Till now,

We have used the Deep Neural Networks to perform the task of :

- Regression
- Binary Classification
- Multi Class classification

From this module,

We are going to learn as how to use the Deep Neural Networks for the task of Computer Vision

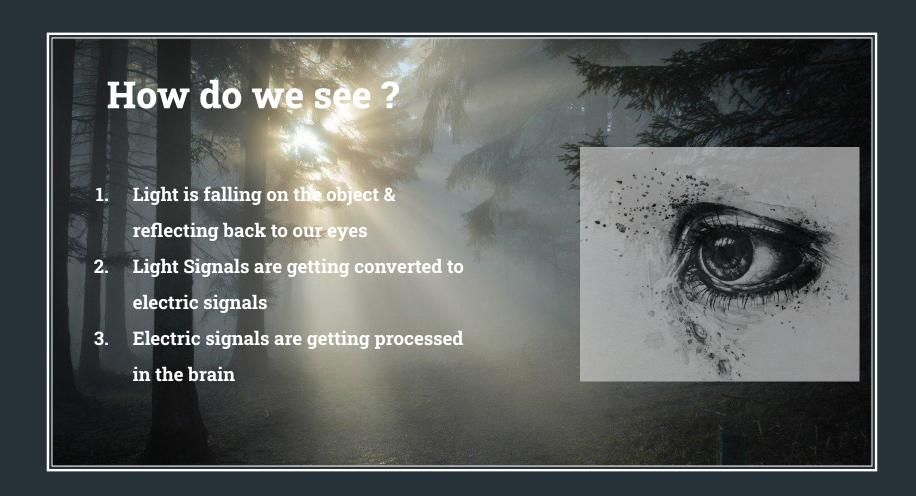
Agenda

- What is Computer Vision?
- Understanding Convolution Operation





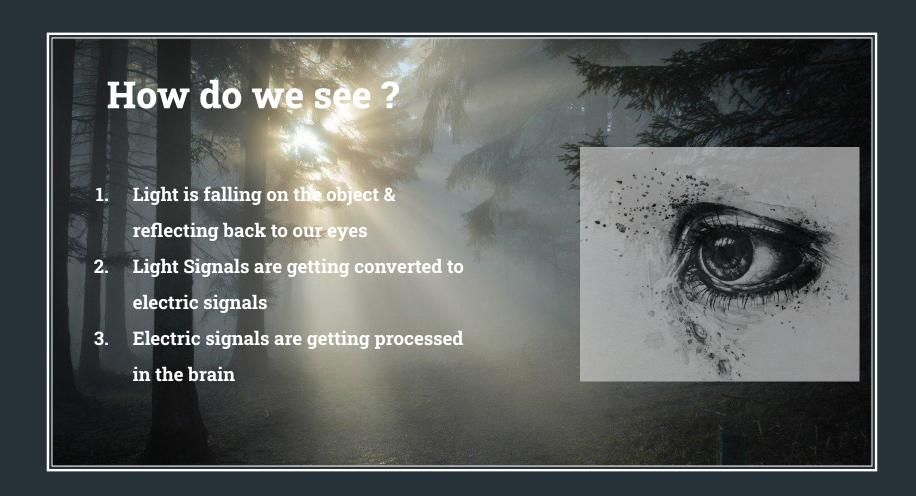
How do we see the objects as Humans?



From Computer's Perspective

- Data is getting captured from outside world
- 2. Data will be transformed
- 3. Data will be processed





What does Computer Vision mean?

Giving the ability to computers to see the outside world, with the use of Hardware, Software and Analytical ability.

Hardware - Webcam, Camera to capture the Outside world

Software - Video/Image processing softwares to convert the image/video to digital format

Analytical Ability - Deep Learning Neural Networks like CNN







Convolution

In mathematics, convolution is a mathematical operation on two functions (f and g) that produces a third function expressing how the shape of one is modified by the other.

In image processing:

Its manipulation of pixel values of an image

What do you when you click a Photo?

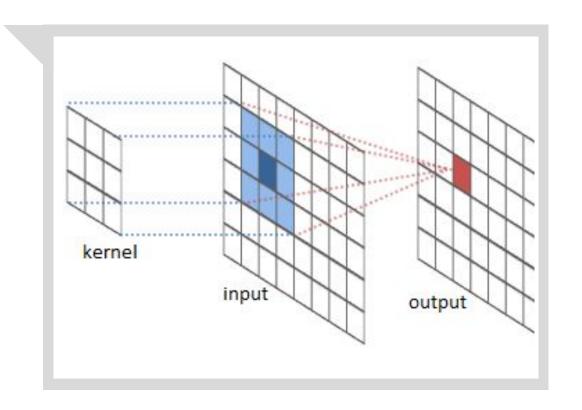


Can you observe the water bodies?





How do we perform mathematically?



Understanding Image & Video Data

Storing the Images

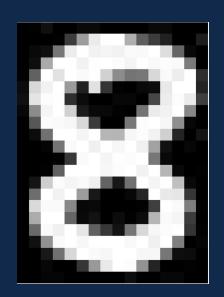


Image



Pixels

Grayscale Image



0	2	15	0	0	11	10	0	0	0	0	9	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0	29
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10	0
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124	1
2	98	255	228	255	251	254	211	141	116	122	215	251	238	255	49
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255	36
16	229	252	254	49	12	0	0		7	0	70	237	252	235	62
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137	C
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6	C
0	13	113	255	255	245	255	182	181	248	252	242	208	36	0	19
	0	5	117	251	255	241	255	247	255	241	162	17	0		C
0	0	0	4	58	251	255	246	254	253	255	120	11	0		C
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0	4
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9	C
0	111	255	242	255	158	24	0	0	6	39	255	232	230	56	C
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125	3
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61	C
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52	4
0	18	146	250	255	247	255	255	255	249	255	240	255	129	0	5
0	0	23	113	215	255	250	248	255	255	248	248	118	14	12	C
0	0	6		0	52	153	233	255	252	147	37	0	0	4	1
0	0	5	5	0	0	0	0	0	14	1	0	6	6	0	0





2d Matrix of shape :

50 X 50

Each Element Represents the Pixel Intensity

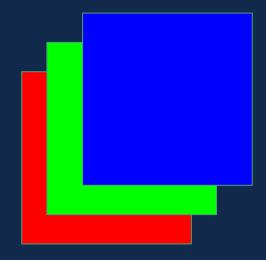
Range: 0 to 255

Image with Size : 50 X 50

Color Image Representation



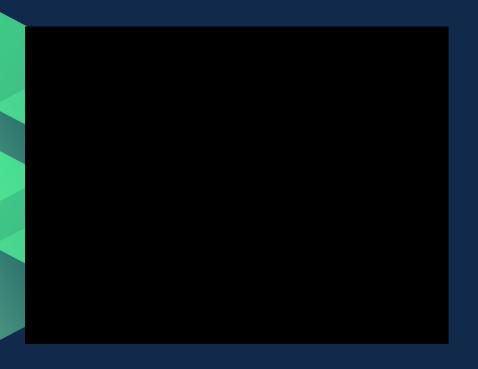
Color Image



Array: 3X50 X 50

Dimension: 50 X 50

Video Data Representation

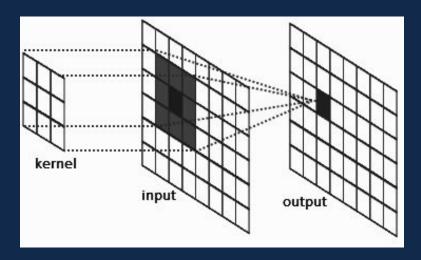


Collection of Images.

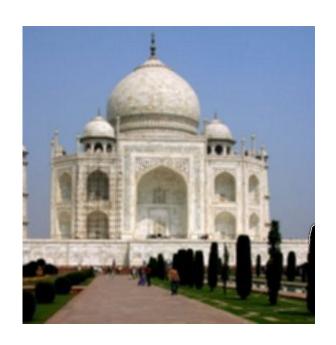
5 Sec Video of 50X50 at 24 fps:

50 X 50 X 3 X 24 X 5

How do we perform mathematically?



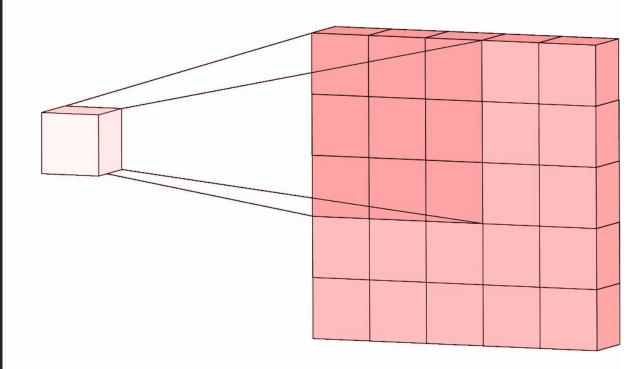
Example Convolution Operation







Intuition of Convolution

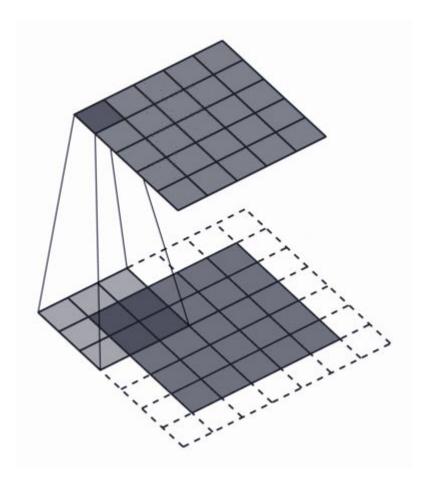


Mathematical Working

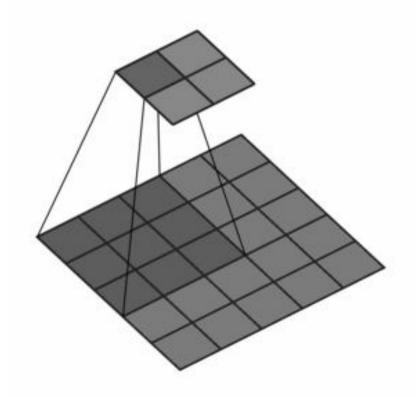
30	3,	22	1	0
02	02	10	3	1
30	1,	22	2	3
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

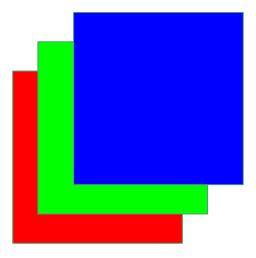
Padding

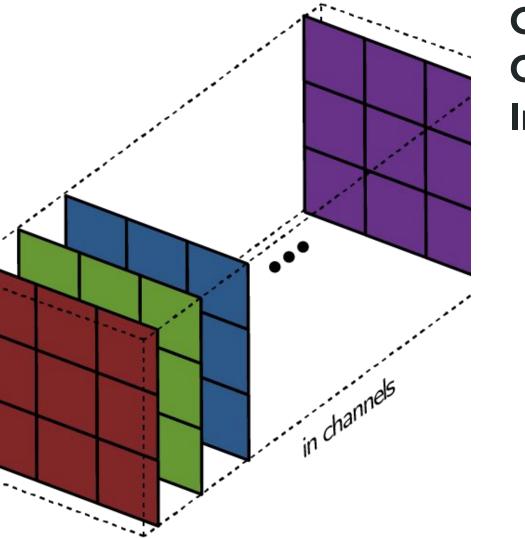


Stride

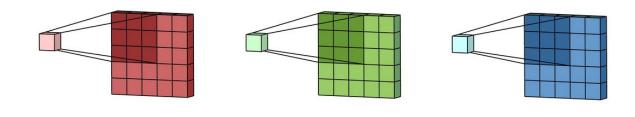


Color Image





Convolution
Operation on Color
Image









Pooling layers

Max pool : Maximum Value

Average Pool : Average Value

Max Pool

20	30
112	37

12	20	30	0
8	12	2	0
34	70	37	4
112	100	25	12



13	8
79	20

