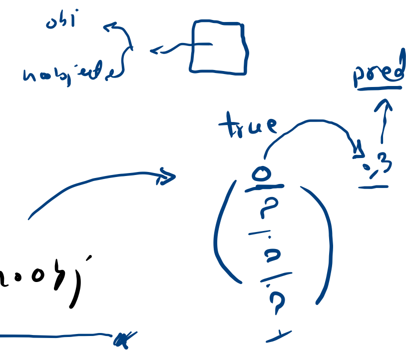


$$L = \sum_{i=1}^{S^2} L_i$$

دو تا loss برای i-ام سلول

$$L_i = 1_i^{obj} \, L_{i-obj} + \lambda \, 1_i^{no-obj} \, L_{i-no\,obj}$$

λ: ,5



$$L_{i-obj} = L^{Bbox}_{i,obj} + L^{conf}_{i,obj} + L^{cls}_{i,obj}$$

Bb Cls

obj

$$L_{i,obj}^{BBox} = \left(\Delta x_i^* - \Delta \hat{x}_i\right)^2 +$$

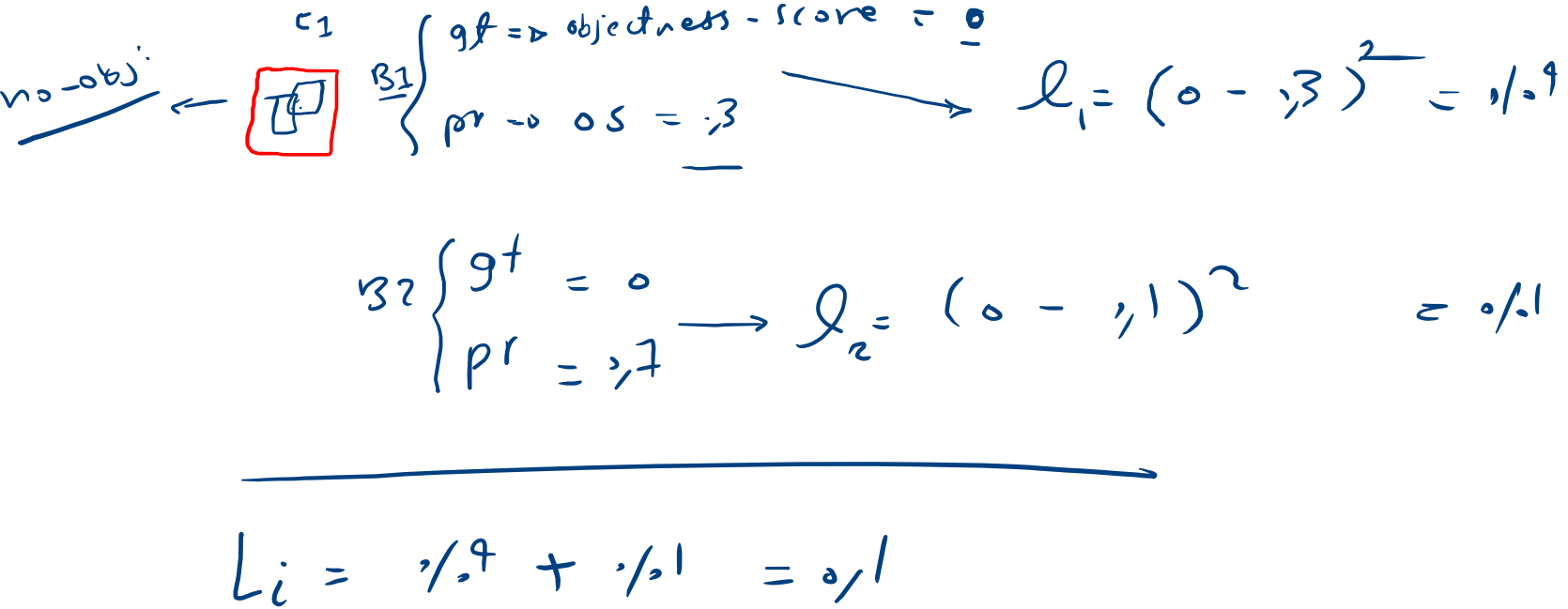
$$\left(\Delta y_i^* - \Delta \hat{y}_i\right)^2 +$$

$$\left(\sqrt{\Delta w_i^*} - \sqrt{\Delta w_i}\right)^2$$

$$+ \left(\sqrt{\Delta h_i^*} - \sqrt{\Delta h_i}\right)^2$$

$$L^{conf} = (c_i^* - \hat{c}_i)^2$$

$$L^{cls} = \sum_{i=1}^{cls} (p_{i,c}^* - \hat{p}_{i,c})^2$$


C1
no-obj
B1
gt => objectness-score = 0
pr-obj = .3
$l_1 = (0 - .3)^2 = .09$
B2
gt = 0
pr = .7
$l_2 = (0 - .1)^2 = .01$
$L_i = .09 + .01 = .1$

The End.