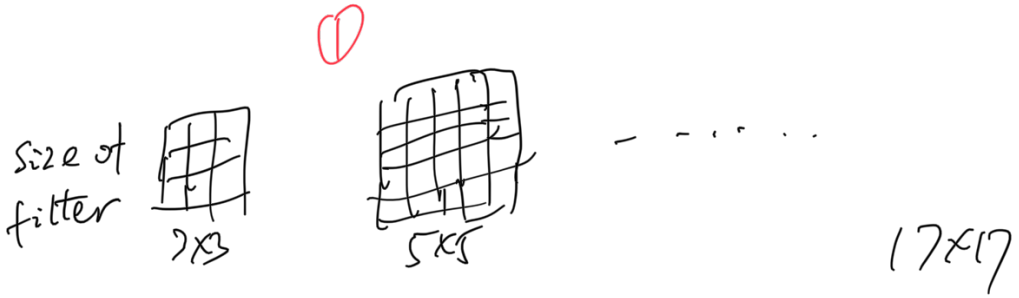
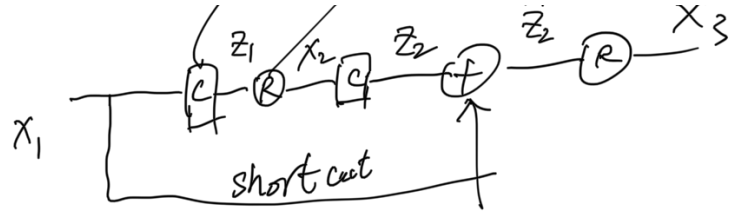
Lecture 6

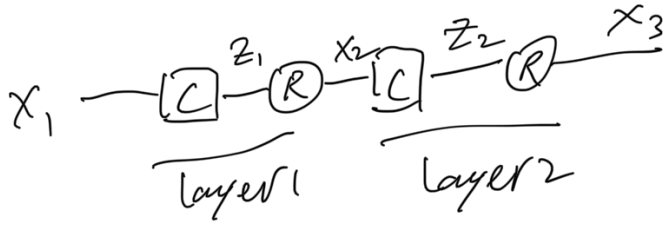
Try :

- Normalization
- Batch Normalization
- Drop out
- Augmentation
  - pixel (location) augmentation
    - affine transformation
      - shift, rotate, mirror, ...
      - range
  - color augmentation
    - brightness, contrast, saturation
- Ensemble



ResNet



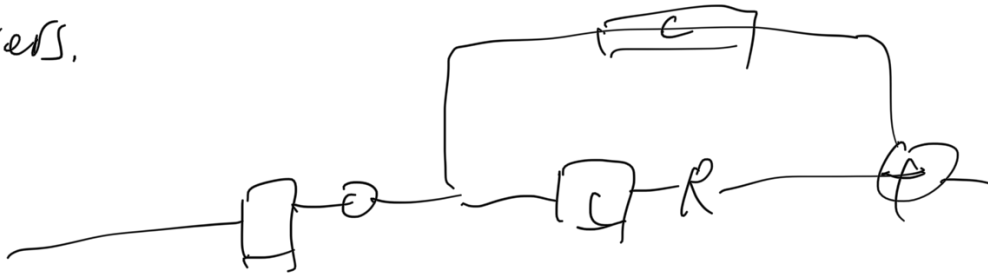


$\checkmark \hat{z}_2 = z_2 + x_1$

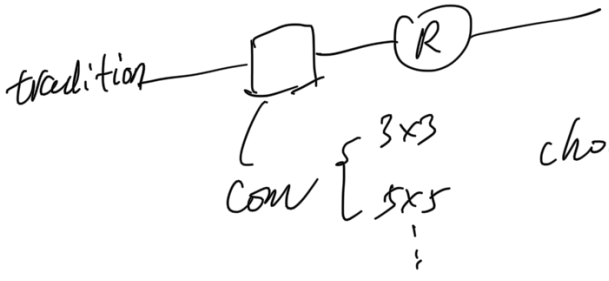
$\Rightarrow z_2 = C(x_2)$

$x_2 = \text{Relu}(z_1)$ $\quad z_1 = C(x_1)$

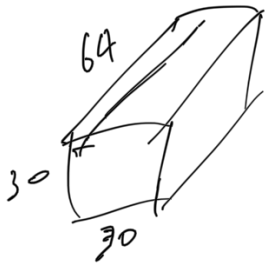
150 layers.



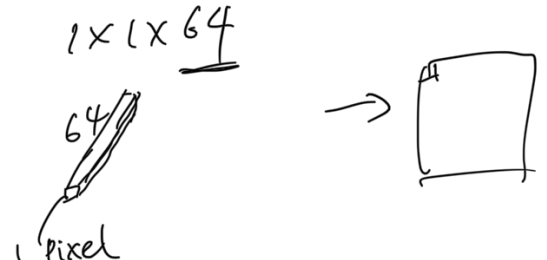
Inception Net

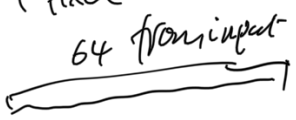


choose one for each layer



30x30x64





= 1

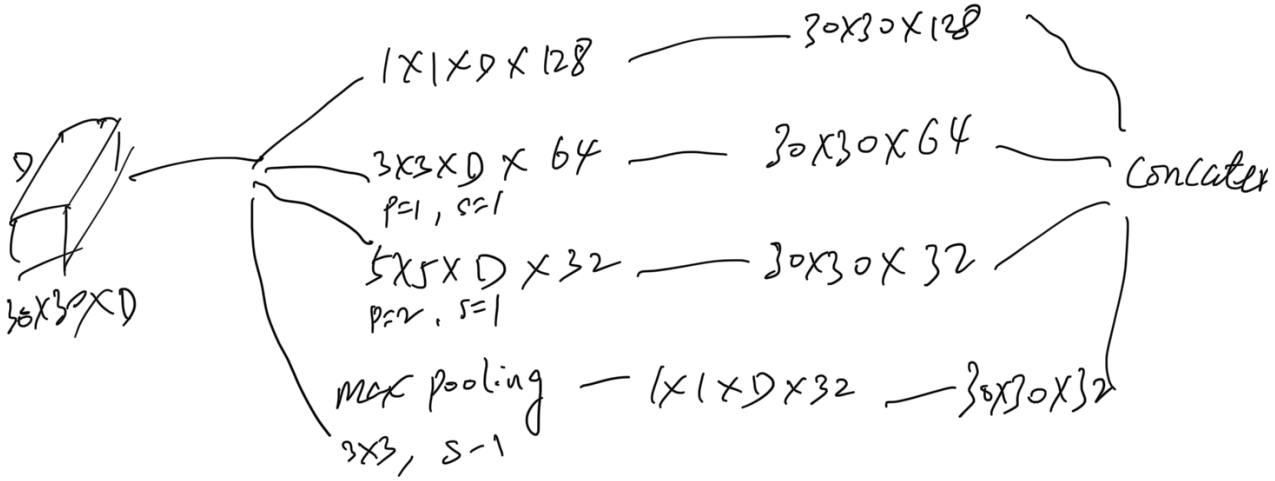
1x64

30x30x64 (*) 3x3x64 = 30x30

pad # = 1

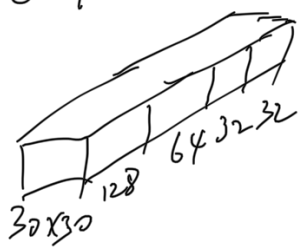
S = 1

128 of them = 30x30x128



lay has 256 filter

concatenation :



Applications:

object recognition, localization, segmentation,

human pose estimation, object pose estimation.

facial recognition.

localization:

output — bounding boxes.

$\{(c_x, c_y, w, h), s_1, s_2, \cdots s_n\}$

0.01 0.05 0.6 0.1

possible object types is n

R-CNN

2,000 bounding boxes.

1. Get 2,000 bounding boxes
2. Run ConvNet on each to recognize.

2,000 $[c_x, c_y, w, h]$

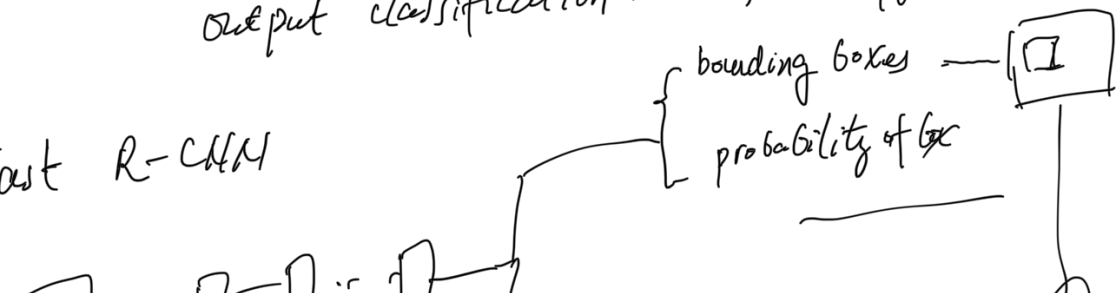
Warp, convert to a fixed size.

step 2 output:

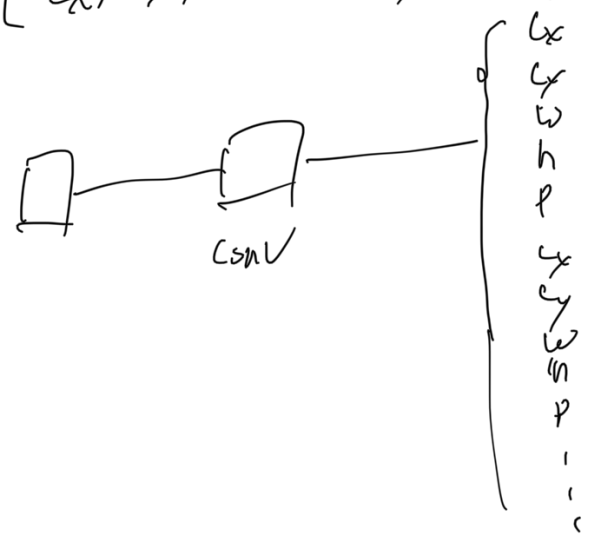
for each bounding boxes.

output classification result, & offsets.

Fast R-CNN



$\{c_x, c_y, w, h, p\}$ output ground truth bounding b



define Loss function. Based on errors in BBoxes.

Yolo :

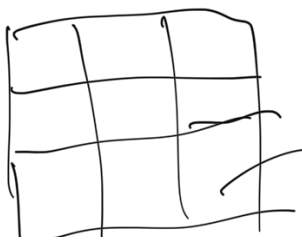
output :

A set of

$[\underbrace{c_x, c_y, w, h}_{\text{BBox}}, \underbrace{p}_{\text{pro\_B-box}}, \underbrace{s_1, s_2 \cdots s_n}_{\text{Classification softmax}}]$

Divide image into smaller patches.

3x3



each does potential BBox # such 2

run in parcellel

2 bounding boxes.

output is

$[\underbrace{c_x', c_y', w', h', p'}, \underbrace{s_1', s_2' \cdots, s_n'}, c_x^2, c_y^2, w^2, h^2, p^2$

$\underbrace{s_{v1}^2 \cdots s_n^2}]$

for box, $(5+n) \times 2$

c_x, c_y $\quad L = \sum_{} \mathbb{1}((c_x - \hat{c}_x)^2 + (c_y - \hat{c}_y)^2)$

box, w, h regression loss

confidence loss

fully overlap with ground th BBox — P=1

proposed BBox, $\hat{P} - 1$

not fully overlap with ground th BBox — P=?

IOU give you P

$$P = \frac{A \cap B}{A \cup B}$$