

Convolution neural network 2

17 October 2023 07:03

Image \rightarrow filter \rightarrow Feature map

6x6 3x3 4x4

$$n - f + 1 = FM$$

Image \rightarrow Filter = 1 feature map.
Image \rightarrow 10 filters = 10 feature maps.

Image \rightarrow filter \rightarrow feature map



padding

padding is nothing but adding image a layer in order to avoid shrinkage.

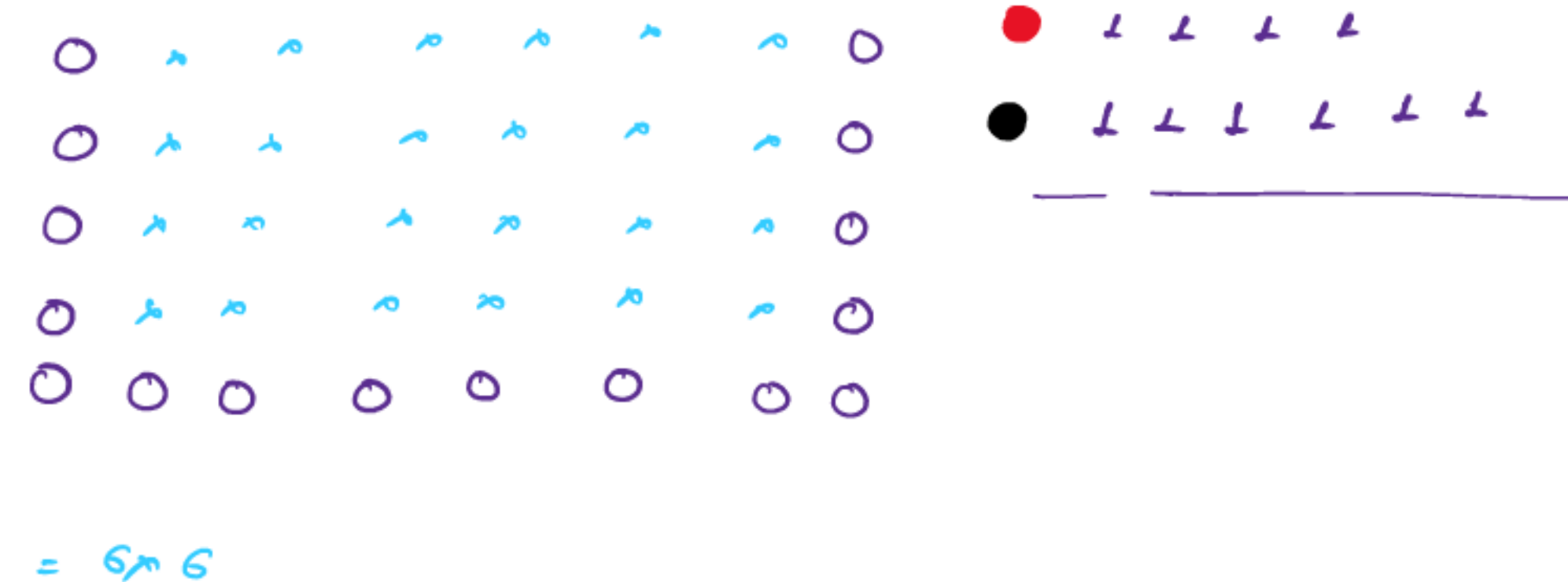


Image = 6x6
padding Image 8x8



Red 1
Black 1 1 1 1

$$n + 2p - f + 1$$

$$8 + 2 \times 1 - 3 + 1 = FM$$

$$= \frac{8 \times 8}{64}$$

$$8 - 3 + 1 = 6 \times 6$$

$$\frac{36}{20}$$

It is simple process of adding layer.



this increases the contribution of pixels which are present at borders.



Stride

generally speaking
larger steps you take
when sliding filter on your
image more details you will
capture in feature map.



Stride 1 feature capture

Image \rightarrow filter \rightarrow FM

add.
1x1 1 1

valid padding
same padding.

valid padding implies no padding.

same padding implies adding a
layer of padding so that output
size of image remains the same.

same

valid

Image 6x6 + 1 = 8x8

filter 3x3

FM

Image 6x6

filter 3x3

FM

$$n + 2p - f + 1$$

$$6 + 2 \times 1 - 3 + 1$$

$$6 \times 6$$

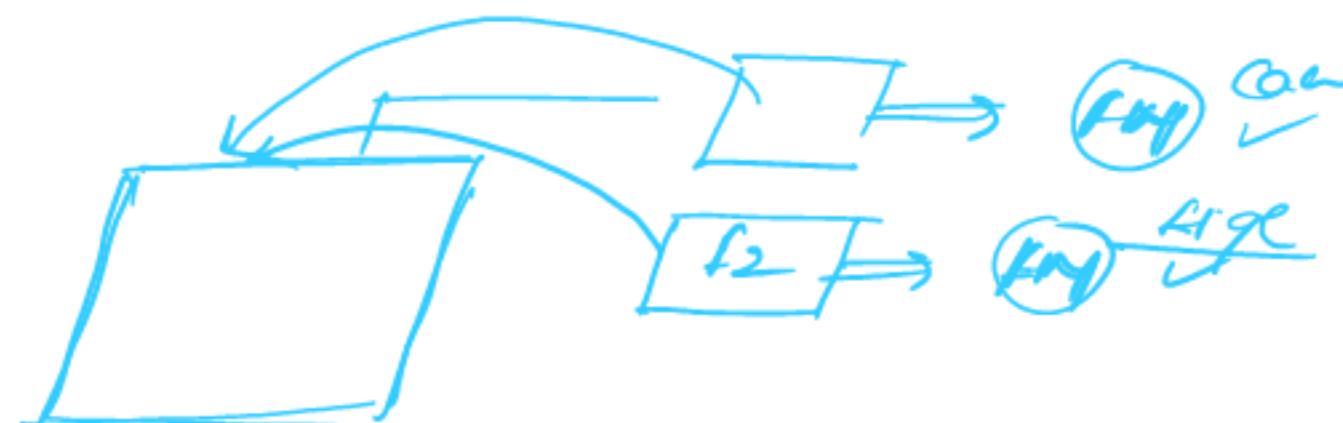
$$36$$

$$n + 2p - f + 1$$

$$6 + 2 \times 0 - 3 + 1$$

$$4 \times 4$$

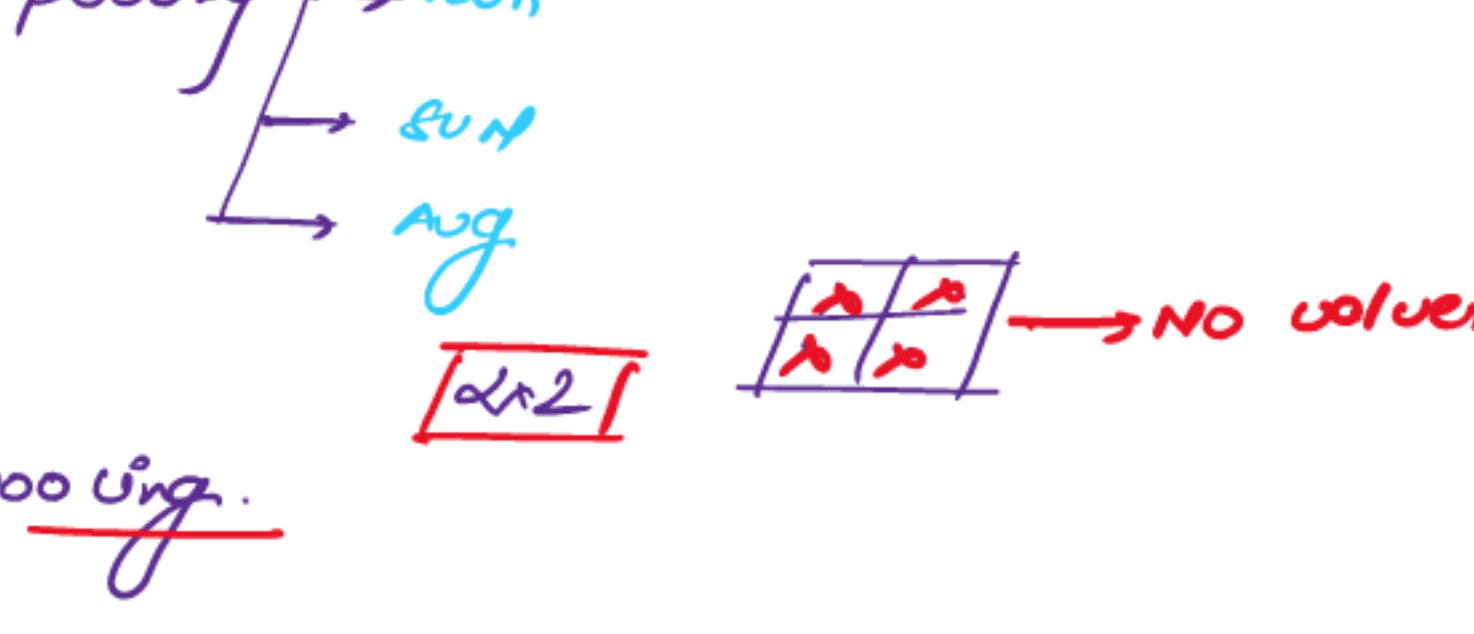
$$16$$



pooling layer

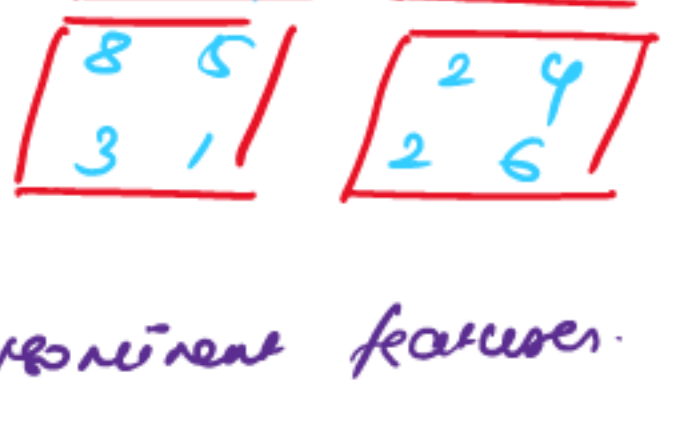
pooling layers involve
sliding a 2 dimensional
filter on each feature map
lying in region.

image
to
filter
to
FM
to
pooling



max pooling

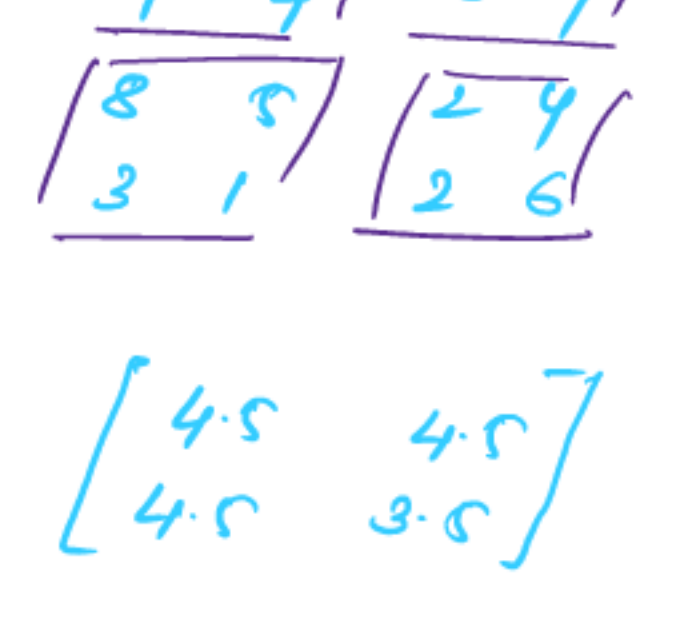
It usually picks
max elements from
the region.



$\begin{bmatrix} 9 & 7 \\ 8 & 6 \end{bmatrix}$ most prominent features.

Avg pooling

avg pooling does
the avg of present
power as region.



min pooling

picks up min
elements from specific
region.

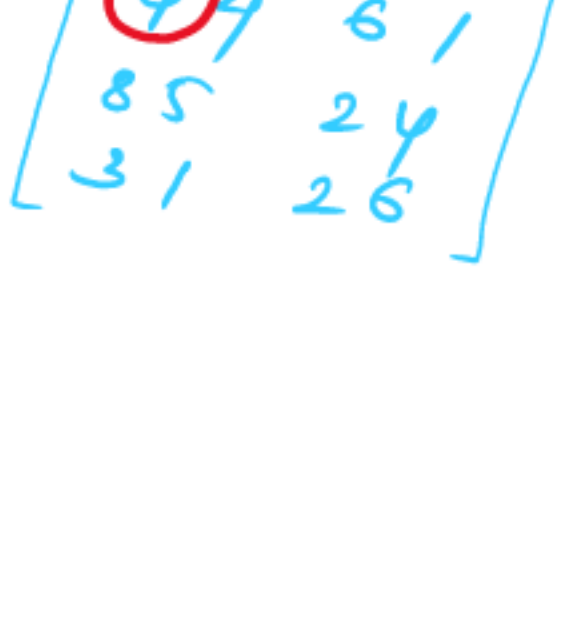


global max pooling

min

sum

avg



9

Spatial information
may be lost if filter
is too big



Flattened

1 max pooled FM



Image

filter



10

10

10

10

10

10

2

10

10

10

10

10

10

6x6

3x3

4x4

4x4

4x4

4x4