CNN

- Convolutional neural network

COMPUTER VISION

- Computer vision is **a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs** — and take actions or make recommendations based on that information.
- It mimics human vision to provide vision to machines to capture images and to make prediction based on that.

-> CNN => data it takes in a visual data

i.e image or a video

image => unstructured data

WHAT IS IMAGE?

- Image is 2-D representation of real world objects made up of Pixels.
- Images are made from pixels .
- Pixels mathematical range is 0-255.

image => pixels => matrix

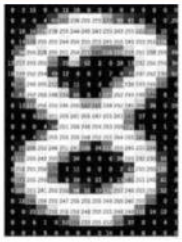
range (0 - 255)

TYPES OF IMAGES

GREY SCALE IMAGE

Grey scale image is made up of black and white pixels.

Pixel value of black is 0 and white is 255.



Black = 0

white = 255

any color : 3 channel i.e R G B

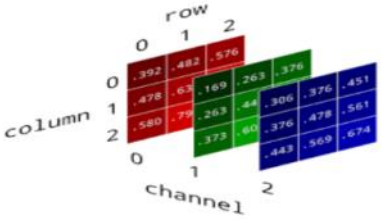
(0 0 0) = Black

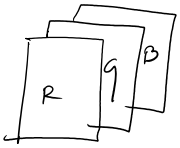
(255, 255, 255) = white

TYPES OF IMAGES

RGB SCALE IMAGE

- RGB stands for red , green and Blue.
- It is made up of three channels i.e red , green and blue channels.
- Overlap of all these three channels is a colorful image.
- Each channel has got the range (0,255).





⟹ X color image

1. Red = (255, 0, 0)

G : (0, 255, 0)

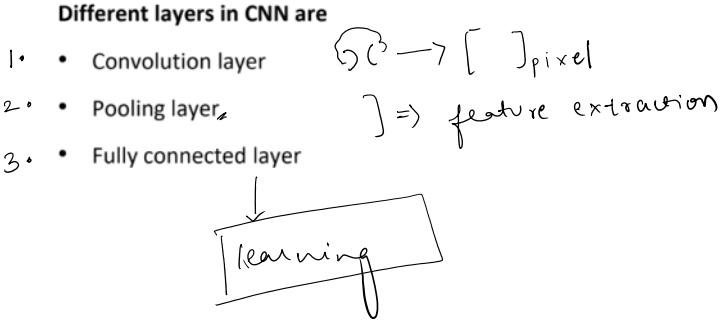
B = (0, 0, 255)

CNN

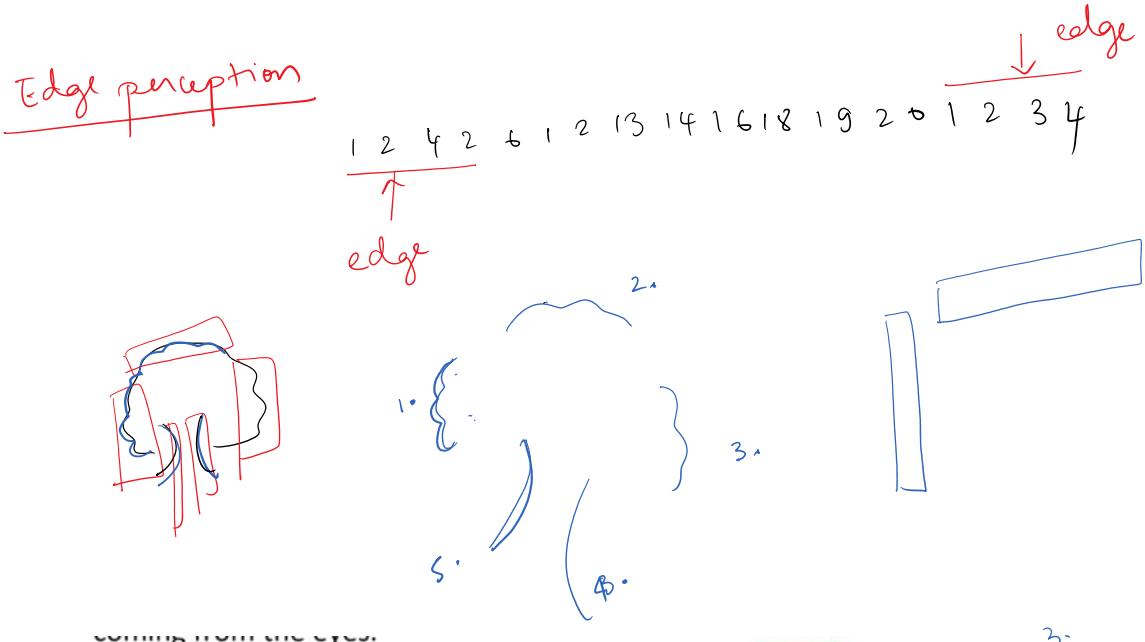
- Convoluted Neural Network is a deep learning algorithm which is widely used for image / object recognition and classification.
- CNN is subset of machine learning highly suitable for computer vision tasks and for applications where object recognition is vital , such as self driving cars and facial recognition.

Different layers in CNN are

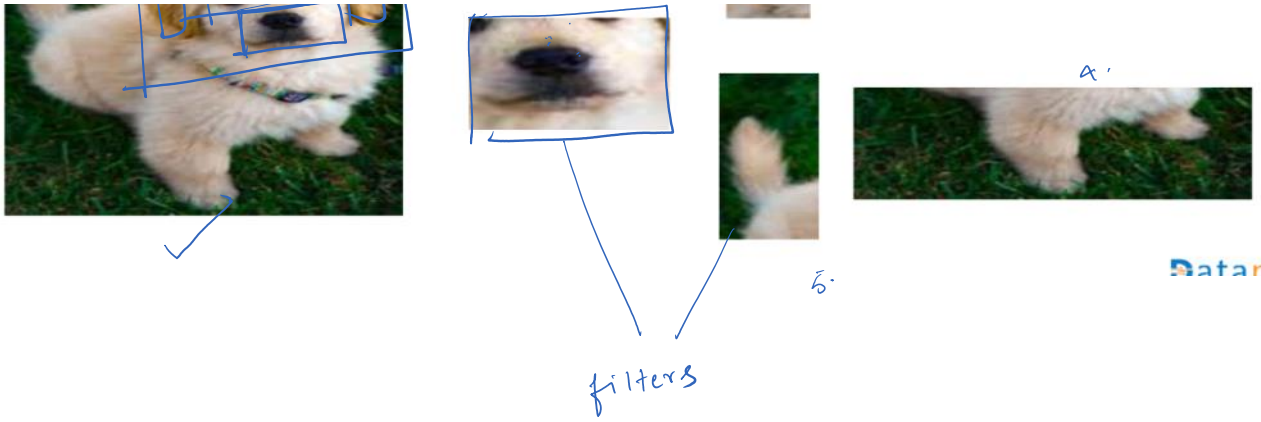
1. • Convolution layer
2. • Pooling layer
3. • Fully connected layer



HOW DO YOU CAPTURE IMAGE?



coming from the eyes.



Datamites

Working

HOW CNN WORKS?

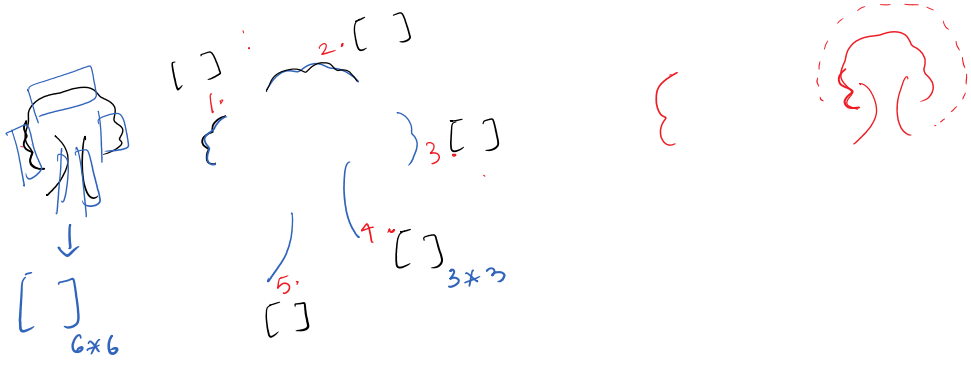
Step1: Take image

2. **Step2: Convert image into pixels**

0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1

6X6

3. Choose the filters | edges



6x6

4. The filters revolve around the original image & forms another image called as convoluted image / feature map

↓

entire info about the extracted filter

convolution layer [convolution ⇓ 1st layer of CNN

0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1

original pixel

6x6

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

filter 3*3

=

0	-4	-4	0

4*4

feature map [holds the entire info]

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \Rightarrow \begin{bmatrix} (0*1)+(0*0)+(0*-1)+ \\ (0*2)+(0*0)+(0*-2)+ \\ (0*1)+(0*0)+(0*-1) \end{bmatrix} \quad 0$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \Rightarrow 0+0-1+0+0-2+0+0-1 = -4$$

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \Rightarrow -4$$

0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1
0	0	0	1	1	1

1.

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

3*3

0	-4	-4	0
0	-4	-4	0
0	-4	-4	0
0	-4	-4	0

4x4

0	0	0	1	1	1

6*6

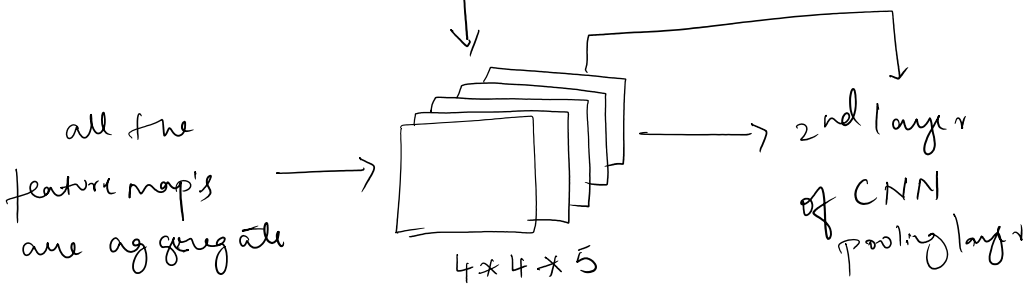
4x4

2. → [ ]4*4

3. → [ ]4*4

4. → [ ]4*4

5. → [ ]4*4



(2)nd layer of CNN => pooling layer

↓

summarize the data and it will reduce the dimension of feature maps

pool_size = 2*2

0	-4	-4	0
0	-4	-4	0
0	-4	-4	0
0	-4	-4	0

4x4

Convoluted imag

pooling

downsize the data

Max → $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ pool_imag 2*2

Avg → $\frac{0-4+0-4}{4} = -\frac{8}{4} = -2$ → $\begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}$ 2*2

3] flatten layer

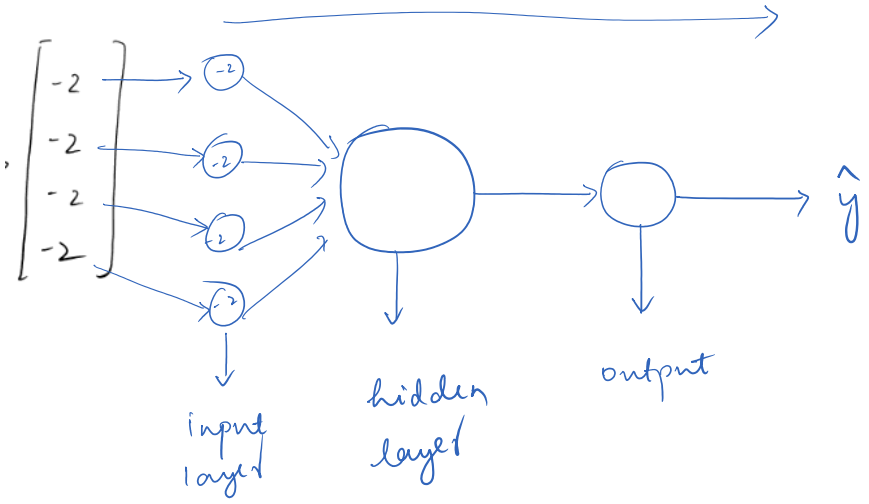
$\begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}$ 2*2 pool_imag

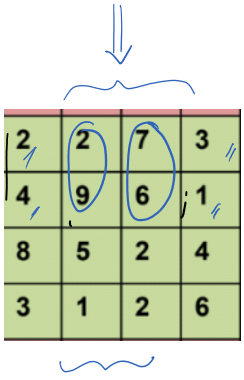
flattening → $\begin{bmatrix} -2 \\ -2 \\ -2 \\ -2 \end{bmatrix}$

pool_image

3rd layer of CNN i.e fully connected layer (FC layer)

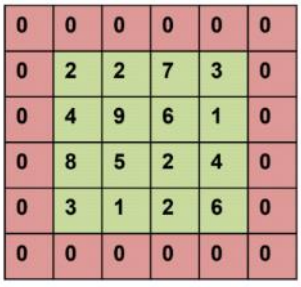
⇒ actual learning of the data





$\begin{bmatrix} 2 & 2 \\ 4 & 9 \end{bmatrix}$ $\begin{bmatrix} 2 & 7 \\ 9 & 6 \end{bmatrix}$ $\begin{bmatrix} 7 & 3 \\ 6 & 1 \end{bmatrix}$

$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$



PADDING

- **Padding** is process of adding new layers to the border of an image.
- The pixel in the corner will only get covers one time but if you take the middle pixel it will get covered more than once that means we have more information from middle pixels than corner pixels.
- Challenges faced are like "**Shrinking output**" and "**loss of information on corners**" and to overcome these challenges we have padding concept.
- Padding helps by adding extra rows and columns on the outer dimension of the data.

0	0	0	0	0	0
0	2	2	7	3	0
0	4	9	6	1	0
0	8	5	2	4	0
0	3	1	2	6	0
0	0	0	0	0	0

TYPES OF PADDING

VALID PADDING(NO PADDING)

It says that new layers are not added

2	2	7	3
4	9	6	1
8	5	2	4
3	1	2	6

ZERO /SAME PADDING

Add zeros at the edges of layers.

CONSTANT PADDING

Adding constant values at the edges of layer

0	0	0	0	0	0
0	2	2	7	3	0
0	4	9	6	1	0
0	8	5	2	4	0
0	3	1	2	6	0
0	0	0	0	0	0

1	1	1	1	1	1
1	2	2	7	3	1
1	4	9	6	1	1
1	8	5	2	4	1
1	3	1	2	6	1
1	1	1	1	1	1