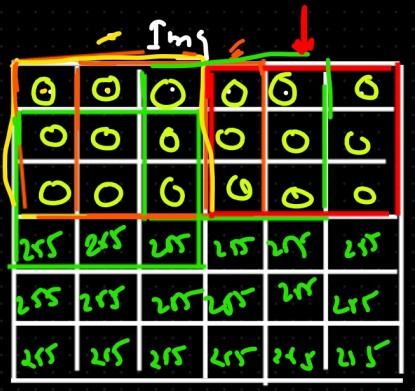
CNN

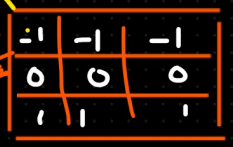
1. Image
2. Filter | Kernel
3. Convolution + Relu
4. Pooling
5. Flatten (2D - 1D)
6. FC

256 = [0-255]

Image → Gray Scale, Color img (3 - Channel) ⇒ R, G, B



6 × 6

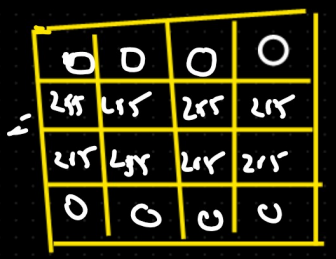


3 × 3

Filter | Feature | Kernel

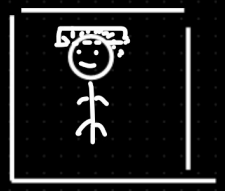
$(0 \times -1) + 0 \times (-1) + 0 \times (-1) + 0 \times 0 + \dots = 0$

=



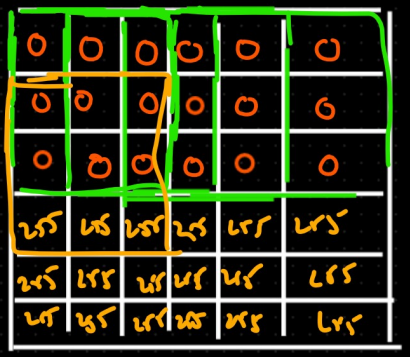
Feature map

Convolution ⇒ ① extract the feature from img.

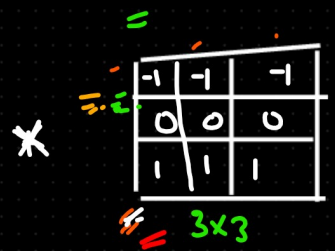


Feature extraction

- Odd.
  - Alexnet
  - VGG16, 19
  - Resnet
  - Inception



6×6



3×3

horizontal edge Detection

feature map

4×4

+ 0x1

LENET

Kernel / filter ⇒ Dimension 1×1, 2×2, 3×3, 4×4, 5×5, 6×6, 7×7

Initially ⇒ (1980-1990)

Sensor ⇒ Image

Hardcoded

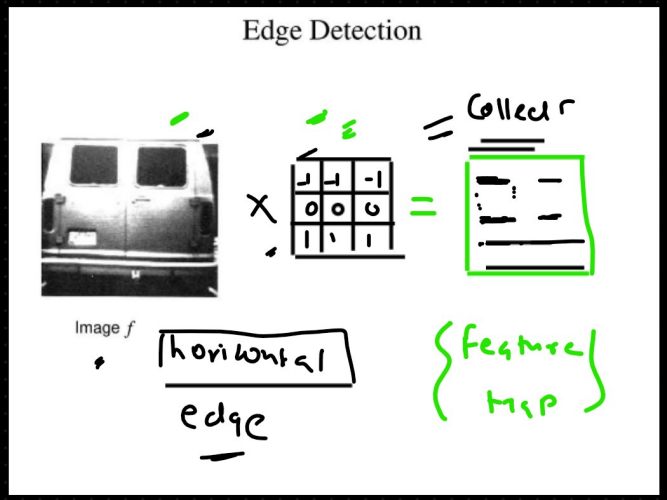
machine

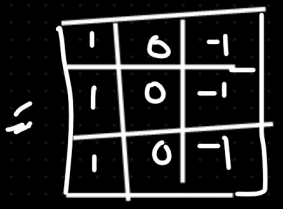
algorithm

Feature

3×3 ⇒ 9

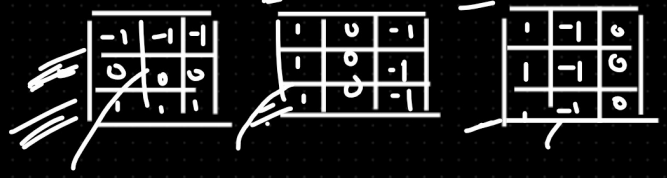
3×3 ⇒ 5×5 ⇒ 7×7 ⇒ 9×9 ⇒ 11×11



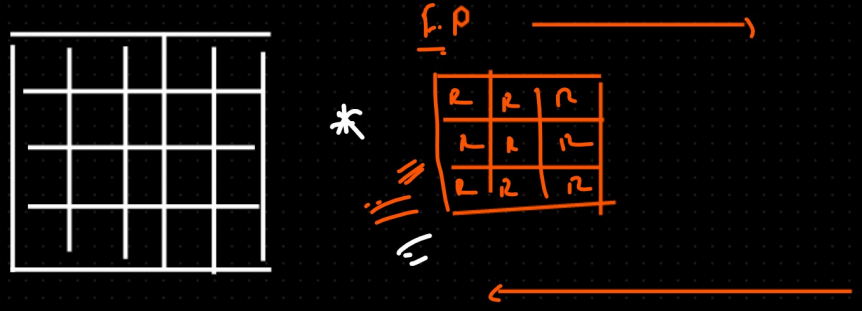


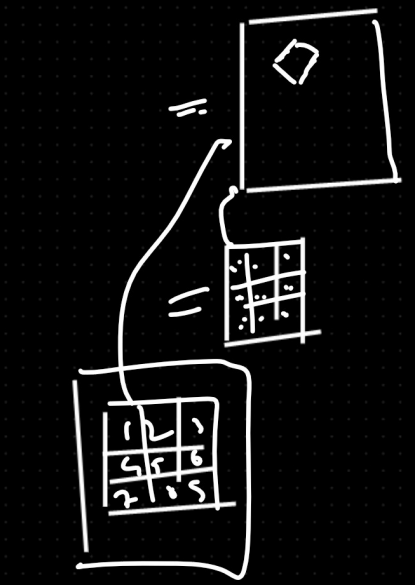
= Vertical edge detection

Filter ⇒ For extracting the feature from the image



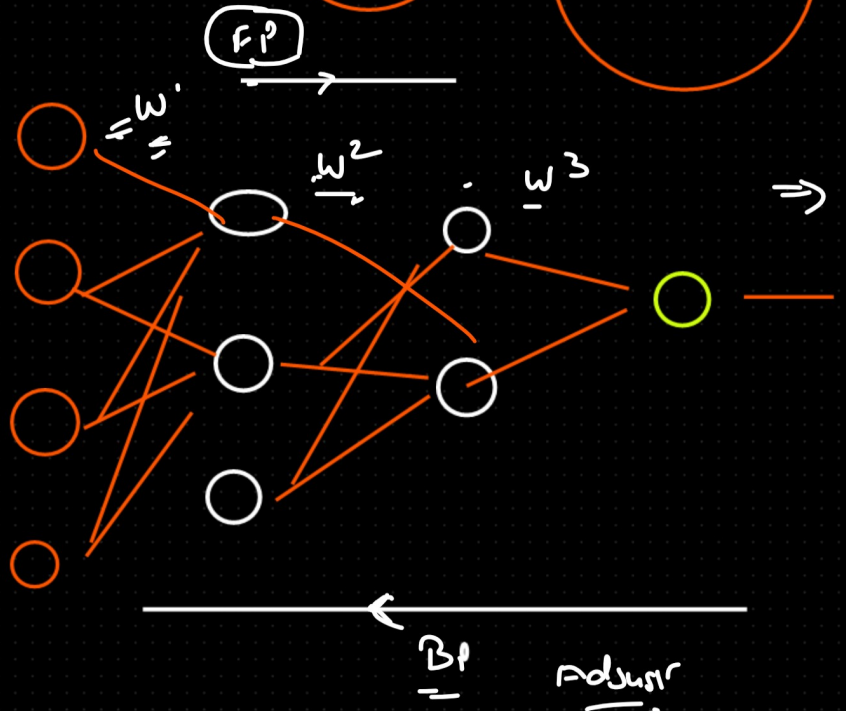
CNN = E-NET ⇒ front Architecture in CNN





Learning

1990 :: Le-Net → CNN ⇒ Learnable



⇒ Random value

NN ⇒ it is learning by itself

it is learning based on data

BP..

$\frac{\partial L}{\partial w}$

$6\times6 * 3\times3 = 4\times4$

$8\times8 * 3\times3 = 6\times6$

$28\times28 * 3\times3 = 26\times26$

$n\times n * m\times m \Rightarrow (n-m+1) \times (n-m+1)$

$6\times6 * 3\times3 = (6-3+1) \times (6-3+1)$

$= (3+1) \times (3+1)$

$= 4\times4$

img_size - filter + 1

64 x 64 * 5 x 5 = 60 x 60

(64 - 5 + 1)

(59 + 1) = 60

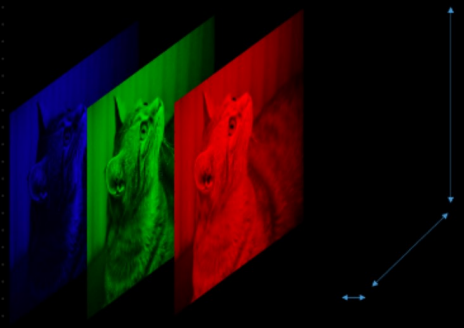
64 x 64 = 60 x 60

Padding

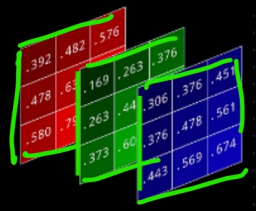
Stride x Padding

= Every feature

low level feature

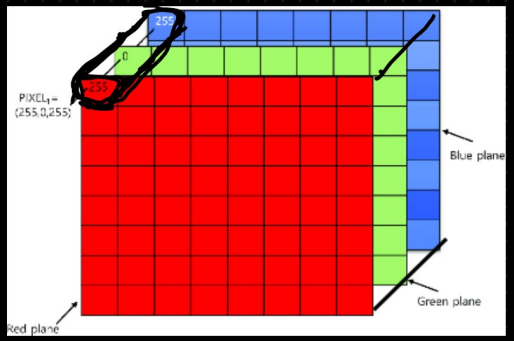


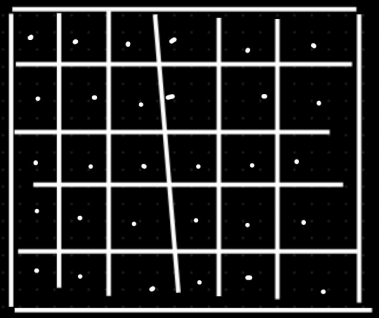
Color (R,G,B)

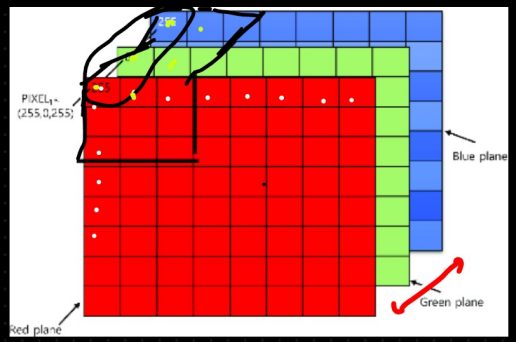


Color → R, G, B

gray scale image



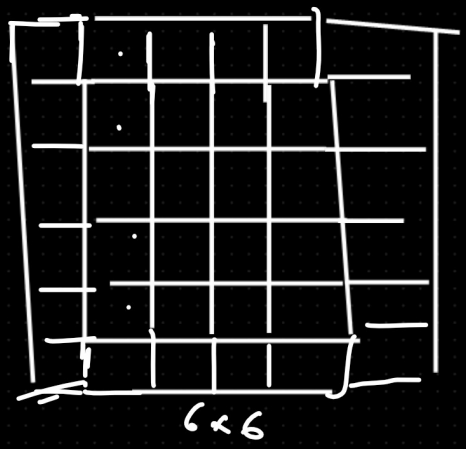




*

(3×3×3)

=



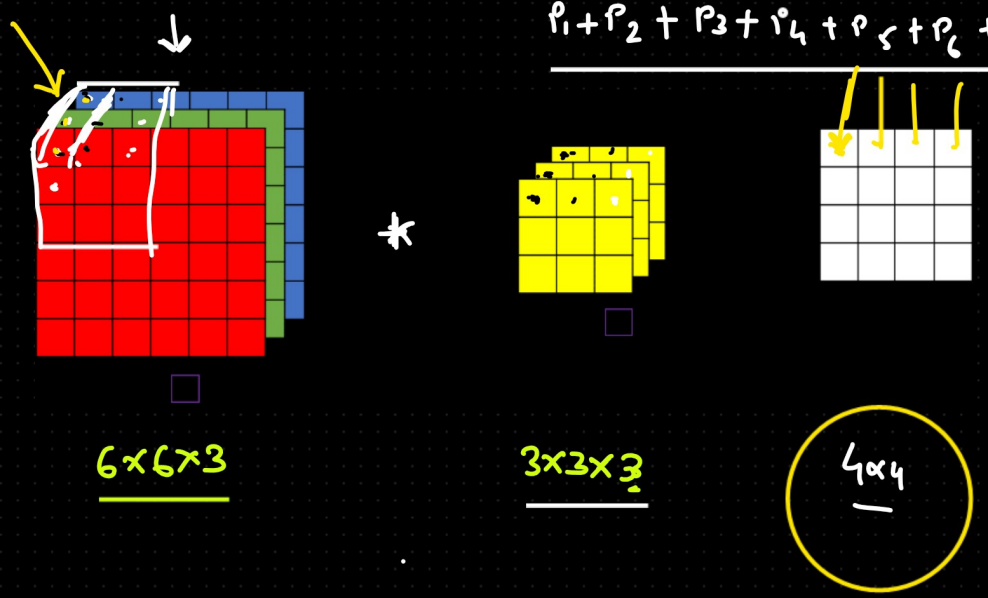
(8×8×3)

Color

(n-m+1)

F.M

Single channel



Practical

image x filter

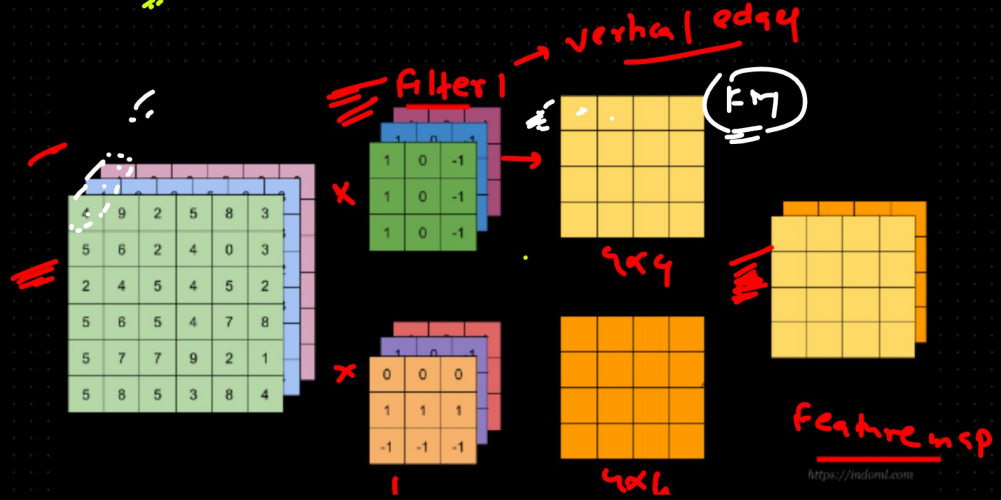


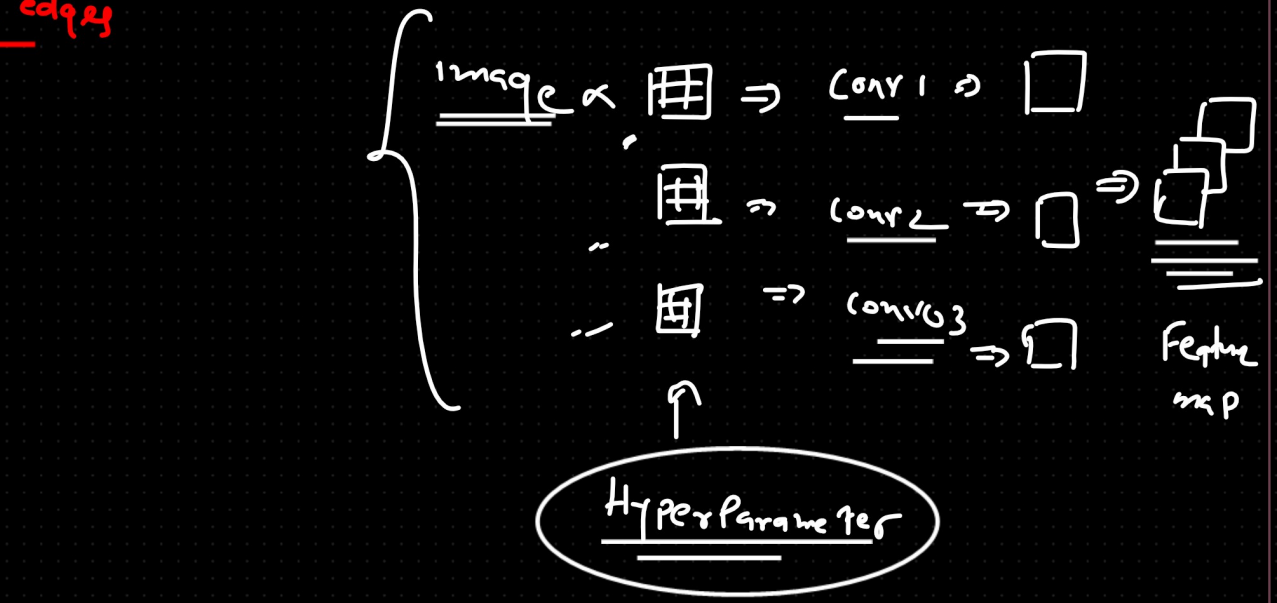
Image x filters

↓

Extract information from img

Image ⇒ side edge
horizontal edge
circle

↘ filter 2 Horizontal edges



Padding | Stride | Pooling

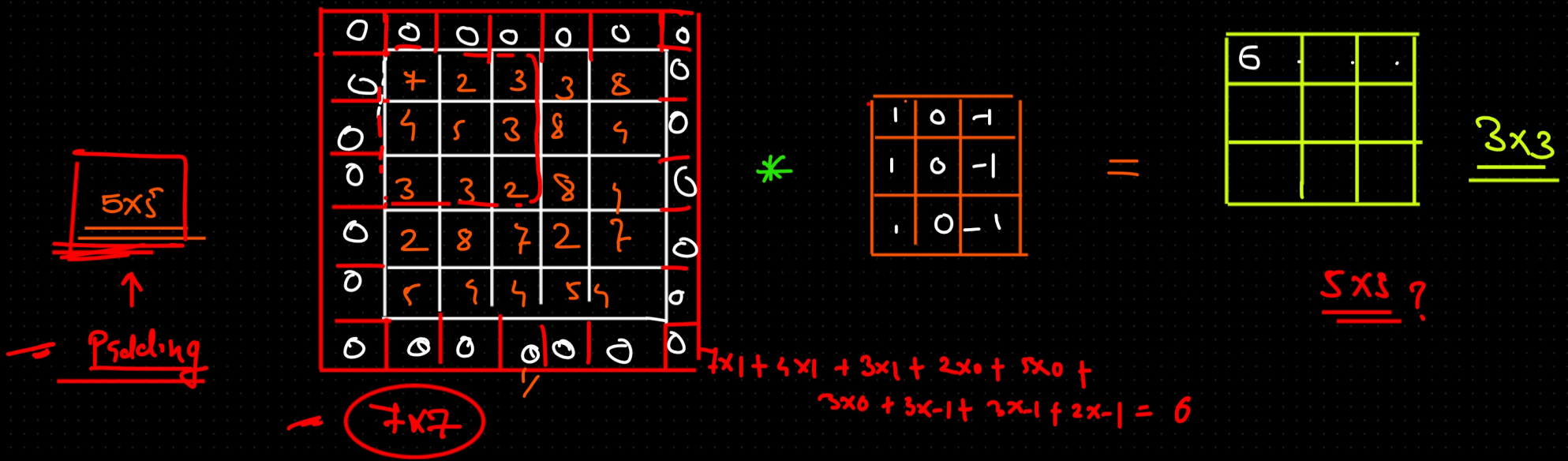
Padding

image x filter ⇒ convolving

6x6 3x3 ⇒ 4x4

n-p+1 ⇒ 6-3+1

3+1 = 4



$n \times n$

$\Rightarrow$ 5x5

$f \times f$

3x3

$= n-f+1$

$5-3+1 =$ 3x3

5x5 $\Rightarrow$ 5x5 ?

$n-f+1 = 5$

$n-f+1$

$7-3+1$

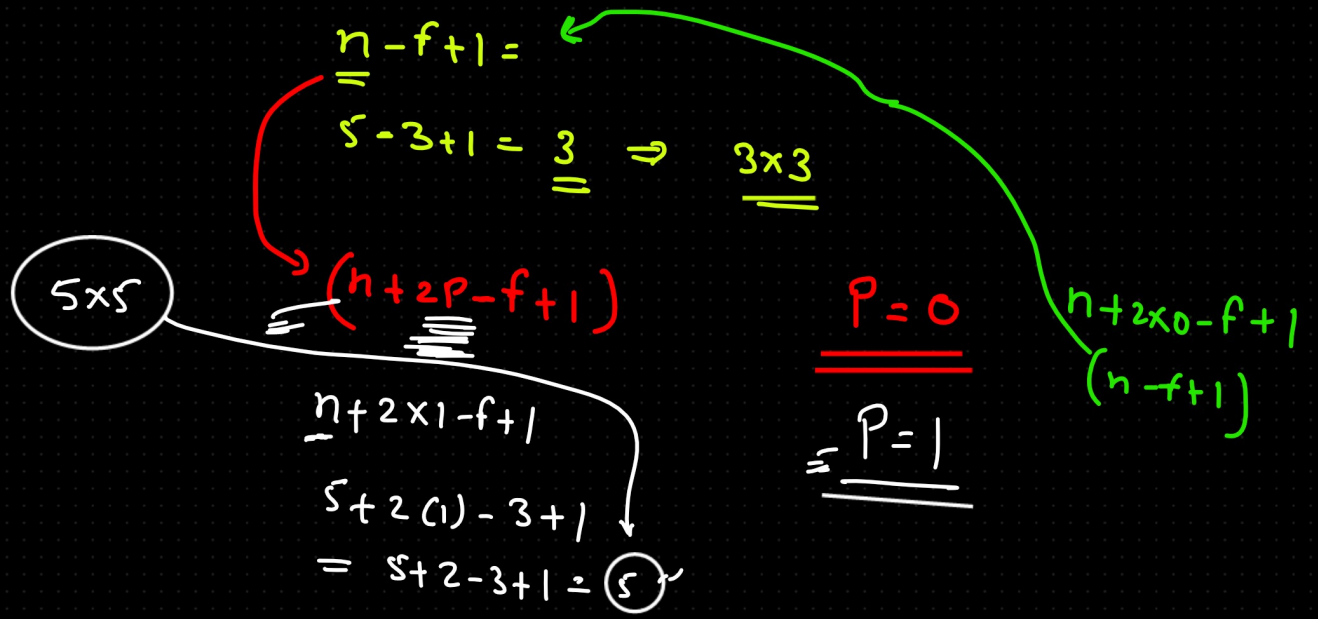
$4+1 =$ 5

$n = 5-1+f$

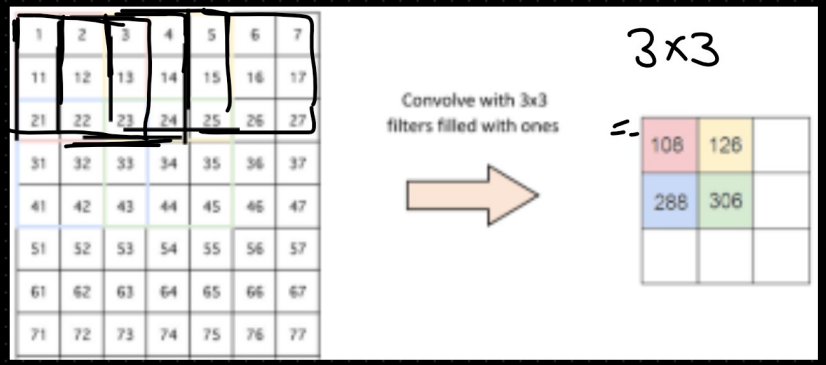
$= 5-1+3$

$n = 7$

$\Rightarrow$ $n=7$ $\Rightarrow$ 5



Stride ⇒ Jump is called Stride



1 Jump ⇒ Stride

While we are iterating over the img

Stride = 1

$\frac{n-f+1}{}$ ⇒ $\left(\frac{n-f+1}{s}\right)$ $s=1$ ⇒ $\left(\frac{n-f+1}{1}\right)$ $n-f+1$

$\frac{n+2p-f+1}{}$ ⇒ $\left(\frac{n+2p-f}{s}+1\right]$ ⇒ $s=1$ ⇒ $(n+2p-f+1)$

6x6 3x3 ⇒ $\left(\frac{6-3+1}{1}\right)$ ⇒ $6-3+1=4$

4x4

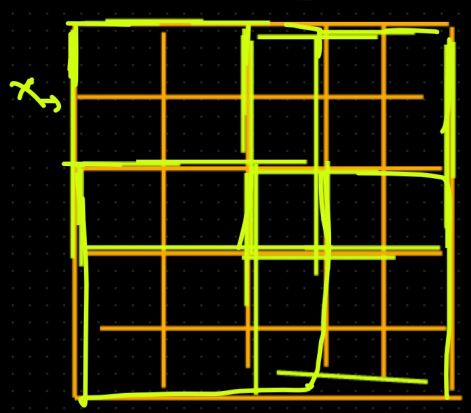
6×6 $\quad 3 \times 3 \Rightarrow \quad \frac{6 + 2\times 0 - 3}{2} + 1$

$\begin{cases} P = 0 \\ S = 2 \end{cases} \Rightarrow \frac{6+(-3)}{2} + 1 \Rightarrow \frac{6-3}{2} + 1 = \frac{3}{2} + 1$

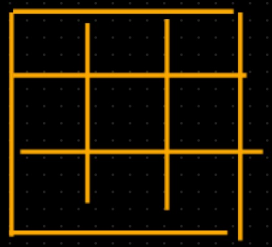
$= 1.5 + 1 = 2.5 \rightarrow 2\times 2$

$6 \times 6 \rightarrow 2 \times 2$

$S = 2$



$*$



$\Rightarrow 2 \times 2$

$5 \times 5 \quad 3 \times 3 \Rightarrow$

$\left(\frac{n+2P-f+1}{S}\right) \Rightarrow \left(\frac{5+2\times 0-3}{2}\right)+1 \Rightarrow \left(\frac{5-3}{2}\right)+1 = 2$

1. High level feature

2. Computation

64 x 64 x 3

↓

low level

$$\left[\frac{n+2p-f}{s}+1\right] \Rightarrow \text{Resultant image}$$

n = Size of image

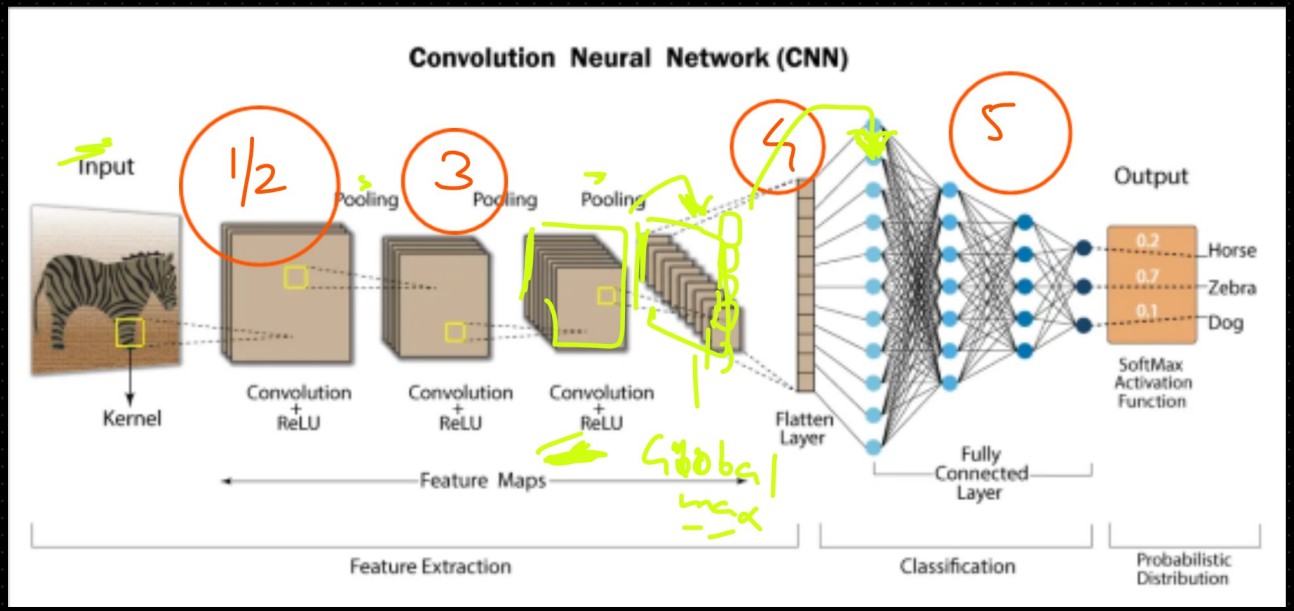
p = Padding

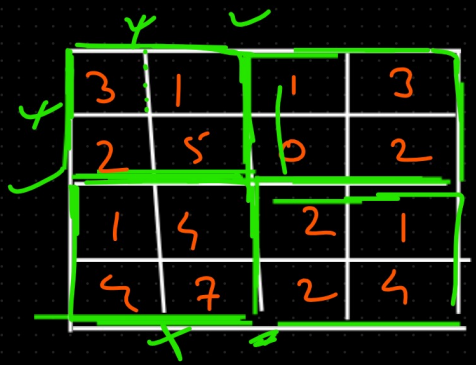
f = filter size

s = stride value

s = 2, 3, 4, ...

1. Convolution
2. Relu (to remove -ve value)
3. Pooling
4. Flatten
5. FC





Feature map

Convoluted image

⇒ Pooling (i want to Pool the imp feature)

1. max Pooling
2. min Pooling
3. averge Pooling
4. Global Pooling

Size = 2x2
Stride = 2
type = max

image ∝ filter

5	3
7	4

2∝2

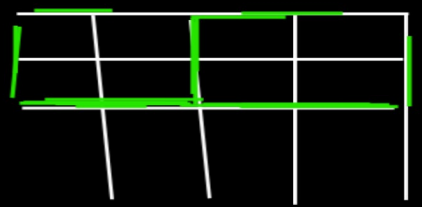
specific feature

intensity of image

Huge

$228 \times 228 \times 3 \;*\; \frac{3 \times 3 \times 3}{(100)} = 226 \times 226 \times 100 \Rightarrow (2 \times 2) \rightarrow 2 \Rightarrow 113 \times 113 \times 100$

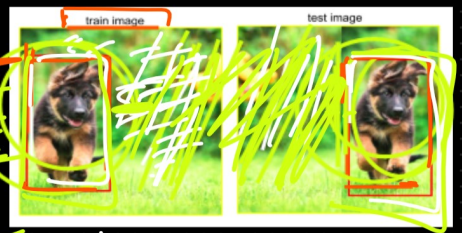
1. reduce the size
2. computational cost
3. focus on the specific feature



4

translation invariance

Pooling



filter

training

fc

filter

Pooling

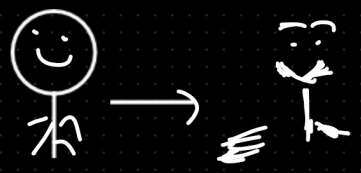
Specific

Max pool, avg

low intensity

high intensity

Enhanced feature



Global Pooling

Global max

Global avg Pooling

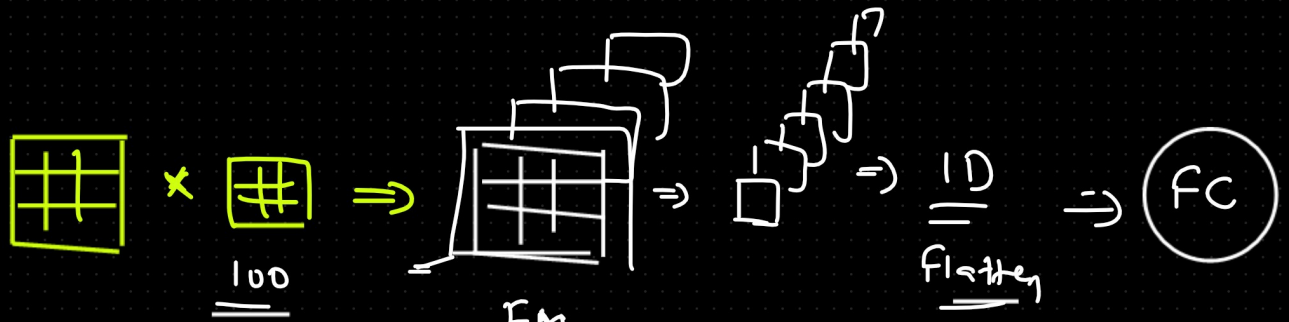
3	1	1	3	3
2	5	0	2	5
1	4	2	1	5
4	7	2	4	6

Convolving

2x2 , 2 , max, min, avg

7

Scalar value



Global max pooling

low intentence

↓

bmm