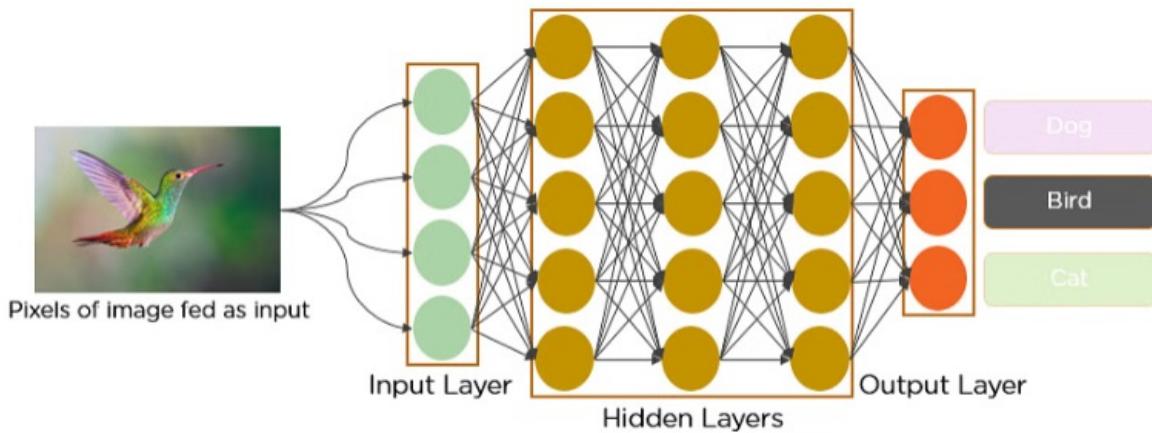


How facial recognition works on social media, or how object detection helps in building self-driving cars, or how disease detection is done using visual imagery in healthcare? It's all possible through convolutional neural networks (CNN).



How it work- lets see at a glance in very short?

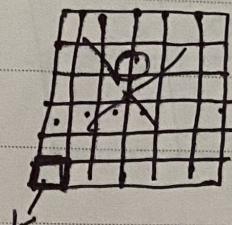
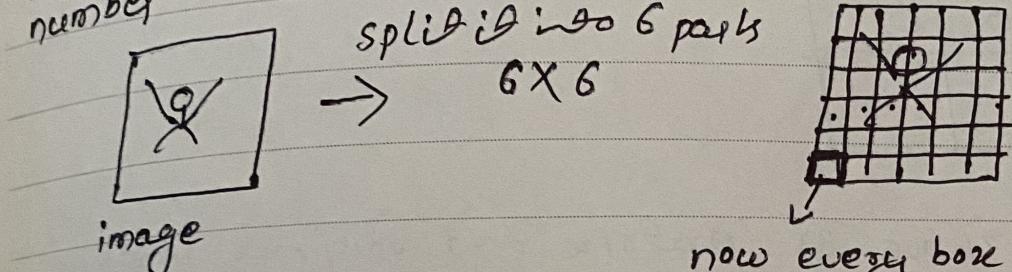
- There's an image of a bird, and you want to identify whether it's really a bird or some other object. The first thing you do is feed the pixels of the image in the form of arrays to the input layer of the neural network (multi-layer networks used to classify things). The hidden layers carry out feature extraction by performing different calculations and manipulations. There are multiple hidden layers like the convolution layer, the ReLU layer, and pooling layer, that perform feature extraction from the image. Finally, there's a fully connected layer that identifies the object in the image.

CNN

CNN

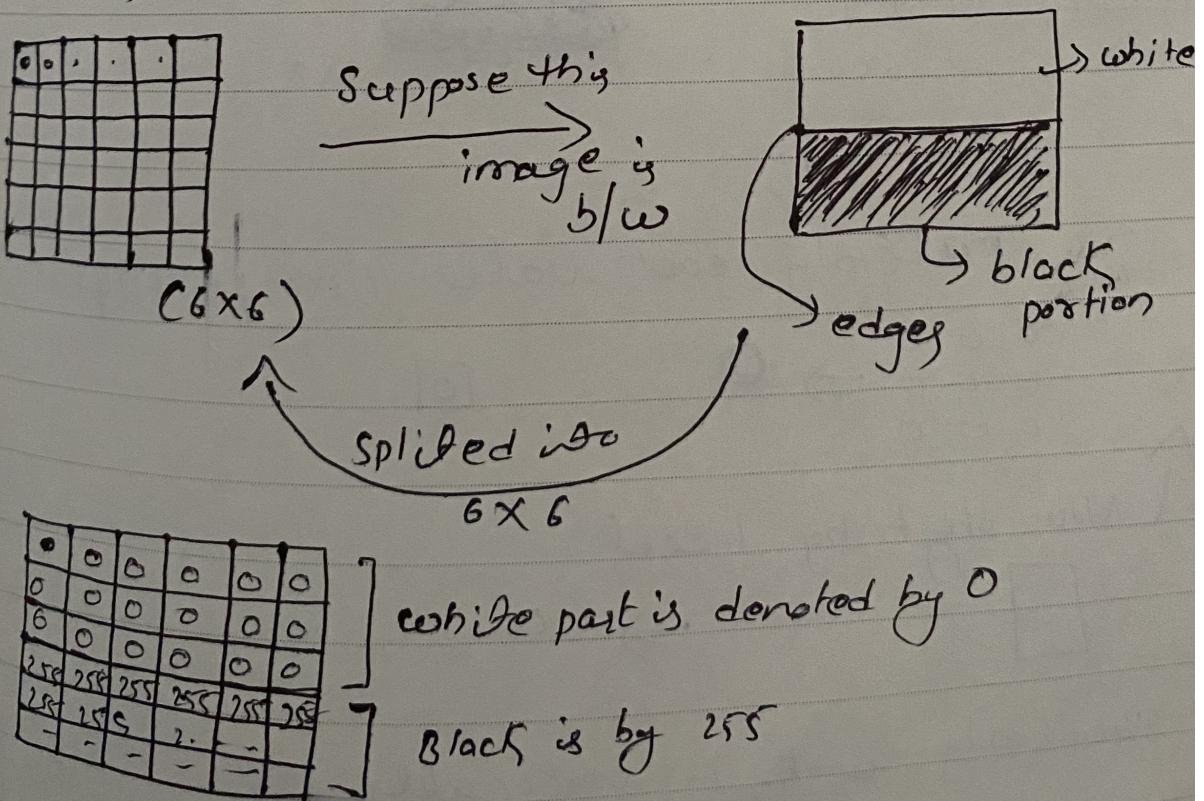
Date: / /

- Suppose we have a image but, if we direct give image as input, computer and Machine will not understand it. So, we have to convert image into numbers.



now every box is called pixels and in computer every box has pixel value and that value determine color of block.

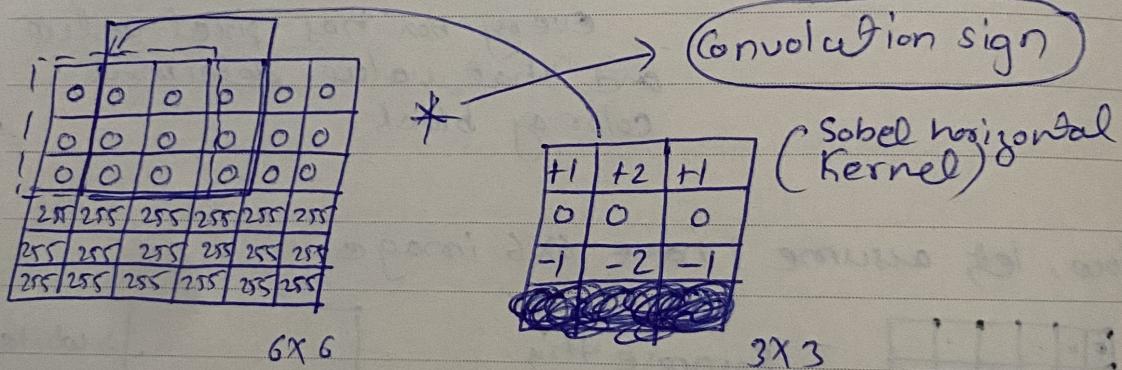
- Now, let assume I have 6x6 image



After various research scientist discovered 3x3 matrix which is known as Kernel or filter.

1	2	1
0	0	0
-1	-2	-1

Suppose I have 3×3 matrix and original image 6×6 . so I add 3×3 into original.



↓
multiply each value $(1 \times 0) + (2 \times 0) + (1 \times 0) \dots$

$$\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \Rightarrow \textcircled{0} \quad \boxed{0}$$

↑ Now shift by 1 cell

$$\square \Rightarrow \textcircled{0}$$

Similarly works in vertically too.

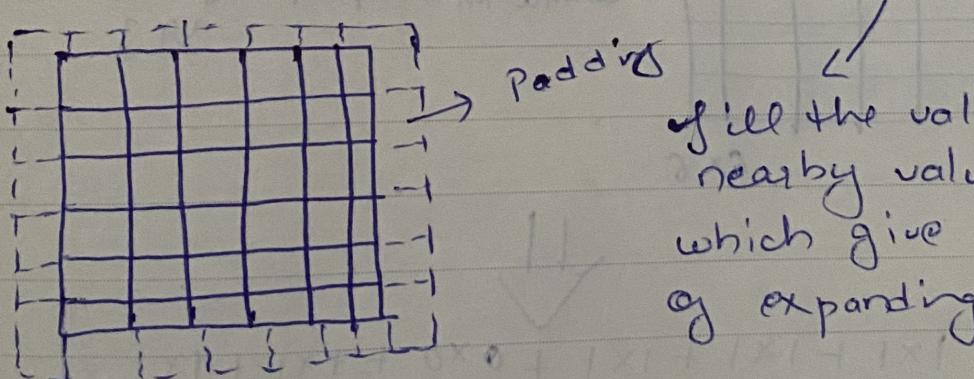
formula \rightarrow

$$\boxed{(n-k+1) \times (n-k+1)}$$

* Now suppose I have 6×6 image and have a Kernel of 3×3 , but I don't want reduction in matrix, instead of 4×4 , I want 6×6 image.

So, now, here, we expand the original image by adding one layer on top, bottom and on sides, This is called Padding.

So, what should be extra layer values.



fill the value with nearby values so which give idea of expanding.

formula \rightarrow

$$(n-k+2p+1) \times (n-k+2p+1)$$

② \downarrow creating layers above and below

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

6×6 (output)

finally we get edges ~~Convolution~~ Convolution.

0	0	0	0
-1020	·	·	-1020
-1020			
0	0	0	0

4x4

But pixel value should be between 0 and 255, minimum value should be 0 and high value 255.

Convert it ↴

255	255	255	255	→ black
0	0	0	0	edge (white)
0	0	0	0	
255	255	255	255	→ black

* Are there any other Kernel to detect vertical edge.

→ Yes.

H1	0	-1
H2	0	-2
H1	0	-1

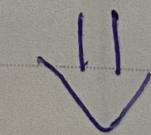
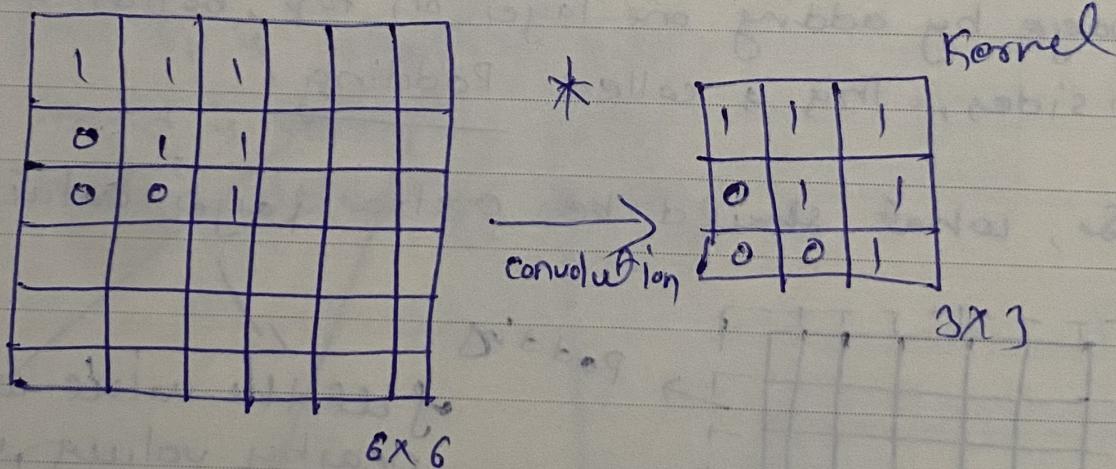
vertical edge detector.

* Strides

Jumping 2 cell horizontally and vertically.

$$\left(\frac{n - ks + 1}{s} \right) \times \left(\frac{n - ks + 1}{s} \right)$$

* Pooling



$$1 \times 1 + 1 \times 1 + 1 \times 1 + 0 \times 0 + 1 \times 1 + 1 \times 1 \\ + 0 \times 0 + 0 \times 0 + 1 \times 1$$

$$= 1 + 1 + 1 + 0 + 1 + 1 + 0 + 0 + 1 = 6$$

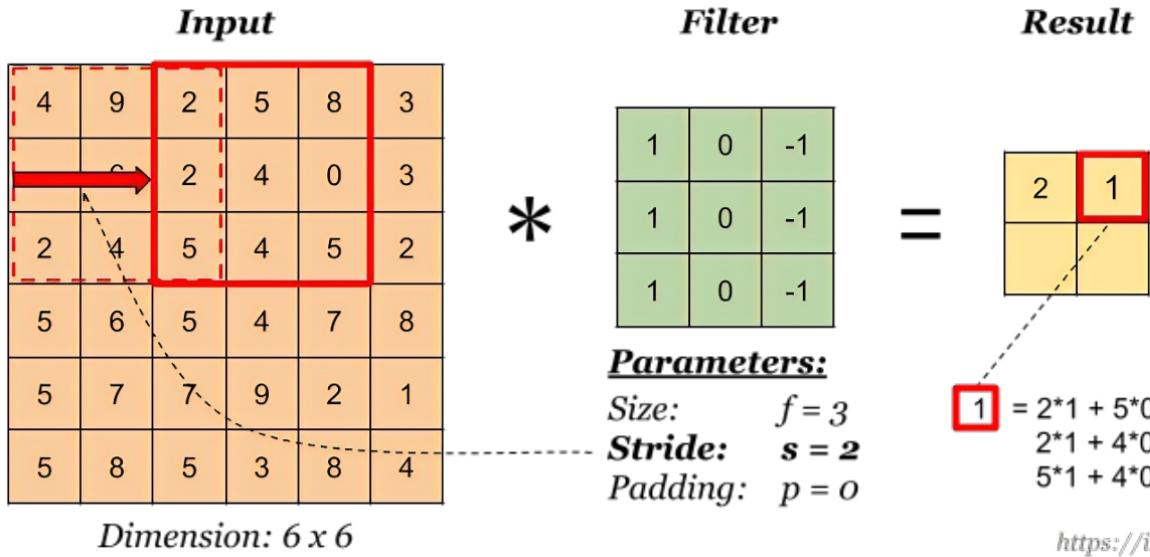
↓ now suppose

4	2	2	3
2	3	4	2
3	2	4	4
2	2	3	3

Pooling

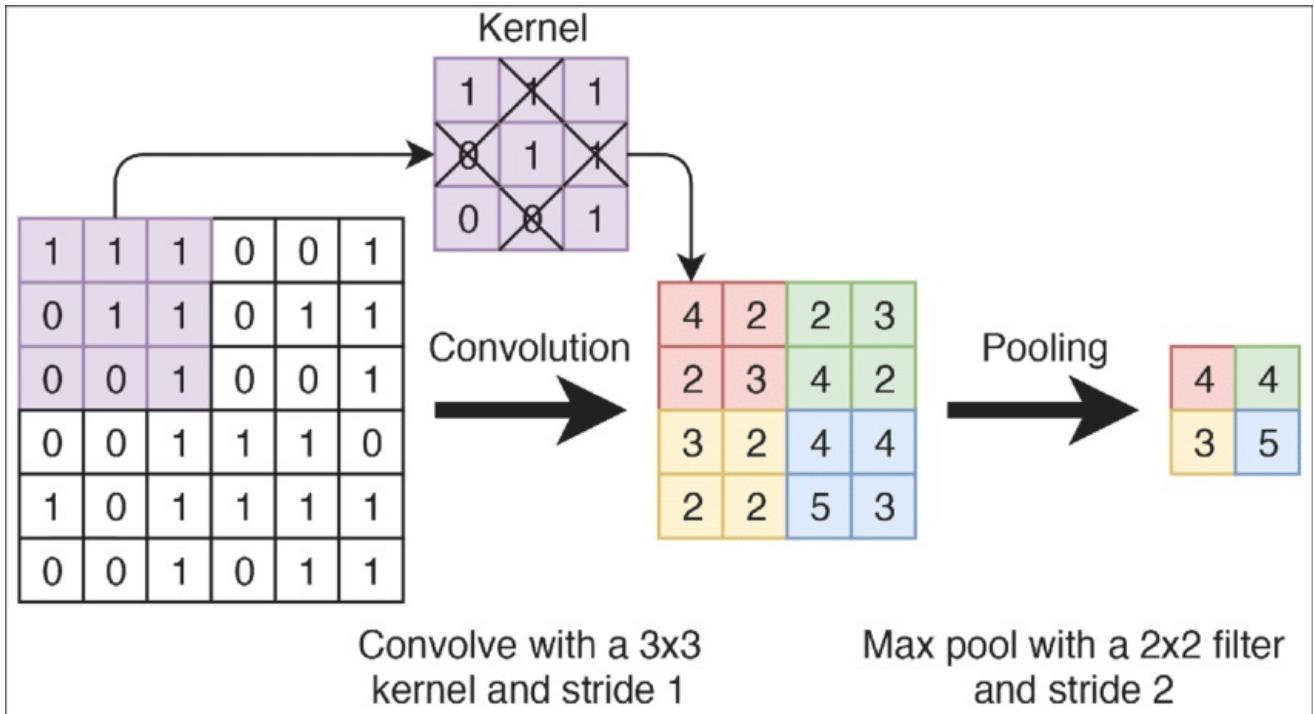
4	4
3	5

(Max pool with 2×2 filter)



<https://indoml.com>

Pooling



Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js