Today

Building a Convolutional NN and image data preprocessing

In the upcoming week you will next learn the concepts that changed the whole course of Machine Learning.

Convolutional Neural Networks, these can detect almost anything from Images and are the most interesting concepts in Deep Learning.

A guide to convolution arithmetic for deep learning
(HIGHLY Recommended)

Today we will learn

Introduction to computer vision

Convolution Operation

Padding

Stride Convolution

Convolution on RGB

What is computer vision

Computer Vision, often abbreviated as CV, is defined as a field of study that seeks to develop techniques to help computers "see" and understand the content of digital images such as photographs and videos

Computer vision

- it is a multidisciplinary field

- subfield of artificial intelligence and machine learning

- involve the use of specialized methods and make use of general learning algorithms
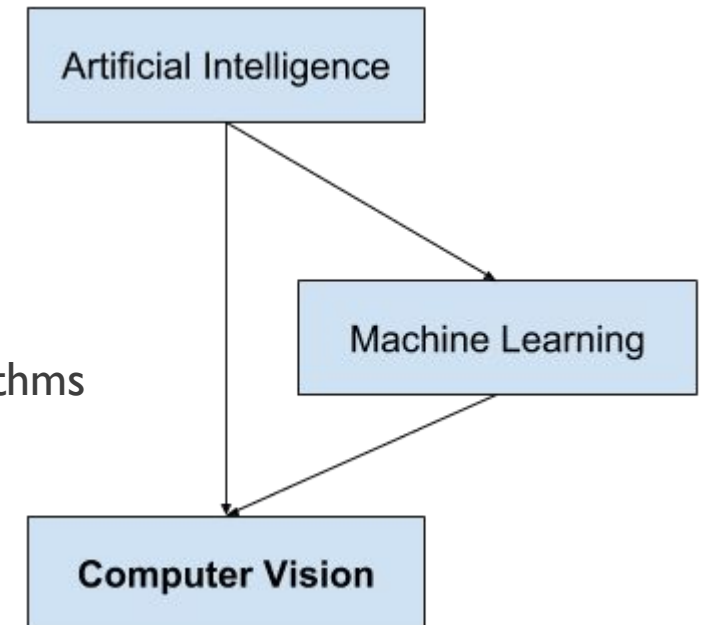
Computer vision - applications

- Optical character recognition (OCR)
- Machine inspection
- Retail (e.g. automated checkouts)
- 3D model building (photogrammetry)
- Medical imaging
- Automotive safety
- Match move (e.g. merging CGI with live actors in movies)
- Motion capture (mocap)
- Surveillance
- Fingerprint recognition and biometrics

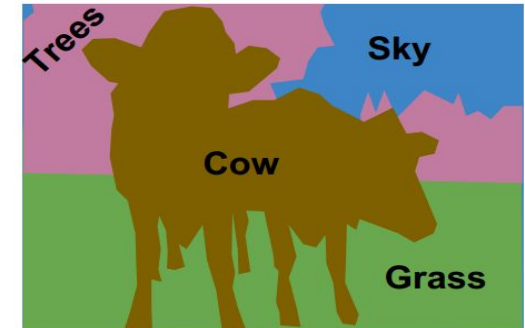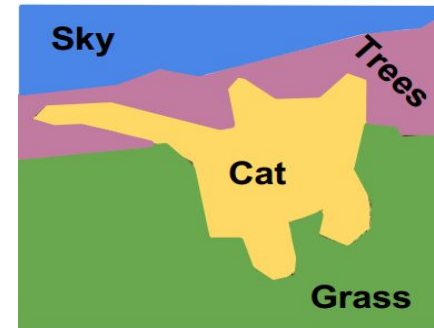# COMPUTER VISION

Image Classification

Cat? (0/1)

64x64

Object detection

Image Segmentation

This image is CC0 public domain

Sky
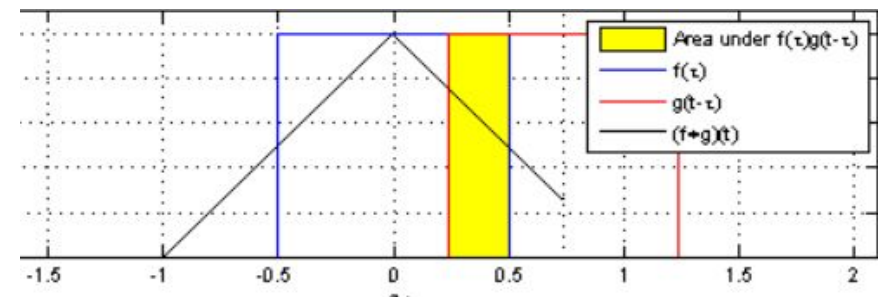
Trees
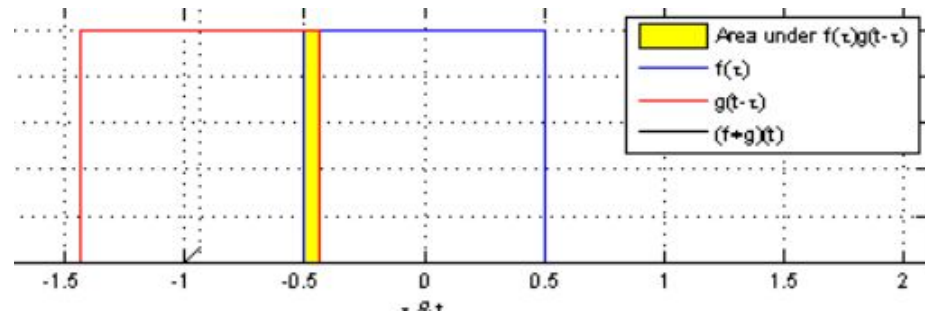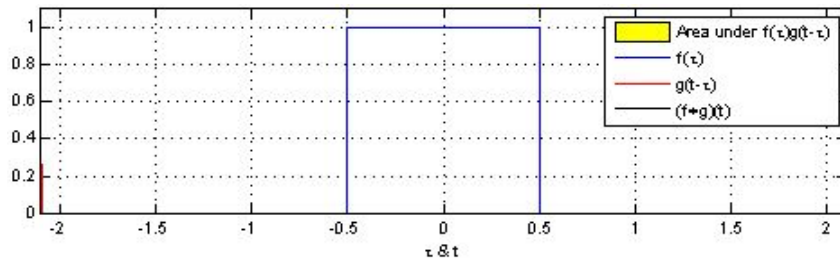
Cat

Grass

Trees

Sky

Cow

Grass

What is convolution ?

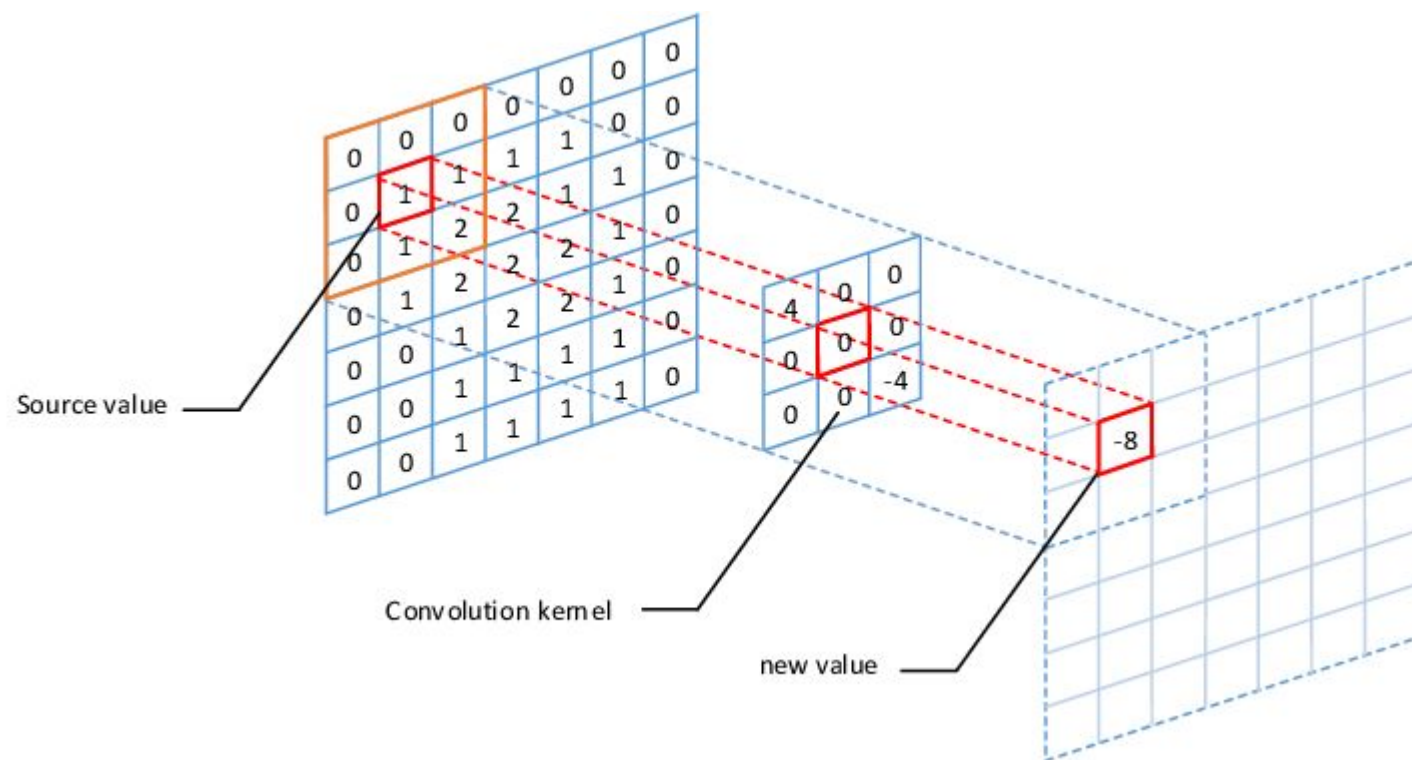In mathematics (in particular, functional analysis),

convolution is a mathematical operation on two functions (f and g)

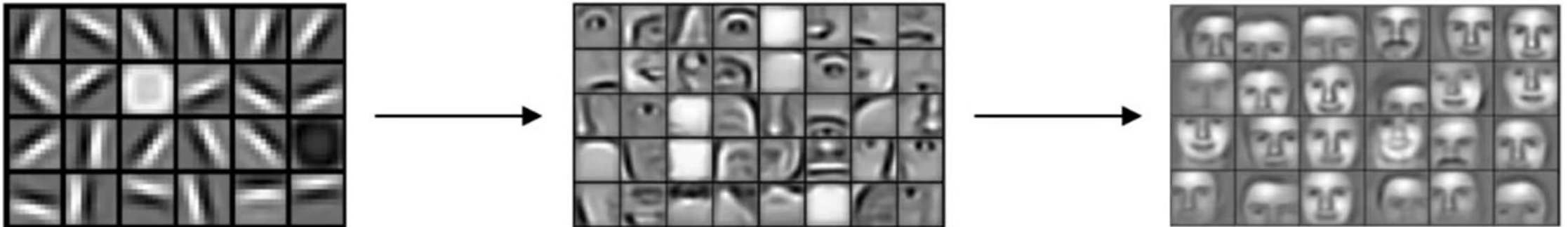that produces a third function that expresses how the shape of one is modified by the other

Example

Source value

Convolution kernel

new value

CONVOLUTION OPERATION

EDGE DETECTION using convolution

Convolution Operation

$= 3*1 + 1*1 + 2*1 + 0*0 + 5*0 + 7*0 + 1*-1 + 8*-1 + 2*-1$
$= 3 + 1+ 2 - 1 - 8 - 2$
$= -5$

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

*

1	0	-1
1	0	-1
1	0	-1

3x3 Filter

=

-5	-4	0	8
-10	-2	2	3
0	-2	-4	-7
-3	-2	-3	-16

Output Image

6x6 Grayscale Image

CONVOLUTION OPERATION

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

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

What is Padding in Machine Learning / convolution ?

Padding is a term relevant to convolutional neural networks

It refers to the amount of pixels added to an image when it is being processed by the kernel of a CNN

For example, if the padding in a CNN is set to zero, then every pixel value that is added will be of value zero

If, however, the zero padding is set to one, there will be a one pixel border added to the image with a pixel value of zero

It extends the area of an image in which a convolutional neural network processes.

https://poloclub.github.io/cnn-explainer/

PADDING

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

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

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

Code example

Types of Padding

There are three types of padding:

Same padding: input_shape=(28, 28, 1))), padding='same'))

Causal padding

Valid padding

Ref: https://github.com/christianversloot/machine-learning-articles/blob/main/what-is-padding-in-a-neural-network.md

STRIDE CONVOLUTION

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

*

3	4	4
1	0	2
-1	0	3

91	100	83
69	91	127
44	72	74

$$\frac{n + 2 * p - f}{s} + 1$$

Stride

Stride is the number of pixels shifts over the input matrix.

When the stride is 1 then we move the filters to 1 pixel at a time.

When the stride is 2 then we move the filters to 2 pixels at a time and so on.

Code example:

```
model.add(Conv2D(1, (3,3), strides=(2, 2), input_shape=(8, 8, 1)))
```

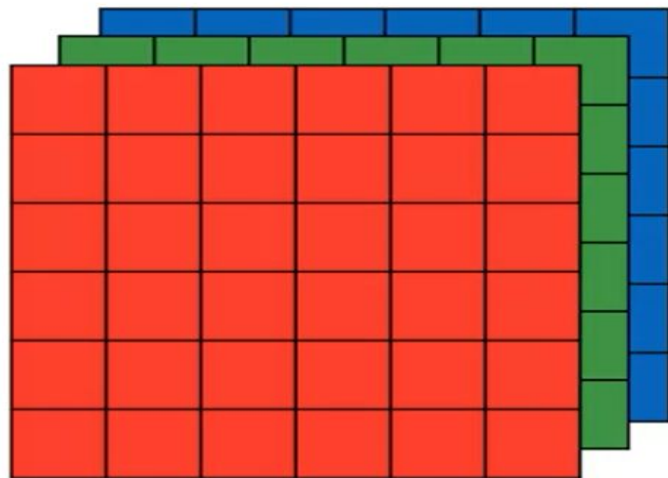# The below figure shows convolution would work with a stride of 2.

1	2	3	4	5	6	7
11	12	13	14	15	16	17
21	22	23	24	25	26	27
31	32	33	34	35	36	37
41	42	43	44	45	46	47
51	52	53	54	55	56	57
61	62	63	64	65	66	67
71	72	73	74	75	76	77

Convolve with 3x3 filters filled with ones

108	126	
288	306	

$$\frac{n + 2*p - f}{s} + 1$$

Ref: https://cs231n.github.io/convolutional-networks/#conv

CONVOLUTION ON RGB

$6 \times 6 \times 3$ * $3 \times 3 \times 3$ = 4×4

$3 \times 3 \times 3$ = 4×4
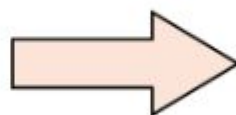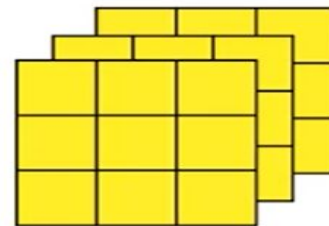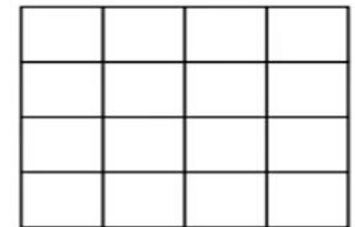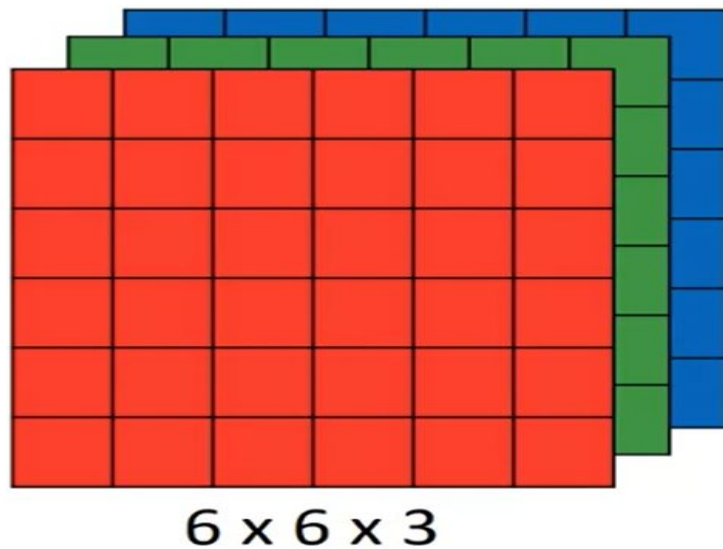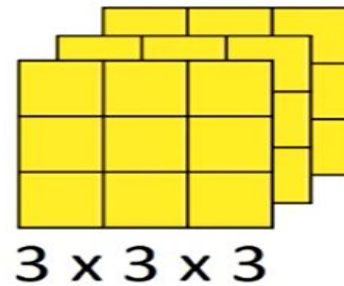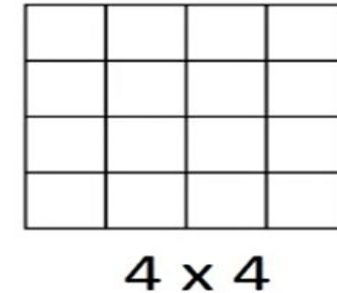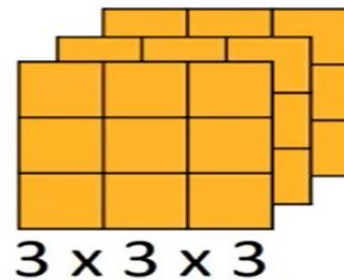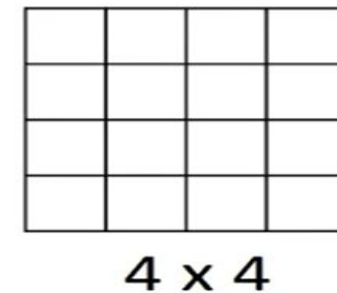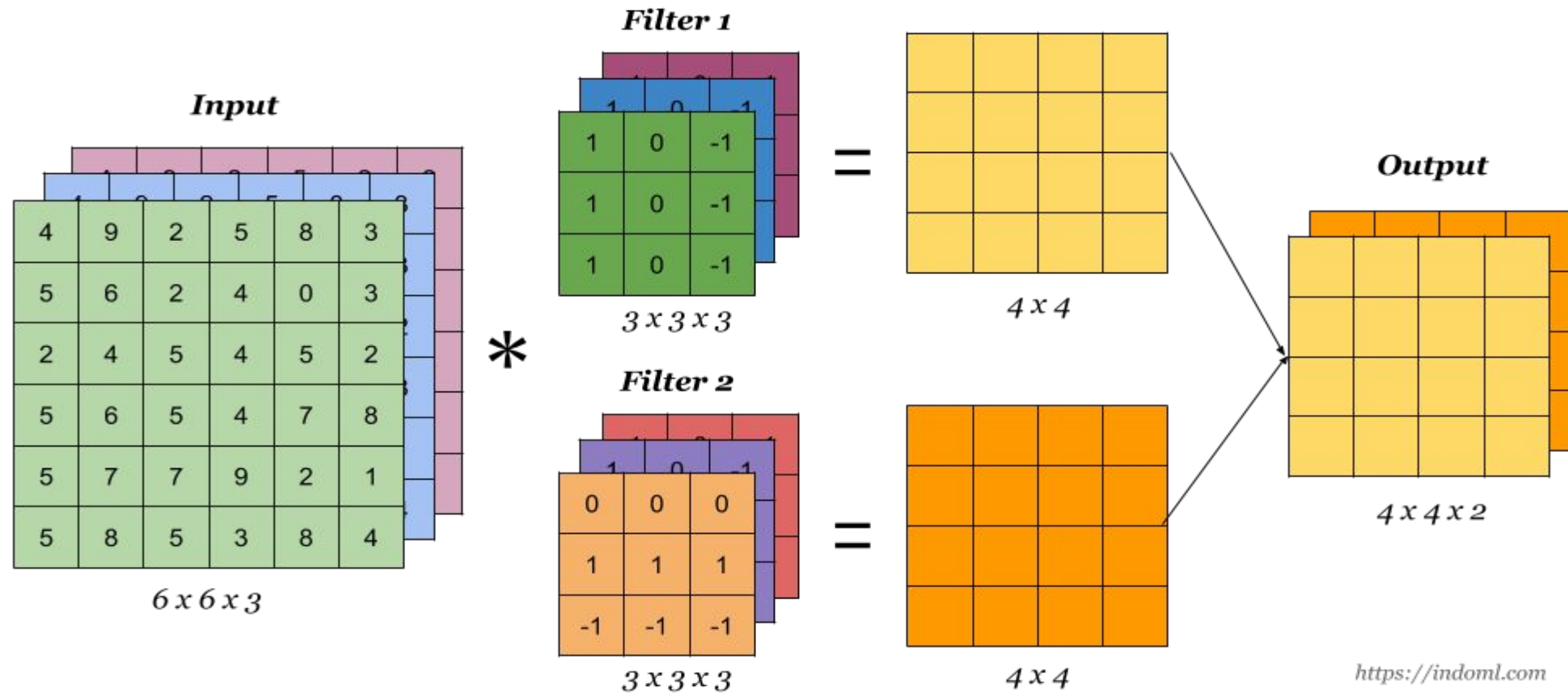
Example

Next session -

- Pooling
- Dropout
- Batch Normalisation
- Data Augmentation

Time to practice some questions!

Q-1 what do you understand by computer vision and is it same as deep neural network? Also, write some applications of CV.

Q-2 What is convolution in CV?

Q-3 What do you mean by the term Padding and Stride? Why do we use padding and stride in convolution of an image? Also write the equation for padding and stride to find dimensions of an output image.