

## Session -2

## CNN -2

## (OVERFITTING)

March 13  
2024



Granny Smith 85.6%

iPod 0.4%

library 0.0%

pizza 0.0%

toaster 0.0%

dough 0.1%



Granny Smith 0.1%

iPod 99.7%

library 0.0%

pizza 0.0%

toaster 0.0%

dough 0.0%

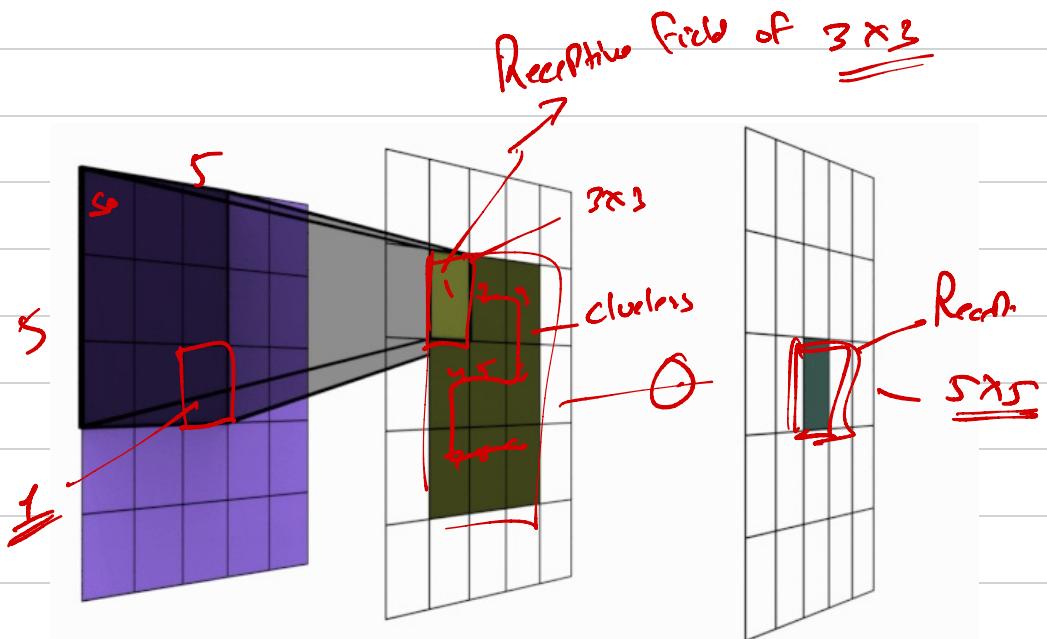
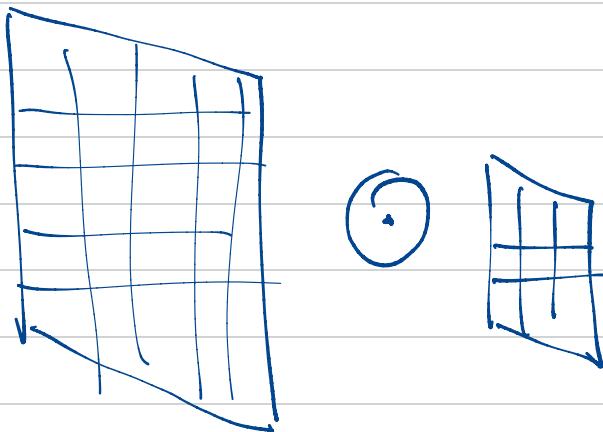
How to confuse  
machine learning:



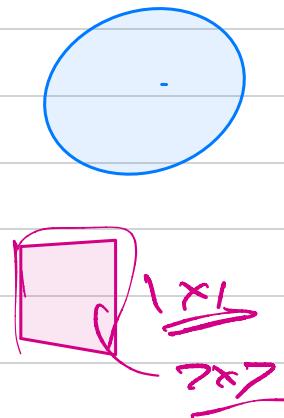
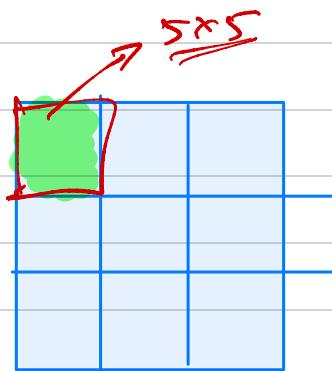
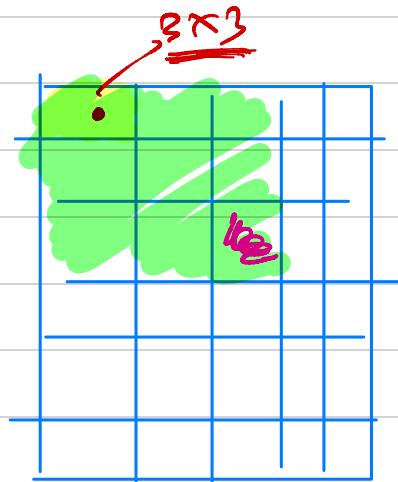
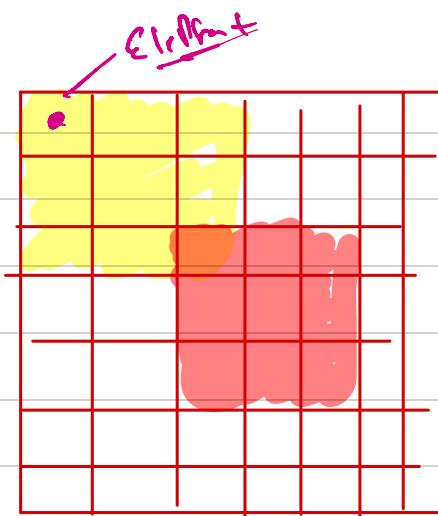
## AGENDA

- ① Receptive Field
- ② How Kernel extracts information  $\leftrightarrow$  with more than 1 channel
- ③ Why not  $5 \times 5$  or  $9 \times 9$  or even sized Kernel
- ④ Non Linearity in CNN ??
- ⑤ Strides ?? Why ?
- ⑥ GAP

# RECEPTIVE FIELD



What RF should you forget??



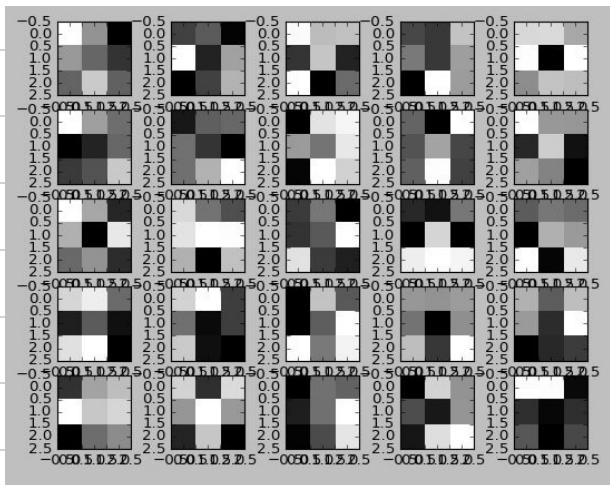
This is wrong

You should have  
at least Receptive Field  
of Entire Image

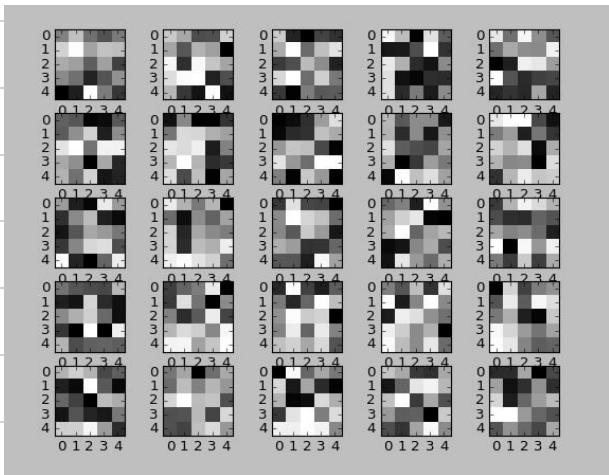
Reidity is your R-Field should  
be more than the size of image.

What kernels  
are trying to  
extract at

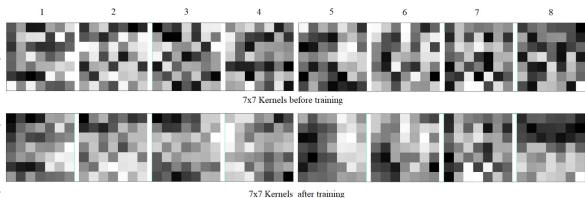
RF of 3x3



RF → 5x5

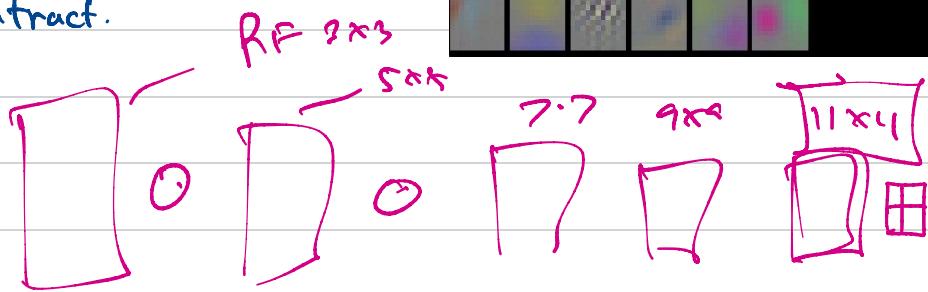


RF → 7x7



RF  $11 \times 11$

Only at this RF  
we're able to make  
sense of WHAT  
Kernel is trying to  
extract.

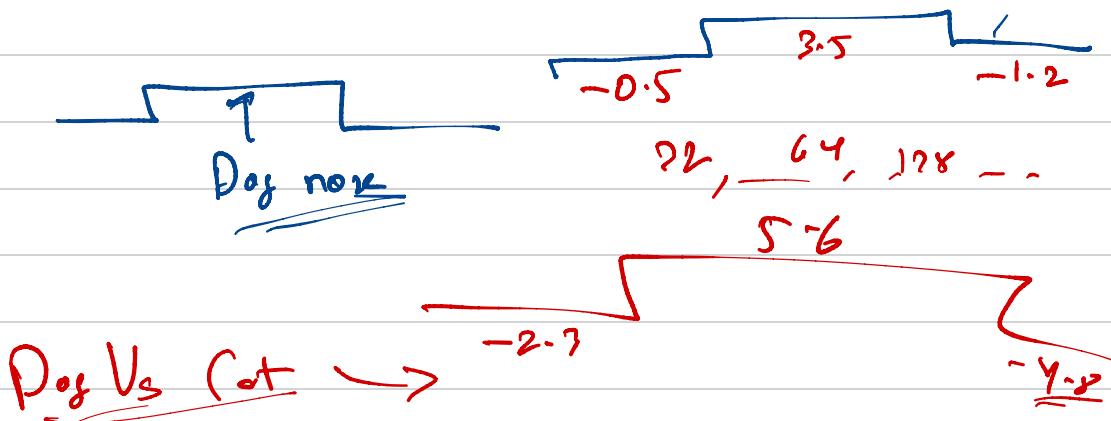


So, don't use Max-pooling for filtering  
off information before you hit  
RF of  $11 \times 11$

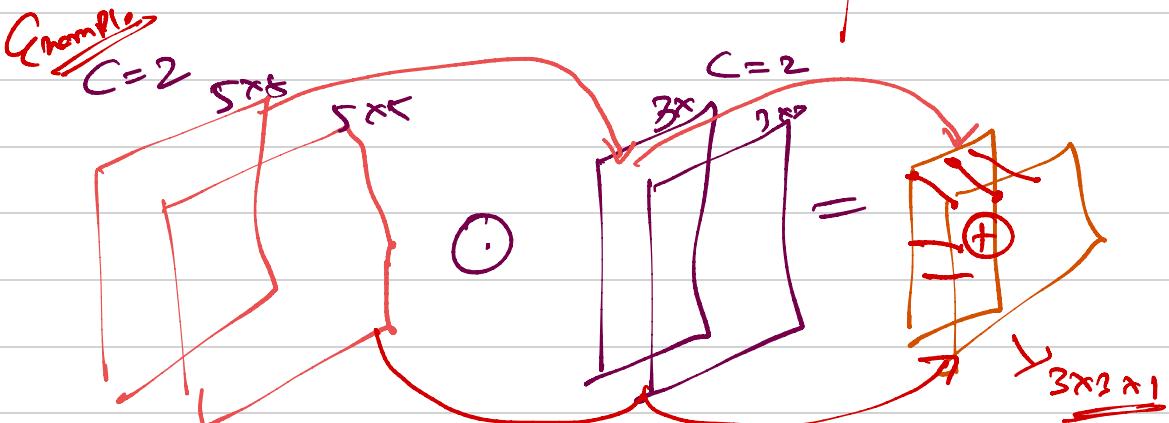
# KERNELS WITH MULTIPLE CHANNELS

$400 \times 400 \times 1$

Dog.n.  
Filter



$$\begin{array}{c}
 \text{Input: } 400 \times 400 \times 1 \\
 \text{Output: } 398 \times 398 \times 32
 \end{array}
 \quad
 \left| \begin{array}{l}
 (3 \times 3 \times 1) \times \underline{32} \\
 (3 \times 3 \times \underline{32}) \times 64 \\
 (396 \times 398 \times 1) \times \underline{64}
 \end{array} \right|
 \quad
 \begin{array}{l}
 R.F = 3 \times 3 \\
 398 \times 398 \times 32
 \end{array}$$



$$400 \times 400 \times 1$$

$$398 \times 398 \times 32$$

$$396 \times 396 \times 64$$

$$(3 \times 3 \times 1) \times 32$$

$$(3 \times 3 \times 32) \times 64$$

$$(3 \times 3 \times 64) \times 128$$

$$R.F = 3 \times 3$$

$$398 \times 398 \times 32$$

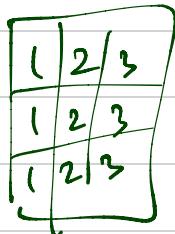
$$(396 \times 398 \times 1) \times 64$$

$$R.F = 3 \times 5$$

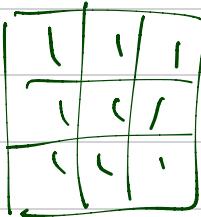
$$394 \times 394 \times 128$$

$$394 \times 394 \times 64$$

$$394 \times 394 \times 1$$

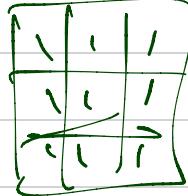
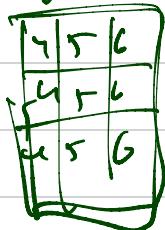


①



=

$$18$$



$$45$$

$$\begin{matrix} 18 \\ 45 \end{matrix}$$

$$= 63$$

$$\begin{matrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{matrix} \times 3$$

②

$$\begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$$

-

$$\begin{matrix} 9 & 9 \\ 9 & 9 \end{matrix}$$

$$\begin{matrix} 9 & 9 \\ 9 & 9 \end{matrix}$$

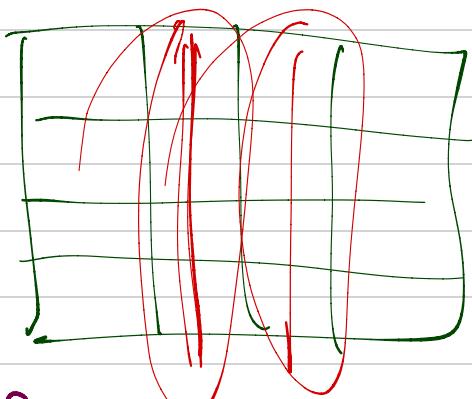
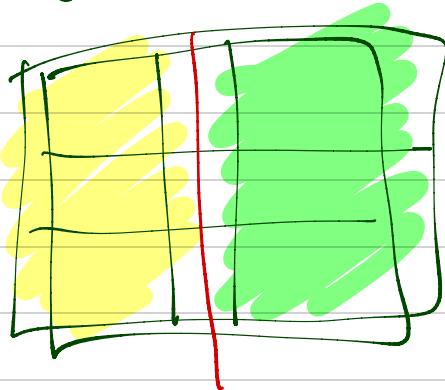
+

$$\begin{matrix} 0 & 2 \\ 2 & 0 \end{matrix}$$

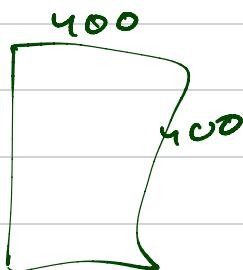
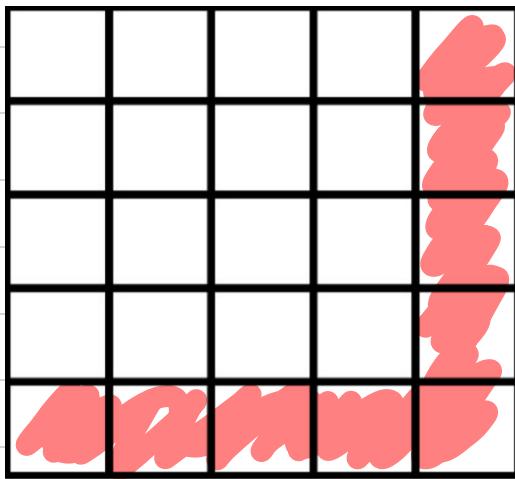
RGB

$$= \begin{matrix} 27 & 27 \\ 27 & 27 \end{matrix}$$

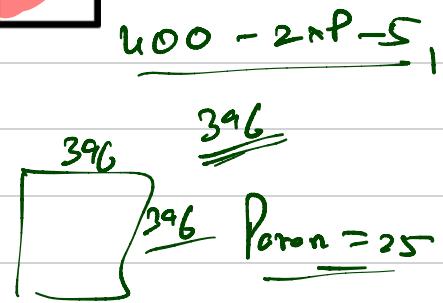
Why not even size Kernel



Lack of units of symmetry



$$5 \times 5 = 25$$



$$400 \quad 9$$

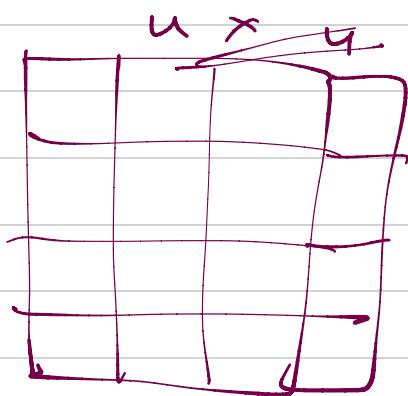
Diagram showing a large rectangle divided into four smaller rectangles. The top-left rectangle has a green border and contains the number 400. The bottom-right rectangle has a green border and contains the number 398. The middle row has a green border and contains the numbers 1 and 2. The middle column has a green border and contains the numbers 1 and 2. A green bracket on the left side groups the top two rectangles, and another green bracket on the right side groups the bottom two rectangles. A green circle contains the number 1.

$$= \boxed{398}$$

~~Selfand~~

Diagram showing a large rectangle divided into four smaller rectangles. The top-left rectangle has a green border and contains the number 1. The bottom-right rectangle has a green border and contains the number 2. The middle row has a green border and contains the numbers 1 and 2. The middle column has a green border and contains the numbers 1 and 2. A green circle contains the number 18.

$$\boxed{396} \quad 2$$

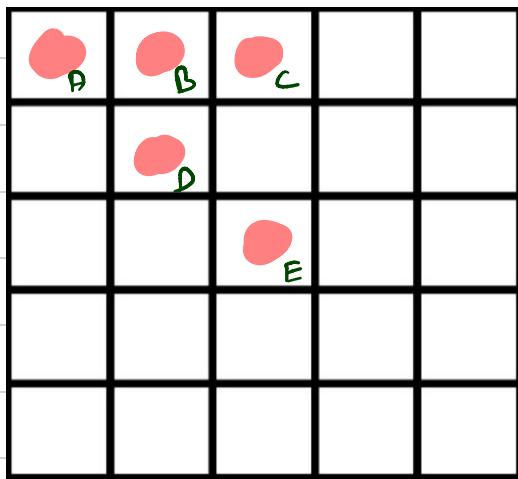


$$\frac{4 - 2 \times 0 - 3}{2} + 1$$

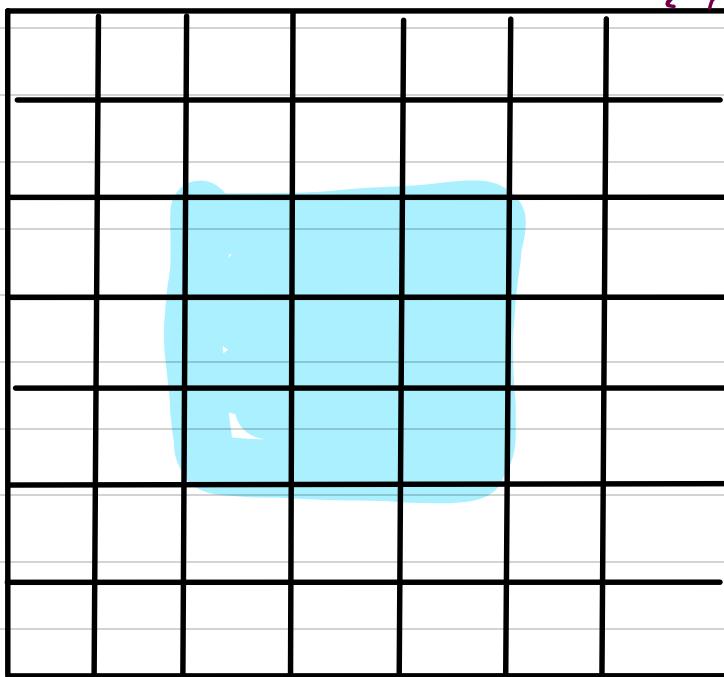
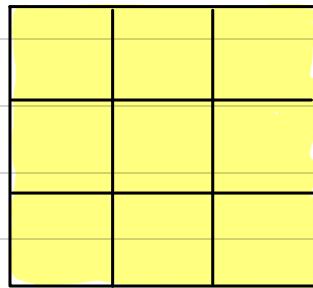
$\frac{1}{2} + 1$

Diagram showing a large rectangle divided into four smaller rectangles. The top-left rectangle has a green border and contains the number 1. The bottom-right rectangle has a green border and contains the number 2. The middle row has a green border and contains the numbers 1 and 2. The middle column has a green border and contains the numbers 1 and 2. A green circle contains the number 1.

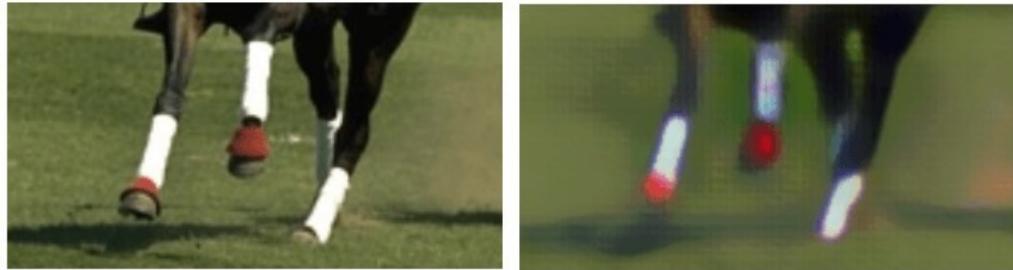
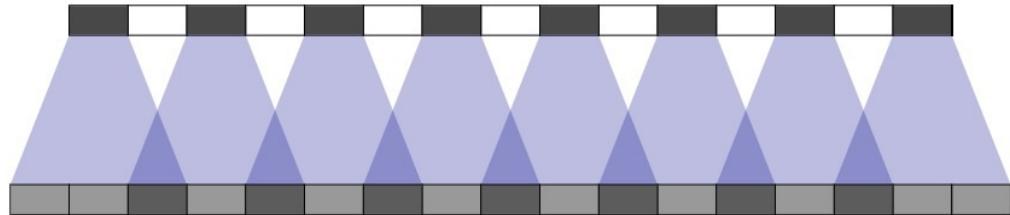
# STRIDES



$\alpha = 1$ ,  $\beta = 2$   
 $c = 3$   
 $d = 4$   
e = 5



$7 \times 7$   
~~400 x 400~~  
 $396 \times 396$   
 9 time



Perceptual Losses for Real-Time Style Transfer and Super-Resolution



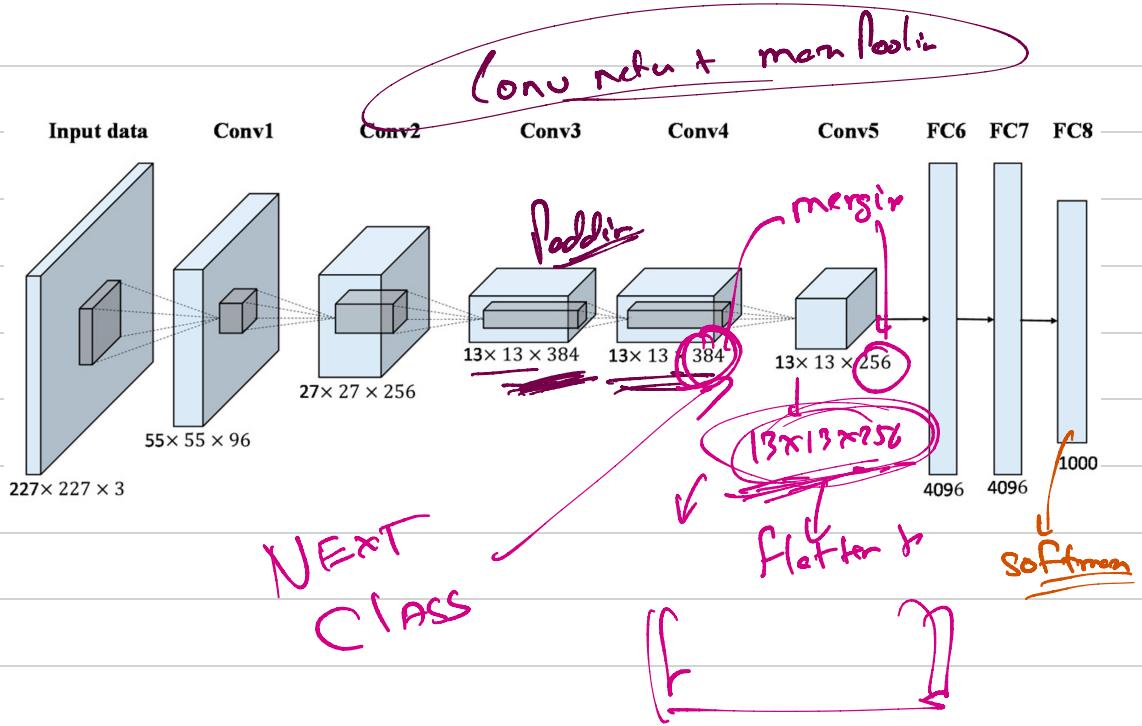
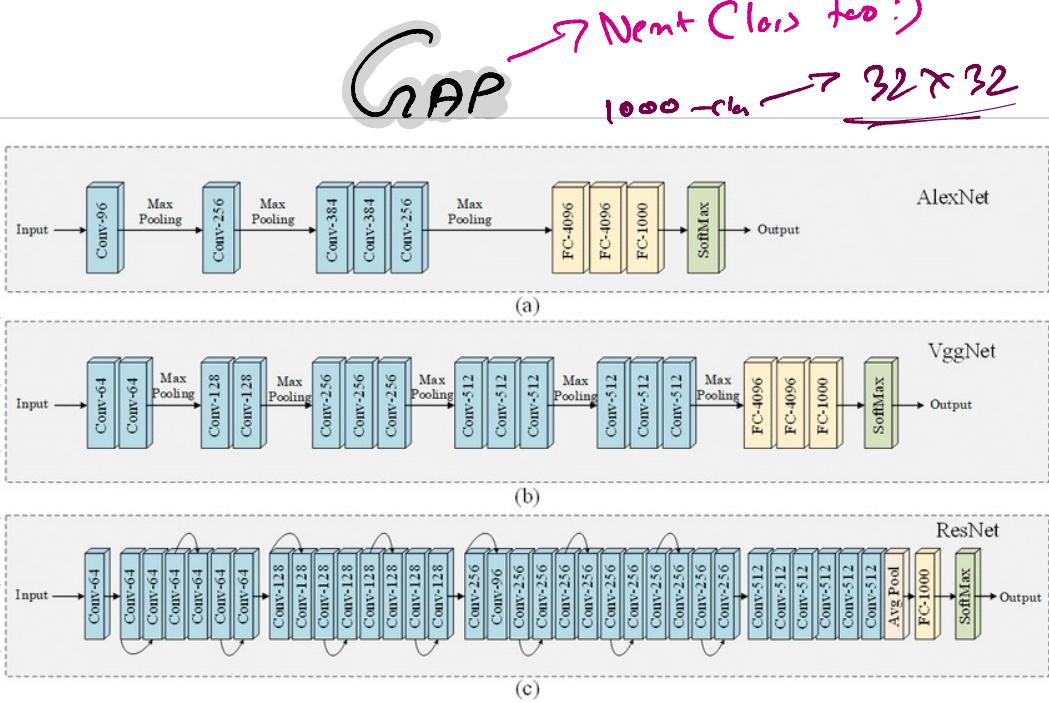
## Benefit ??

①

Inference → Faster

②

Model → easily run on Edge device



| Layer Name  | Tensor Size | Weights    | Biases | Parameters |               |
|-------------|-------------|------------|--------|------------|---------------|
| Input Image | 227x227x3   | 0          | 0      | 0          |               |
| Conv-1      | 55x55x96    | 34,848     | 96     | 34,944     |               |
| MaxPool-1   | 27x27x96    | 0          | 0      | 0          |               |
| Conv-2      | 27x27x256   | 614,400    | 256    | 614,656    |               |
| MaxPool-2   | 13x13x256   | 0          | 0      | 0          |               |
| Conv-3      | 13x13x384   | 884,736    | 384    | 885,120    |               |
| Conv-4      | 13x13x384   | 1,327,104  | 384    | 1,327,488  |               |
| Conv-5      | 13x13x256   | 884,736    | 256    | 884,992    |               |
| MaxPool-3   | 6x6x256     | 0          | 0      | 0          |               |
| FC-1        | 4096x1      | 37,748,736 | 4,096  | 37,752,832 | explaining 21 |
| FC-2        | 4096x1      | 16,777,216 | 4,096  | 16,781,312 |               |
| FC-3        | 1000x1      | 4,096,000  | 1,000  | 4,097,000  |               |
| Output      | 1000x1      | 0          | 0      | 0          | GRAP          |
| Total       |             |            |        | 62,378,344 |               |

(Layr. lagen. Den (4096, activation = ReLU))