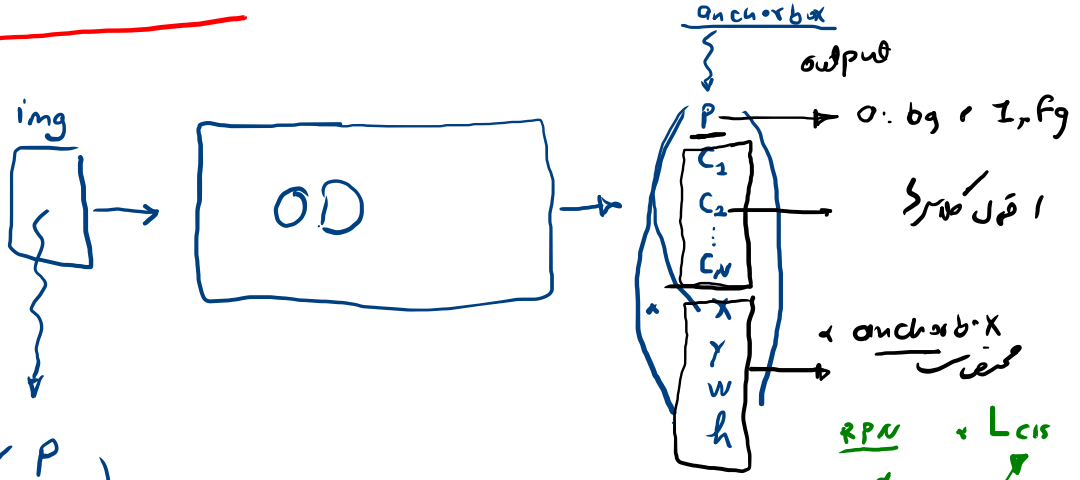
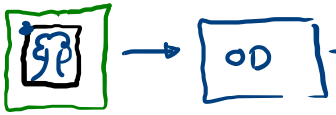
Faster RCNN → object Detection →



gt-bb: $\begin{pmatrix} P \\ C_1 \\ \vdots \\ C_N \\ x \\ y \\ w \\ h \end{pmatrix}$

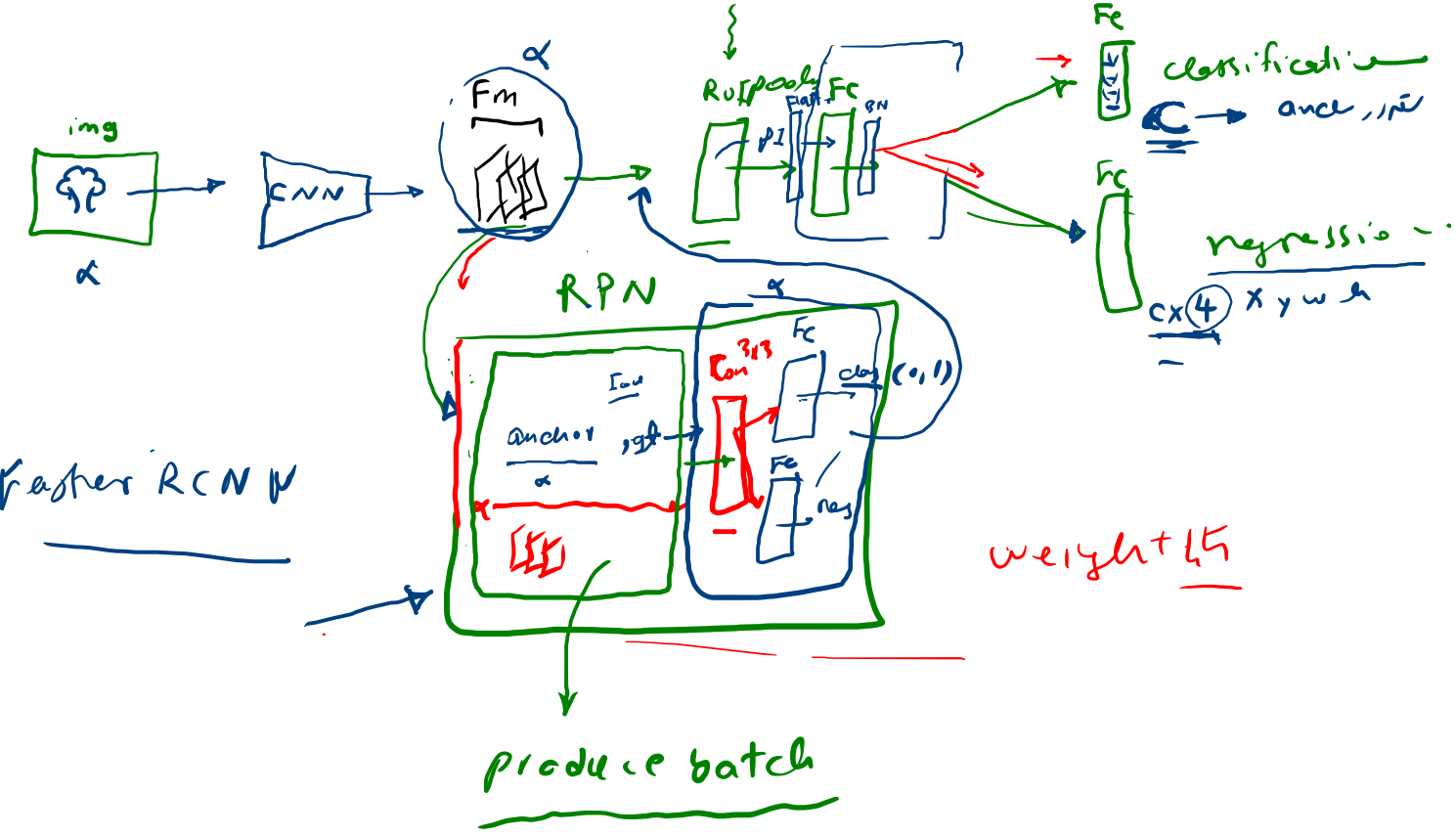


RPN : L_{cls} , L_{reg}

out : [1, 1, 8, 1, 2, 300, 100, 50, 80]

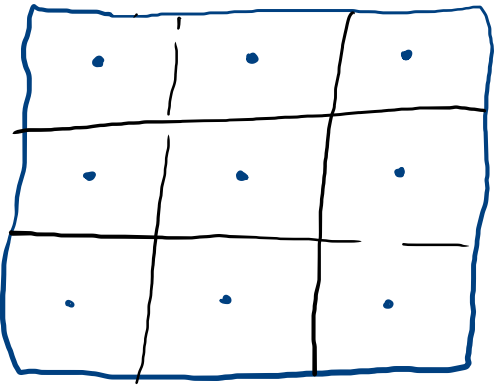
gt : [1, 1, 0, 290, 105, 53, 79]

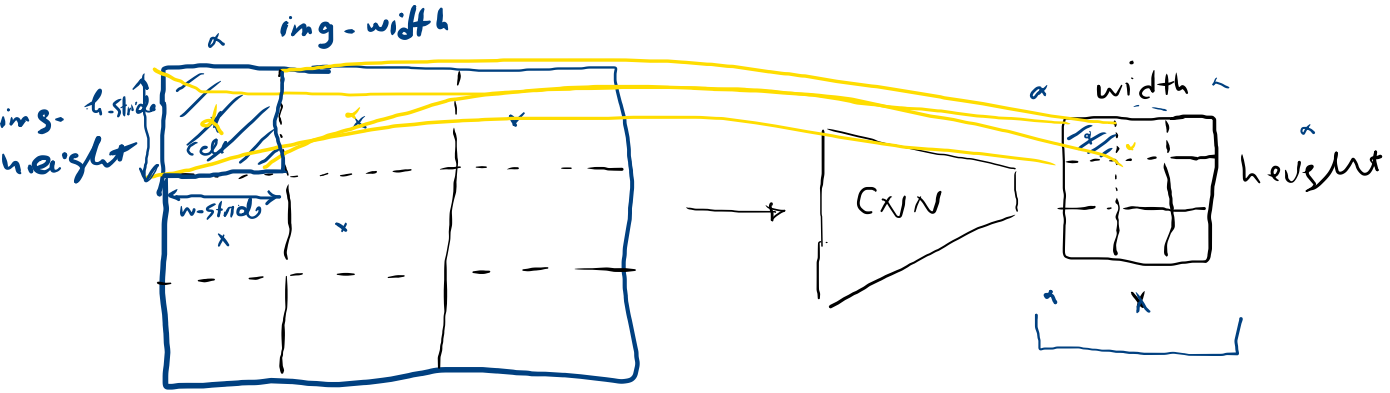
out ≈ gt

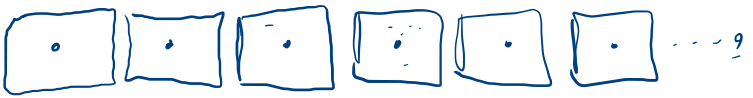

img
CNN
Fm
RoIPool
Flatt
Fc
BN
Fc
Fc
RPN
Conv 3x3
Fc
Fc
anchor
produce batch
weight + 45

End

sliding window





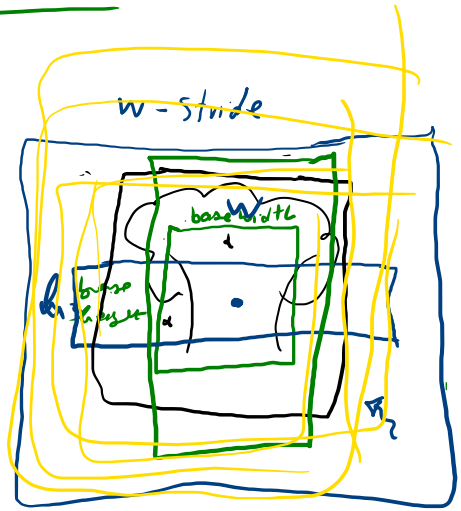


→ num - feature - map = width x height

$$w\text{-}stride = \frac{img\text{-}width}{width}$$

$$h\text{-}stride = \frac{img\text{-}height}{height}$$

generate_anchors



- base_width
- base_height
- ratio : aspect ratio $\frac{h}{w}$ [.5, 1, 2]
- scales → [3, 6, 12]

w = base_width = 1

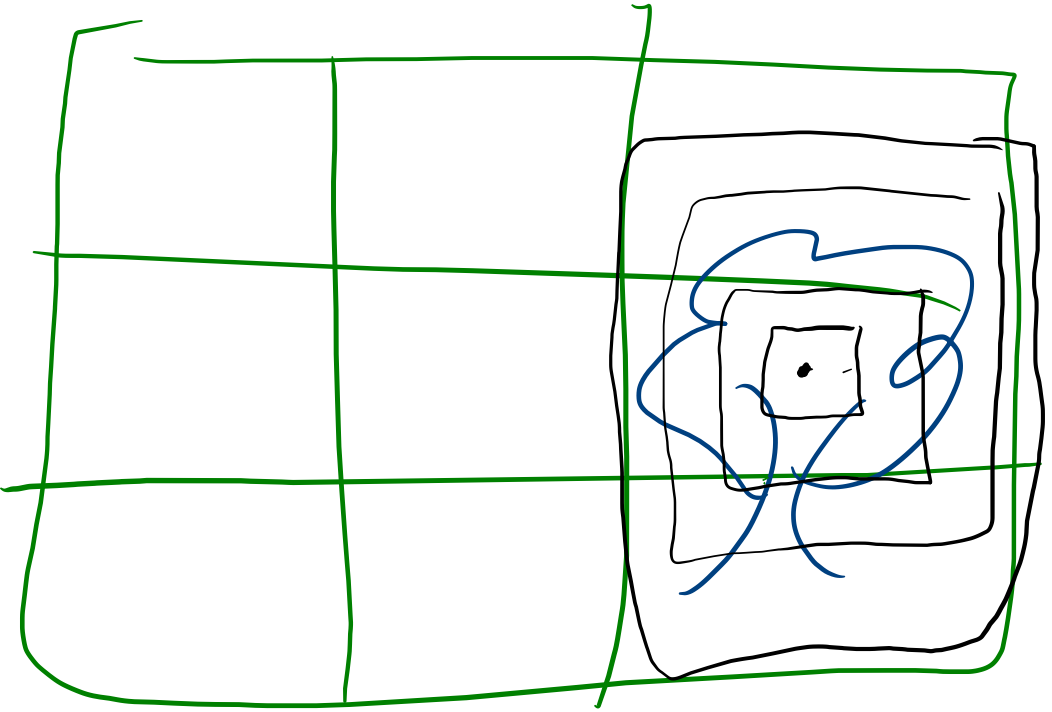
h = base_height = 1 → $w_s = \left(\frac{1 \times 1}{.5}\right)^2 = 4$

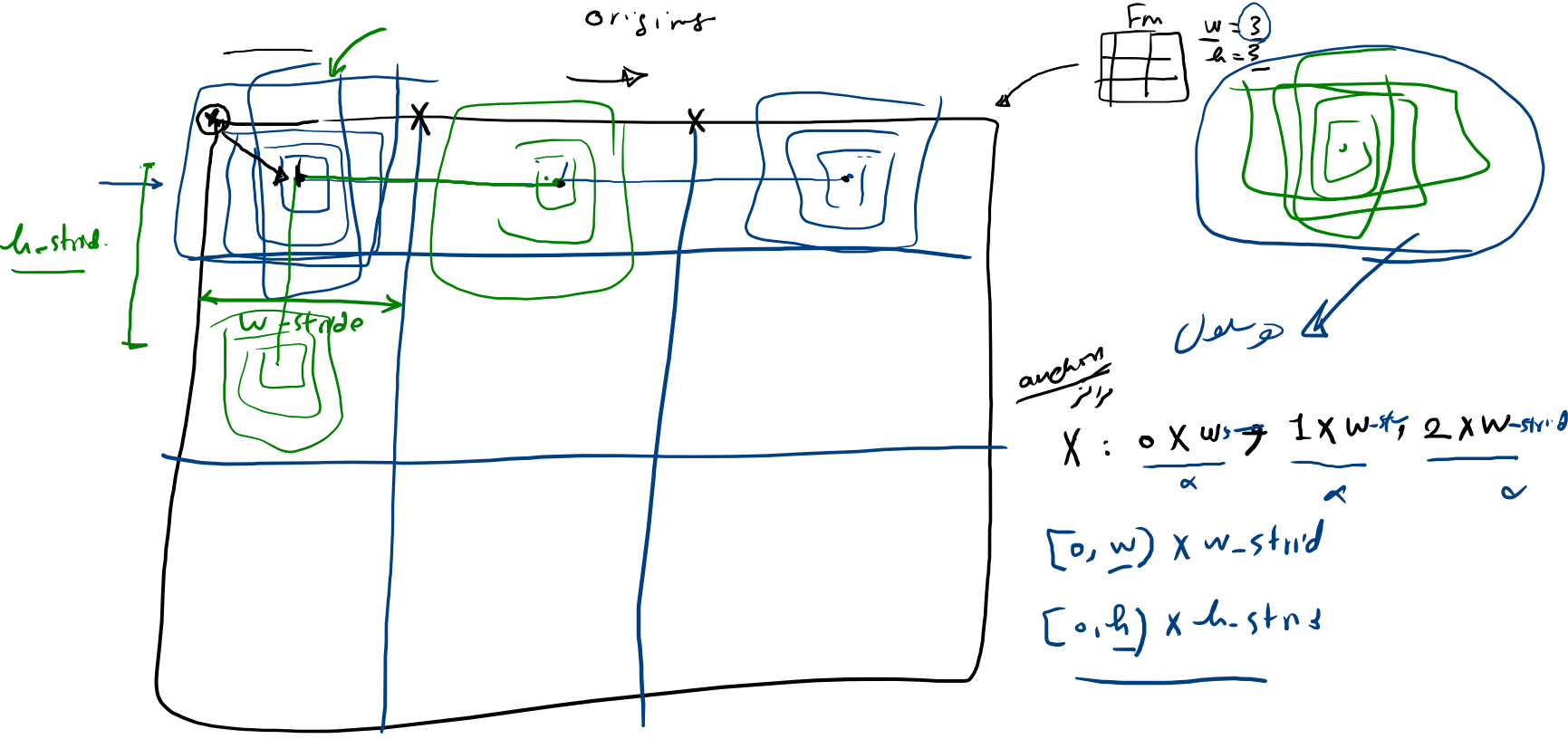
$h_s = \sqrt{4} \times .5 = 1$

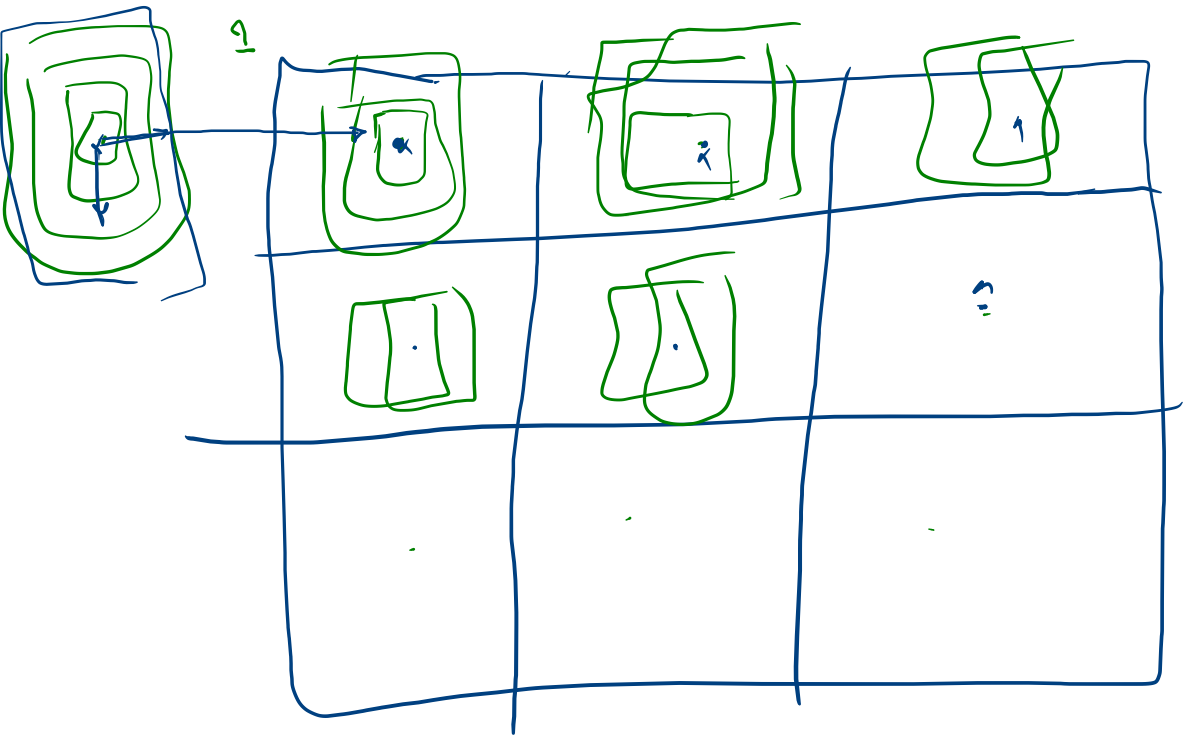
for ar = .5 → $w_s = \left(\frac{w \times h}{ar}\right)^{1/2}$

→ $h_s = w_s \times ar$

$\frac{h_s}{w_s} = \frac{1/\sqrt{2}}{\sqrt{2}} = \frac{1}{2}$




origins
Fm
w = 3
h = 3
h_stride
w_stride
anchors
X : 0 X w_s 1 X w_st 2 X w_stride
[0, w) X w_stride
[0, h) X h_stride

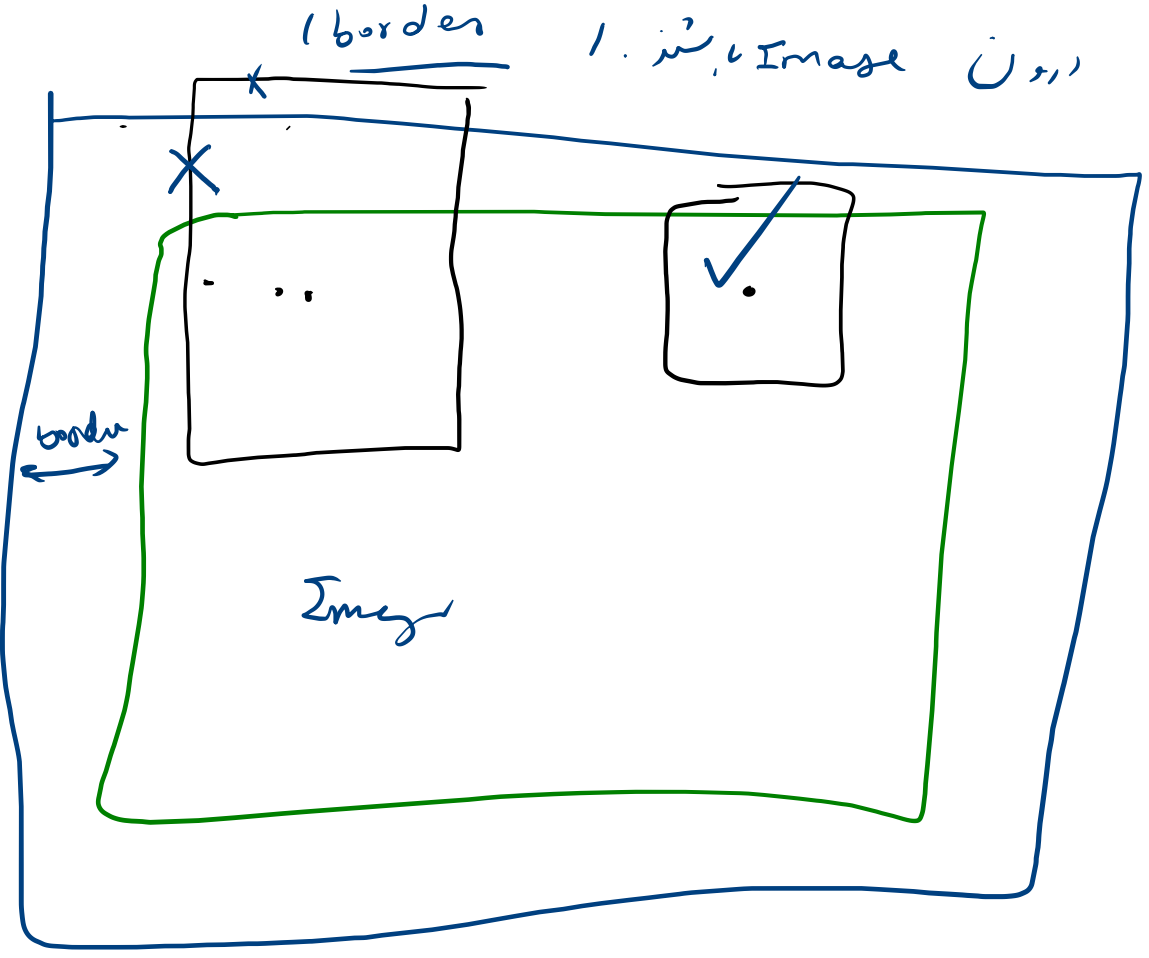


num_feature_map

9 x 9 = 81

Anchors توی شبکه: برای هر سلول‌ها! ، حالا چه anchor هایی

باید انتخاب بشوند؟ را استخراج anchor هر سلول)


border
Image
Anchor
border
Image

a. Max-IOU

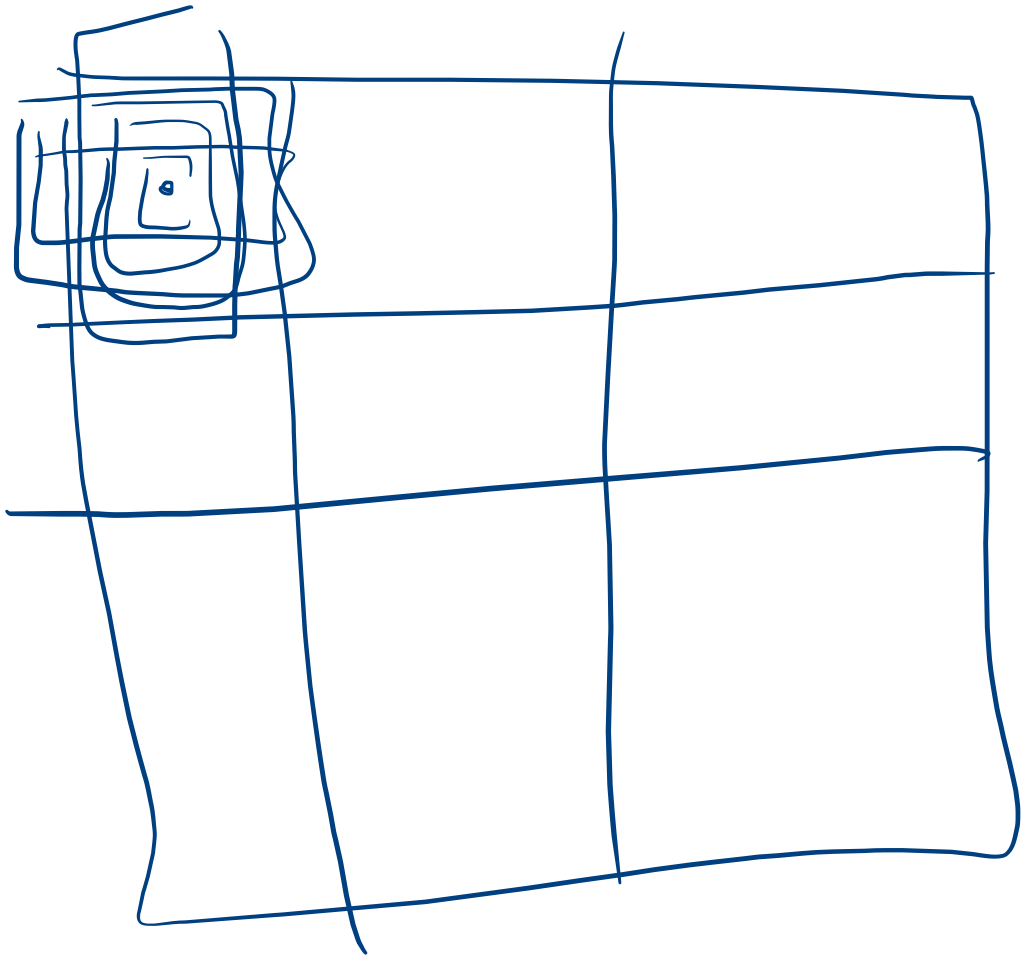
IOU وسط هر Bb را حساب کنم → Max ها را برداریم → (بیشترین تعداد anchor ها) که درون هستند

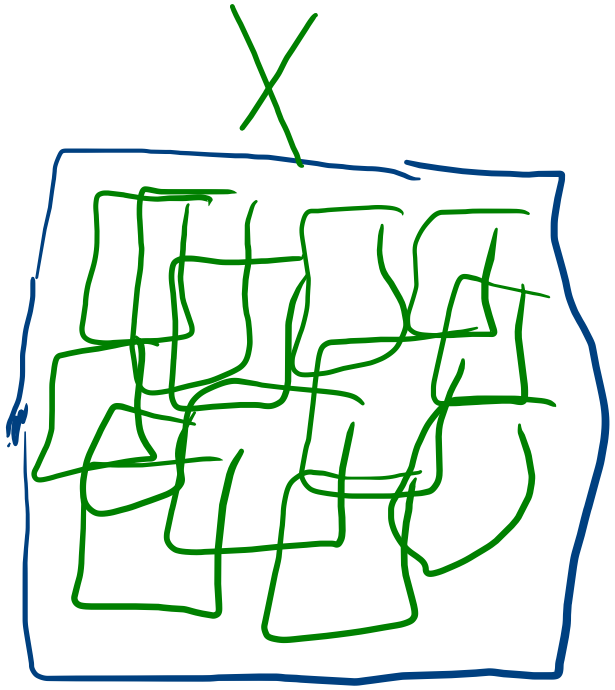
IoU > .7 → object → 1

(A) ⇒ IoU < .3 → Background → 0

$$Num\text{-}BG = \frac{Num\ fg}{2}$$

(Num-BG - A) choi's (Randomly -1)





$\bar{B}\bar{g} - \bar{F}\bar{q} - Frac.$

q