Dhirubhai Ambani University

(Formerly known as DA-IICT)

Topic: Deep Learning: CNN

Course: Programming Lab

Course Code- PC503

Dr. Ankit Vijayvargiya

Assistant Professor

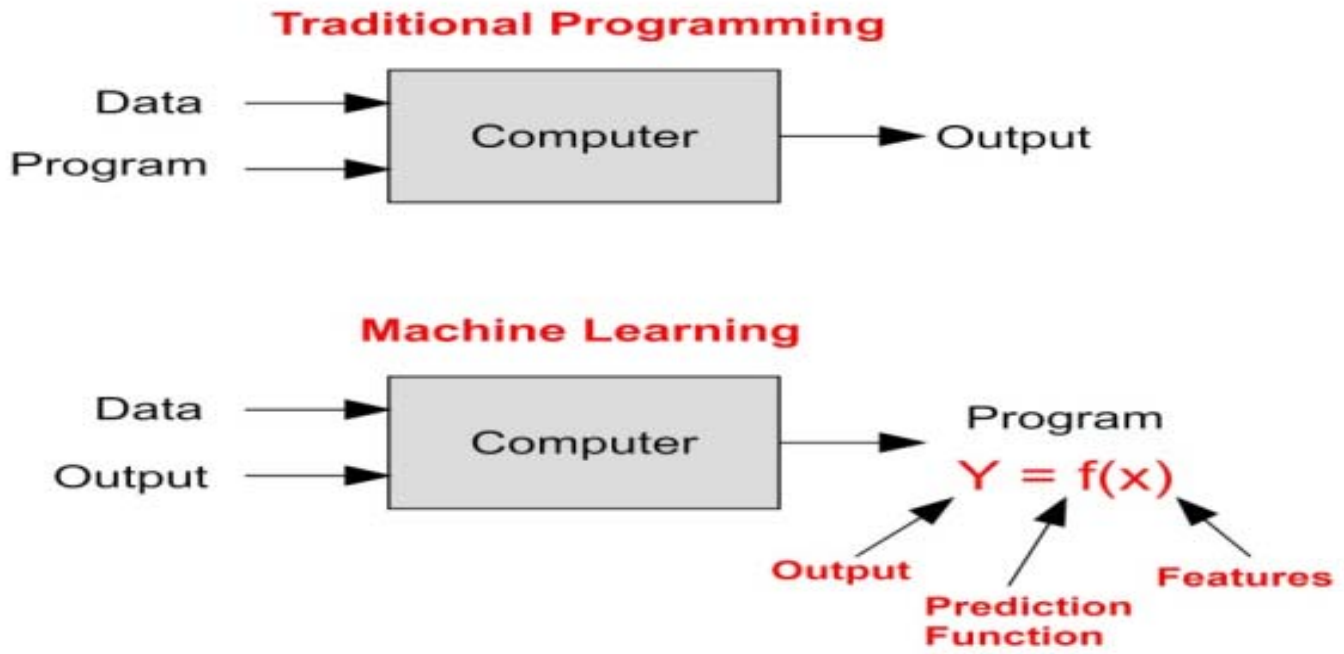
Room No. 4205, Faculty Block 4

Email: ankit_Vijayvargiya[at]dau.ac.in

Phone: 079-68261628(O), 7877709590(M)

Dhirubhai Ambani University
Formerly known as DA-IICT

Machine Learning (ML)

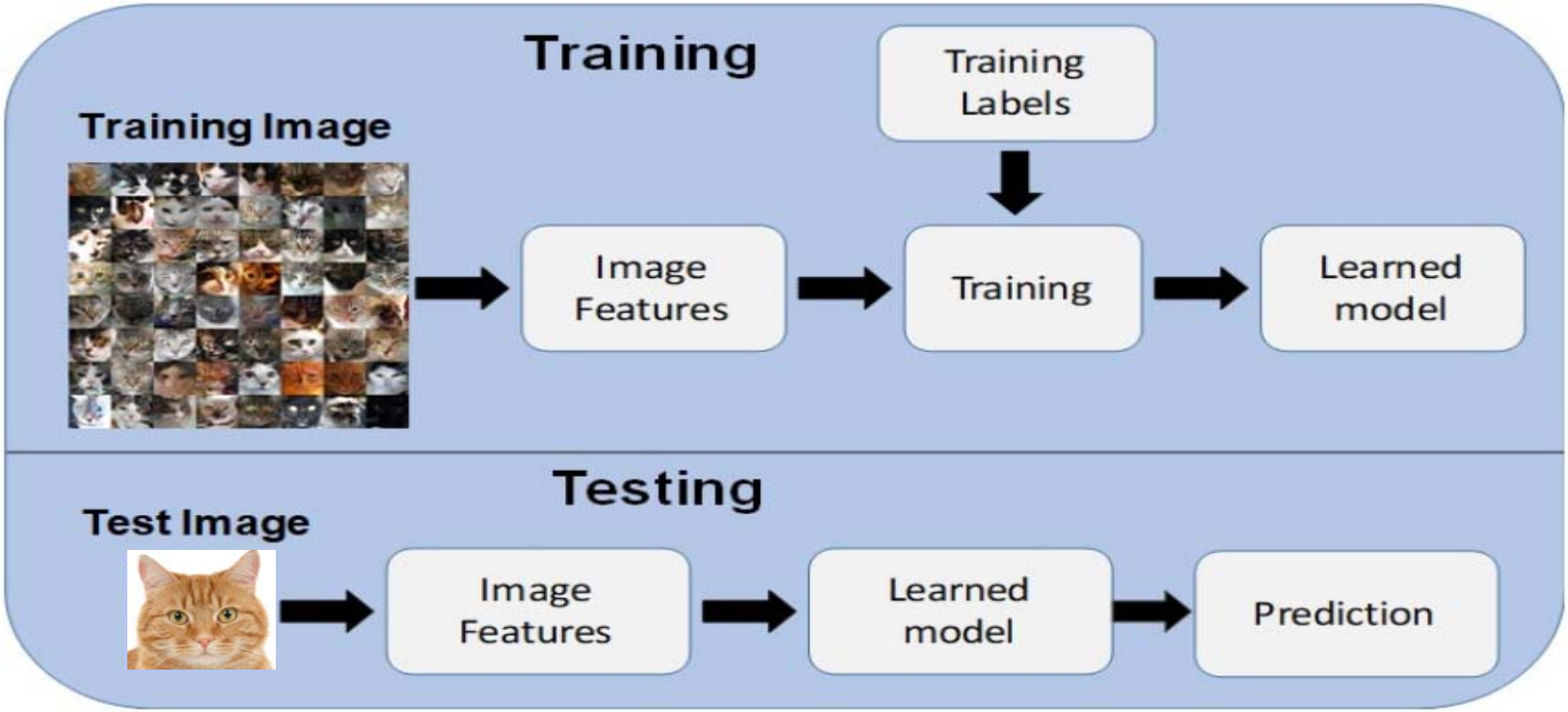


- Training: given a training set of labeled examples $(X_1, Y_1), \ldots, (X_n, Y_n)$, estimate the prediction function $f(x)$ by minimizing the prediction error on the training set.
- Testing: apply f to a never before seen test example x and output the predicted value $Y = f(x)$.

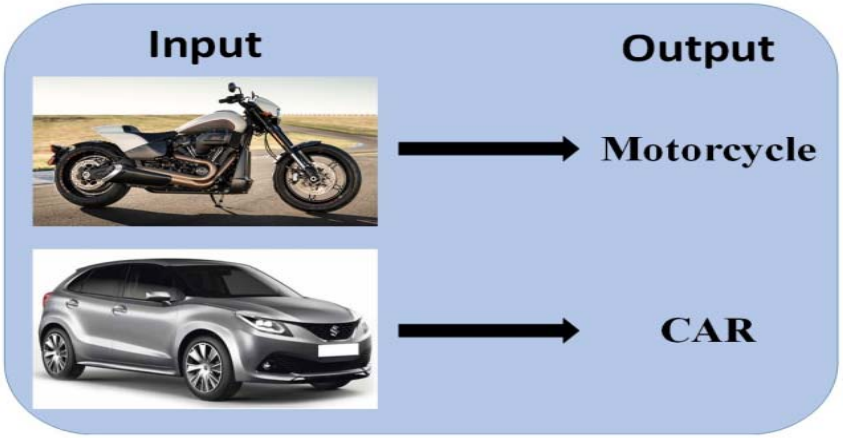
Why ML?????????

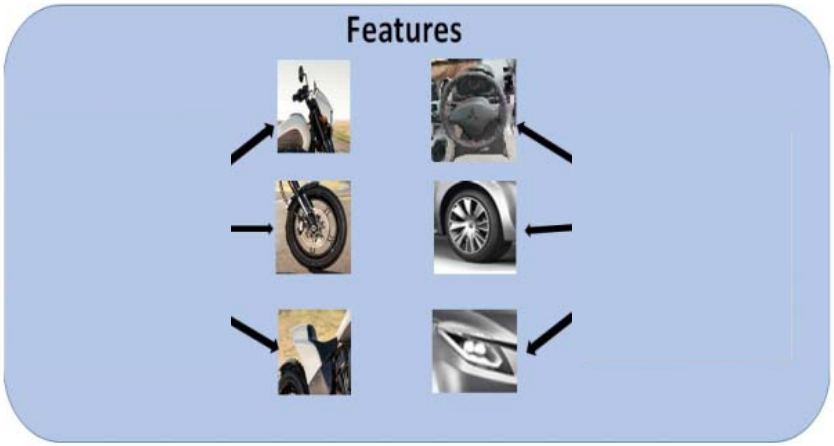
- Large amount of data available!
- Analysis of this data is not possible manually
- ML Algorithms manage the large amount of data
- Solve complex problems!

Steps of ML


Training
Training Labels
Training Image
Image Features
Training
Learned model
Testing
Test Image
Image Features
Learned model
Prediction

Features





But still classification is not easy……

Chihuahua or Muffin

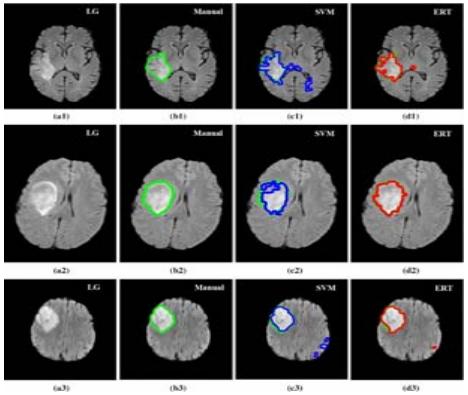
Sheepdog or Mop

Barn owl or Apple

Labradoodle or Fried-chicken

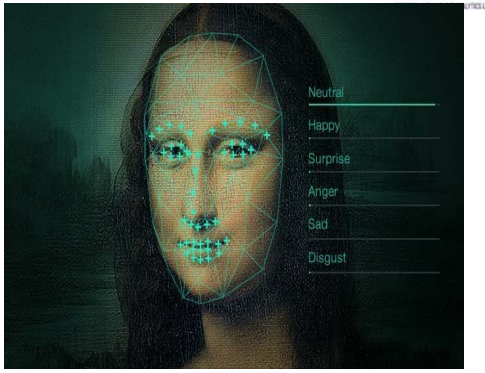
Applications

Healthcare



Disease Diagnosis

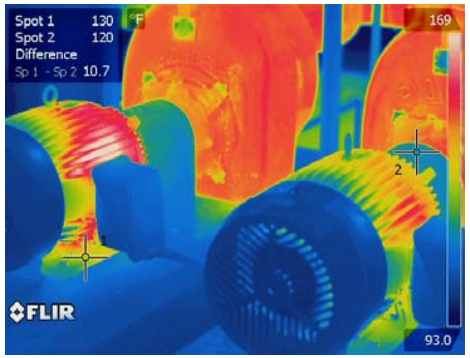
Robotics



Sentiment Analysis

Self Driving Cars

Price Prediction

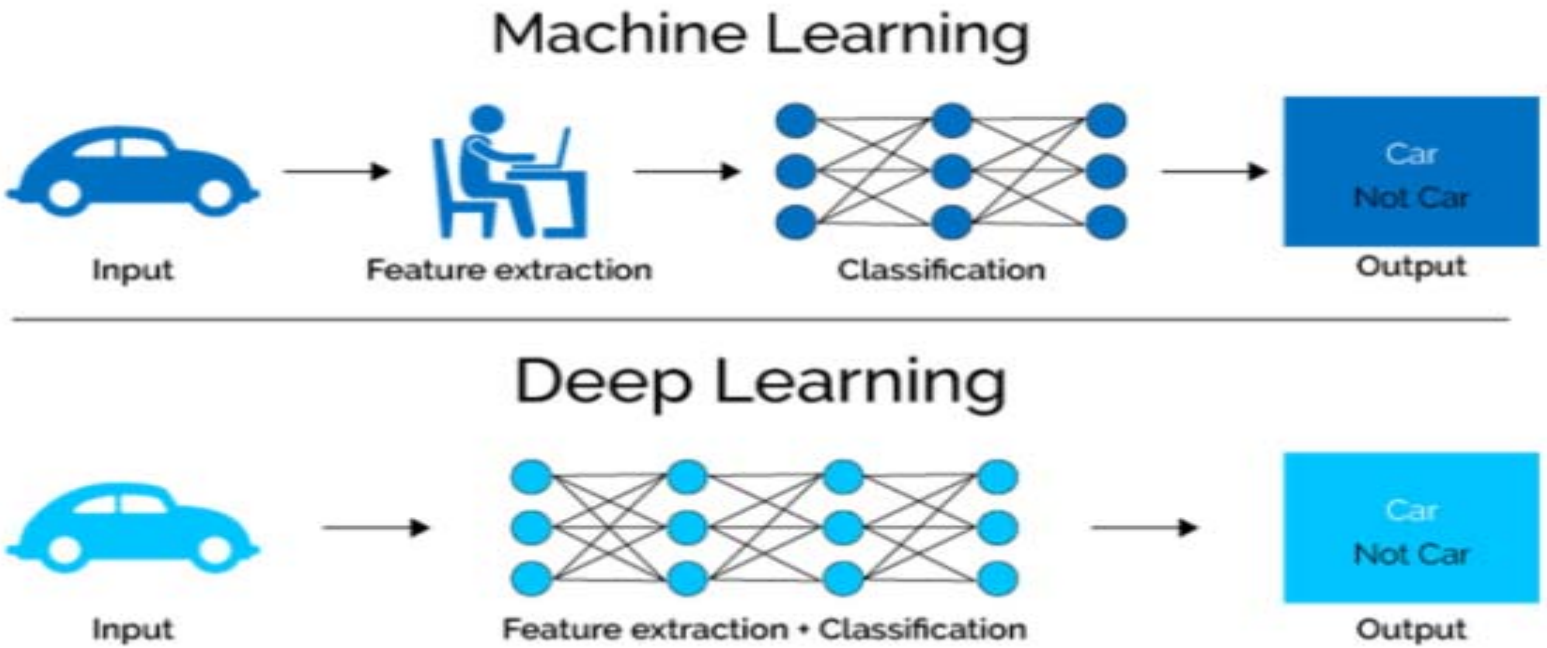


Engineering

Digital Marketing

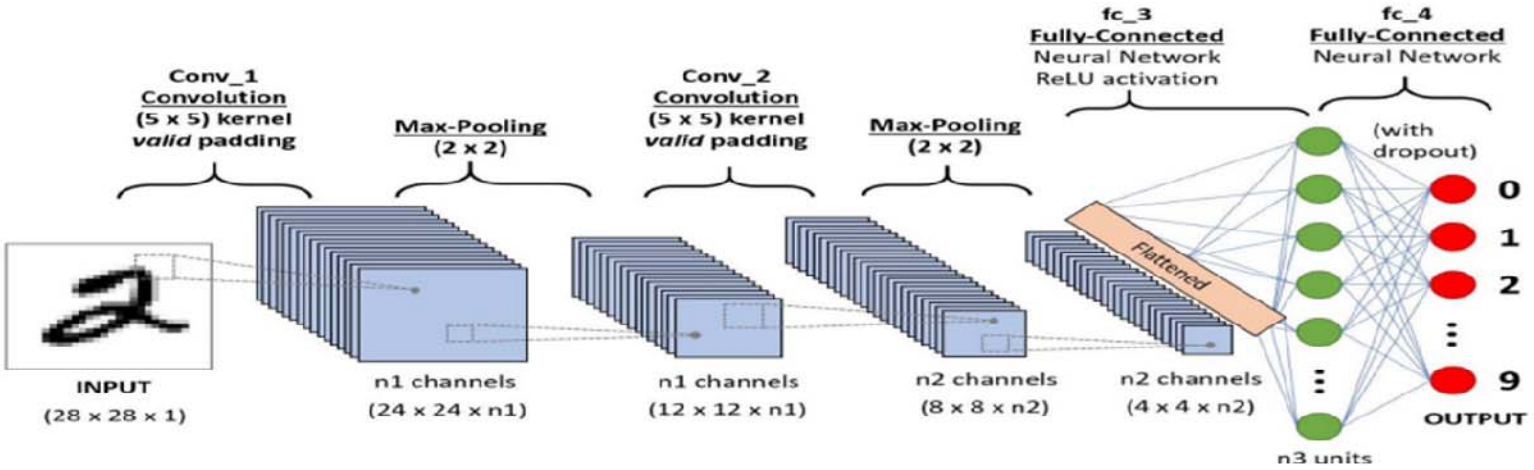
Dhirubhai Ambani University
Formerly known as DA-IICT

What does Deep Learning offer….

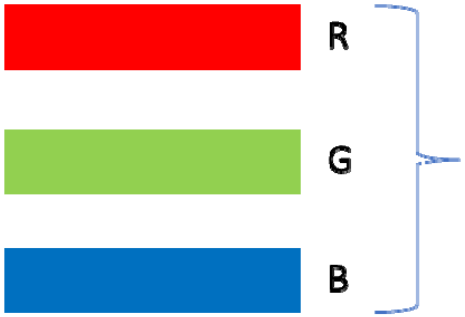


Convolutional Neural Network: Image Classification

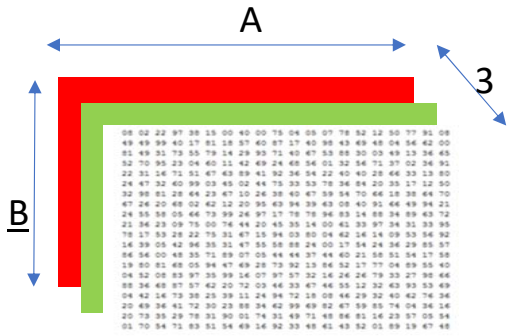
- How the Computer Reads an Image?
- What is the Fully Connected Network?
- Why not Fully Connected Networks for Image Recognition?
- What is Convolutional Neural Network?
- How do Convolutional Neural Networks work?



How computer reads an image



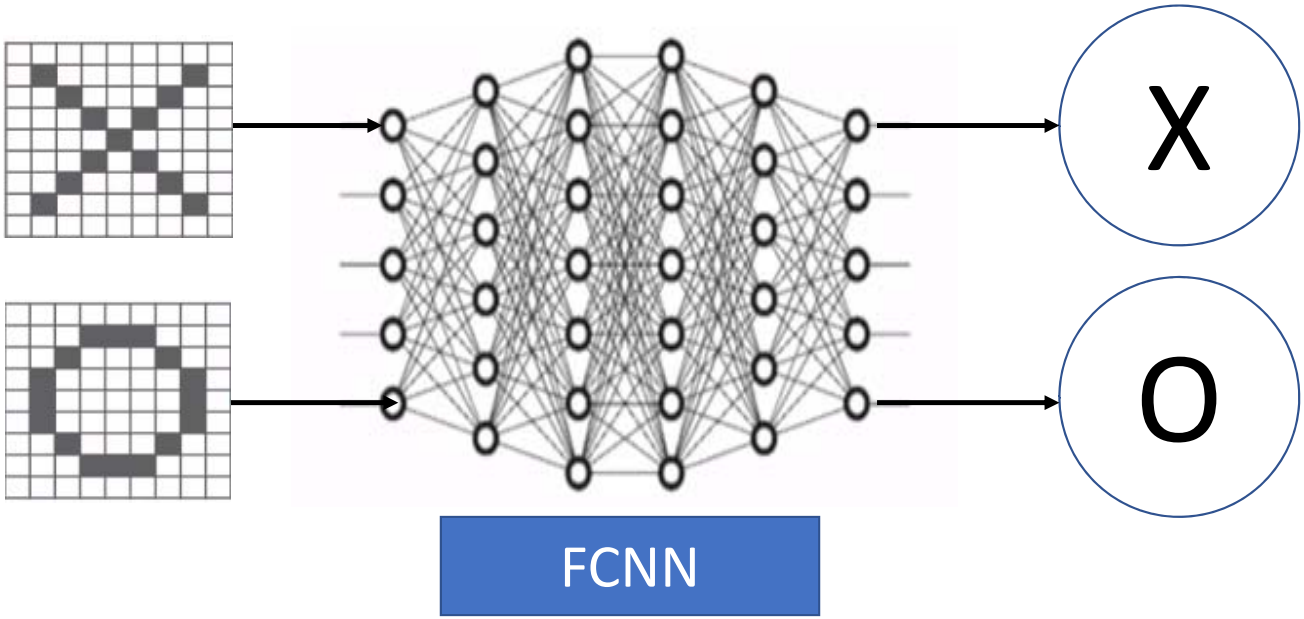
Channels



Pixels value of RGB

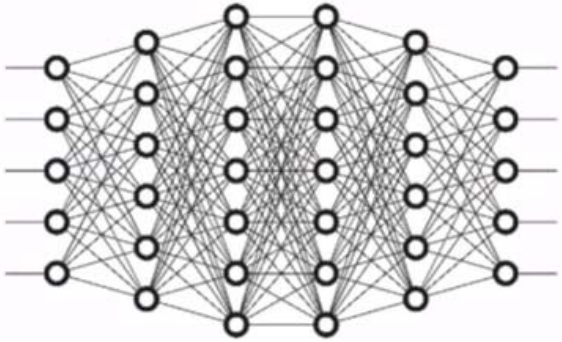
B: Rows | A: Columns | 3:

Fully Connected Neural Network



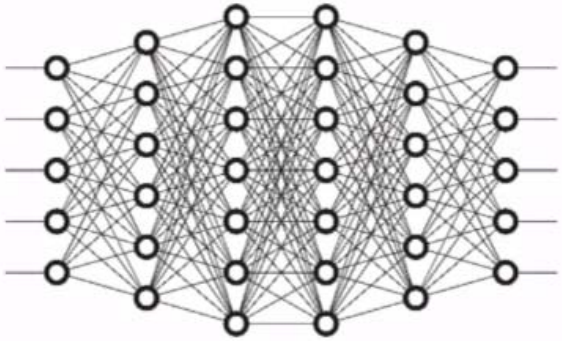
Why not fully connected networks

Image with 28 x 28 x 3 Pixels (MNIST)



Number of weights in the first hidden layer will be 2352

Image with 200 x 200 x 3 pixels



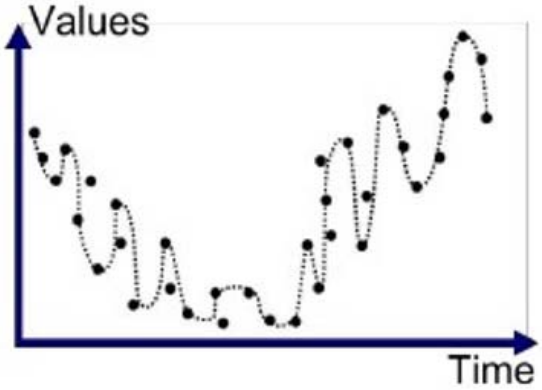
Number of weights in the first hidden layer will be 120,000

Dhirubhai Ambani University
Formerly known as DA-IICT

Why not fully connected networks

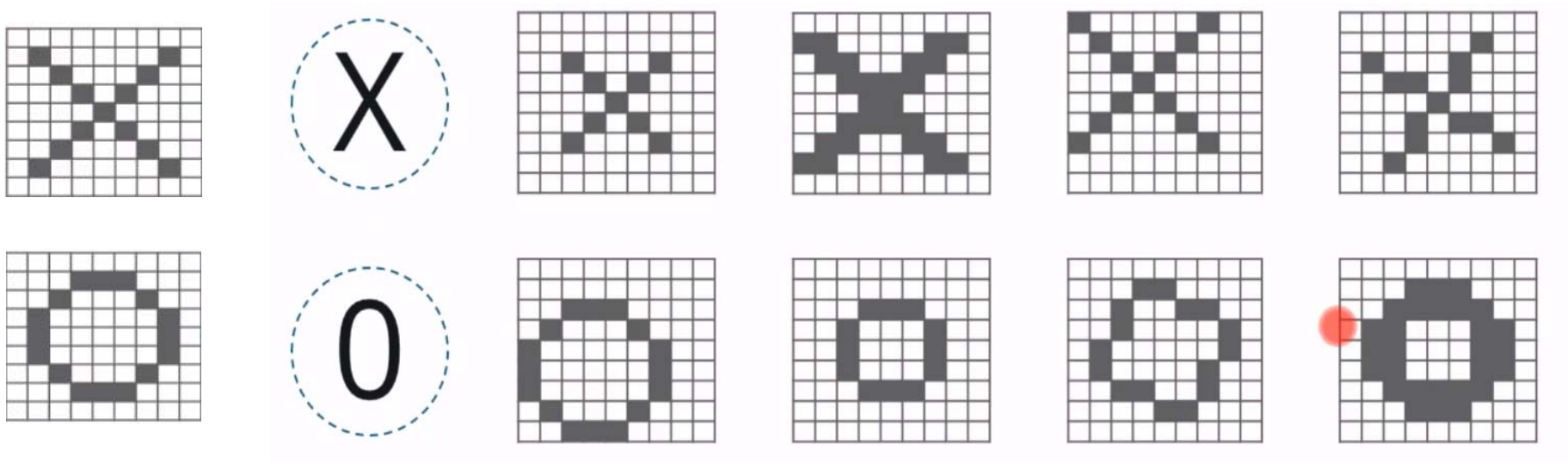
High computational Resources

Overfitting



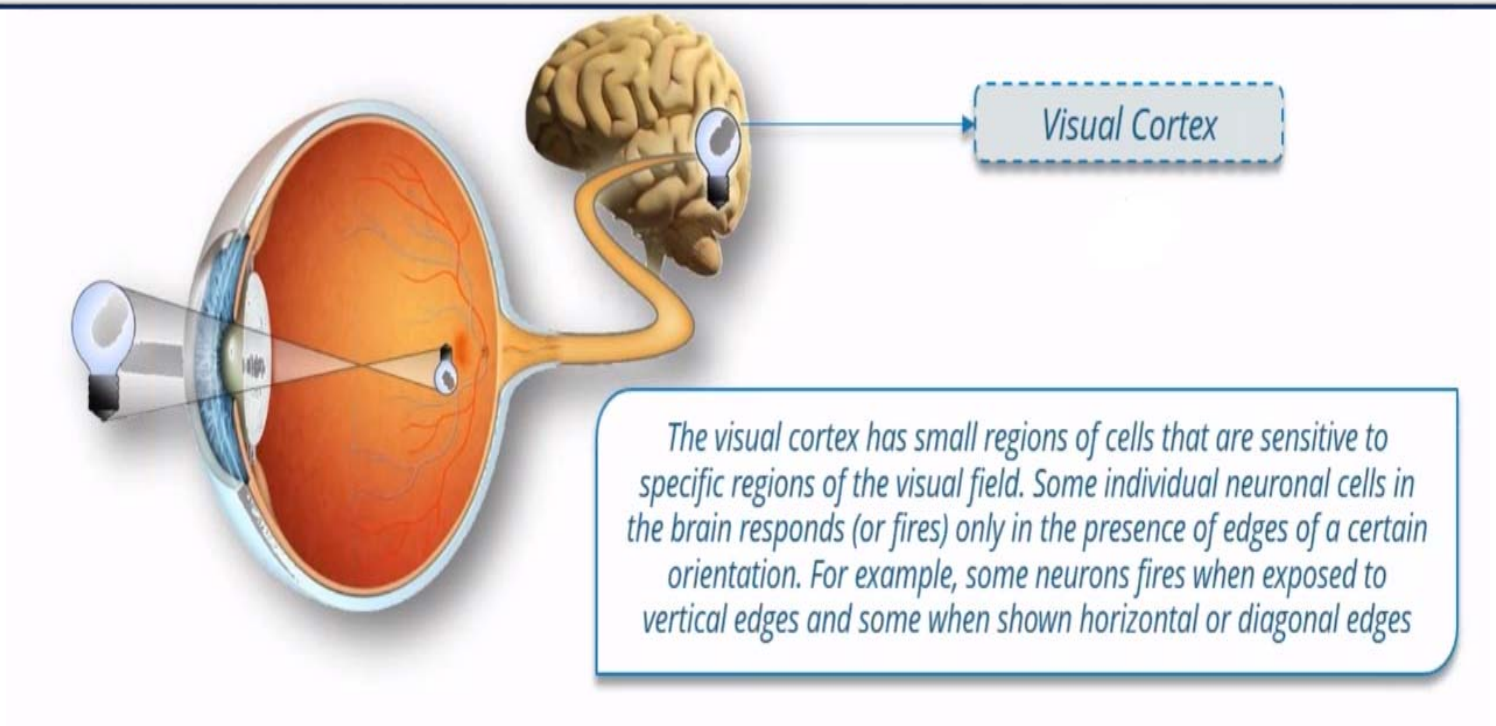
Why not fully connected networks

Here we will have some problems because X and O images won't always have the same images.
There can be certain deformations. Consider the diagram below



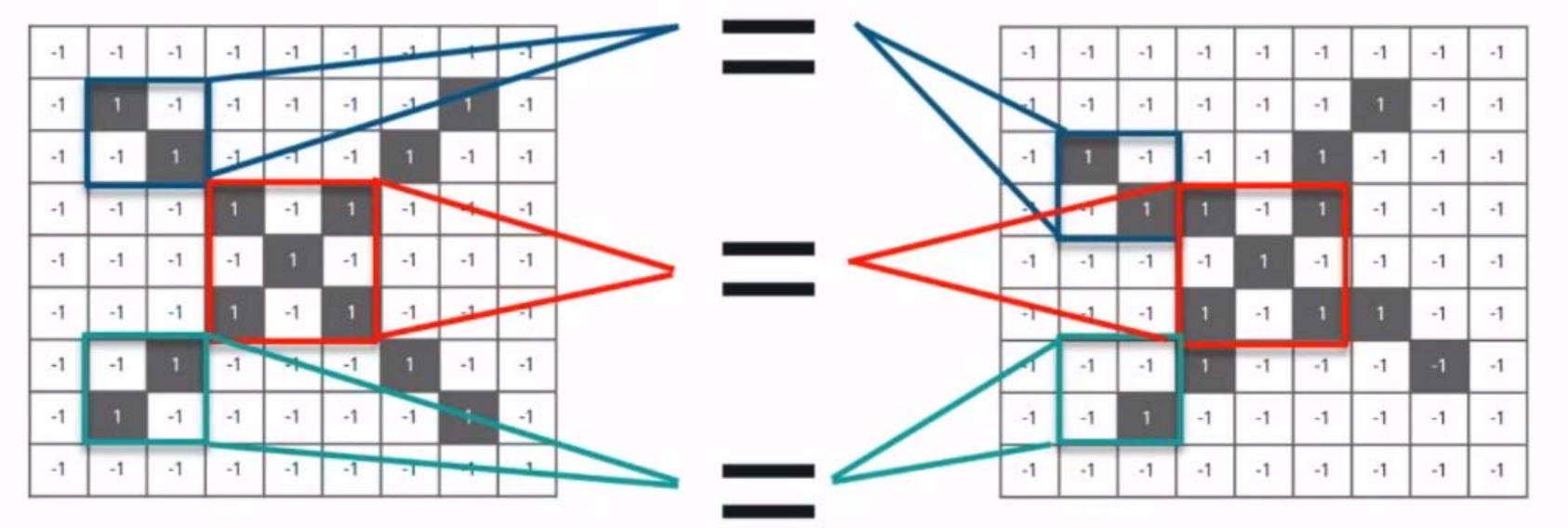
What is Convolutional Neural Network

Convolutional Neural Network (CNN) is a type of feed-forward artificial neural network in which the Connectivity pattern between its neurons is inspired by the organization of the animal visual cortex.



How CNN works?????

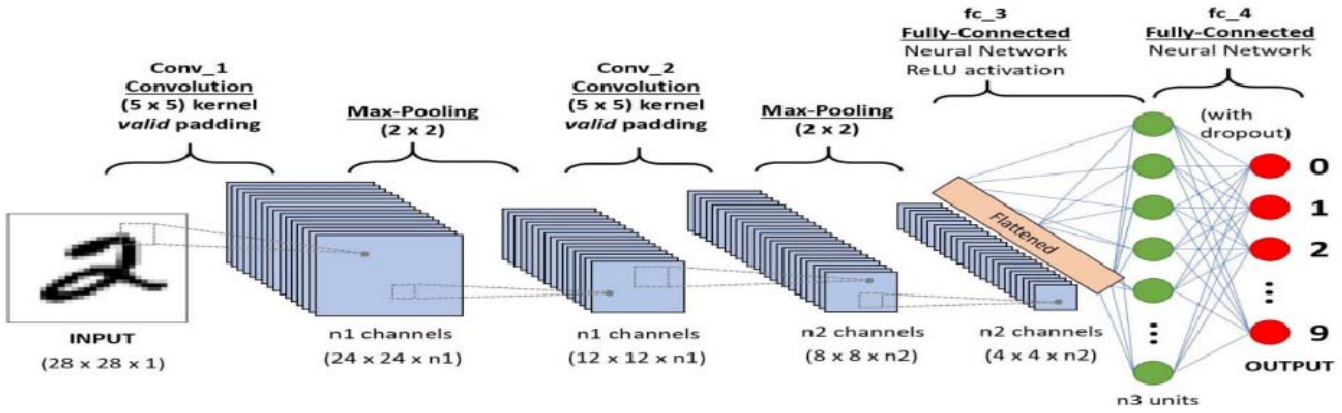
CNN compares the images piece by piece. The pieces that it looks for are called features. By finding rough feature matches in roughly the same position in two images, CNN gets a lot better at seeing similarity than the whole-image matching scheme.



Layers of CNN

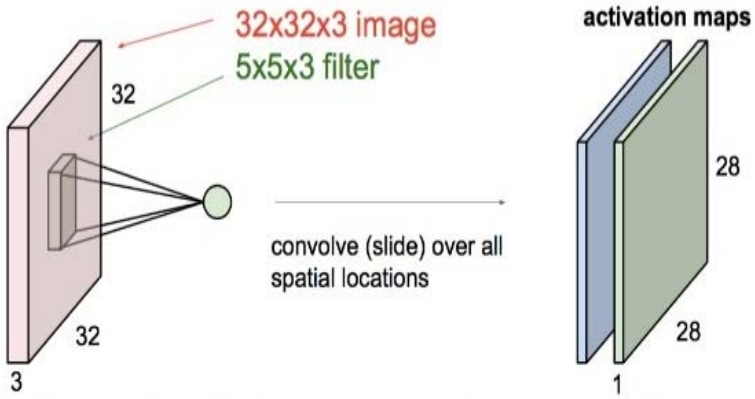
A Convolutional Neural Network consists of an input and an output layer,as well as multiple hidden layers. the hidden layers of a CNN typically consists:

- Convolutional Layer
- Pooling Layer
- Flattering
- Fully Connected Layer



Convolutional Layer

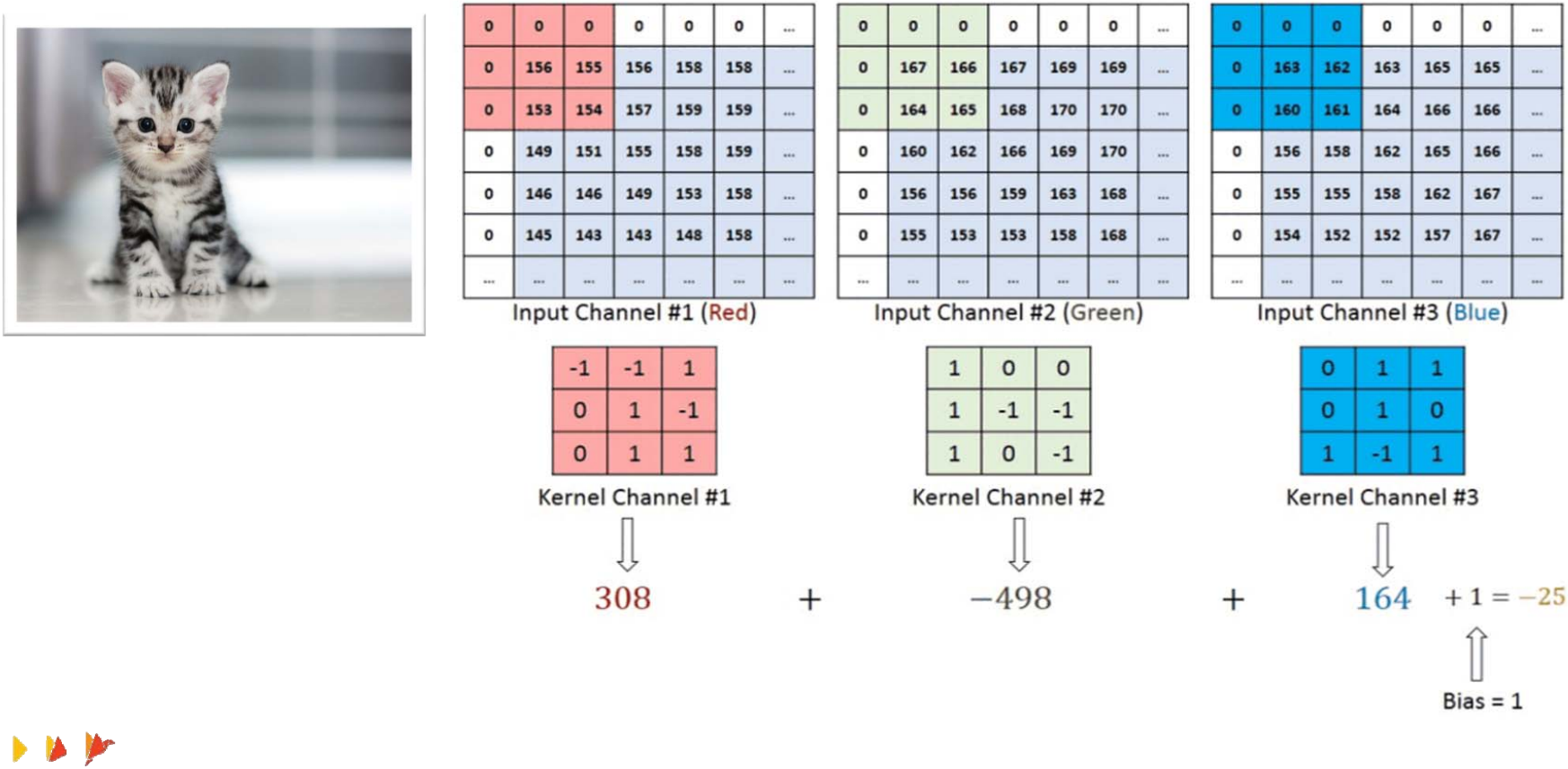
- The primary purpose of the convolution layer is to extract the features from the input image.
- They detect patterns regardless of their location in the input image.
- Reduce the number of parameters compared to fully connected layers by sharing weights (filters).
- Focus on small regions of the image, preserving spatial relationships between pixels.
- Convolutional layers can handle high-dimensional input data like large images effectively.



- Convolution (3-dim dot product) image and filter
- Stack filter in one layer (See blue and green output, called channel)

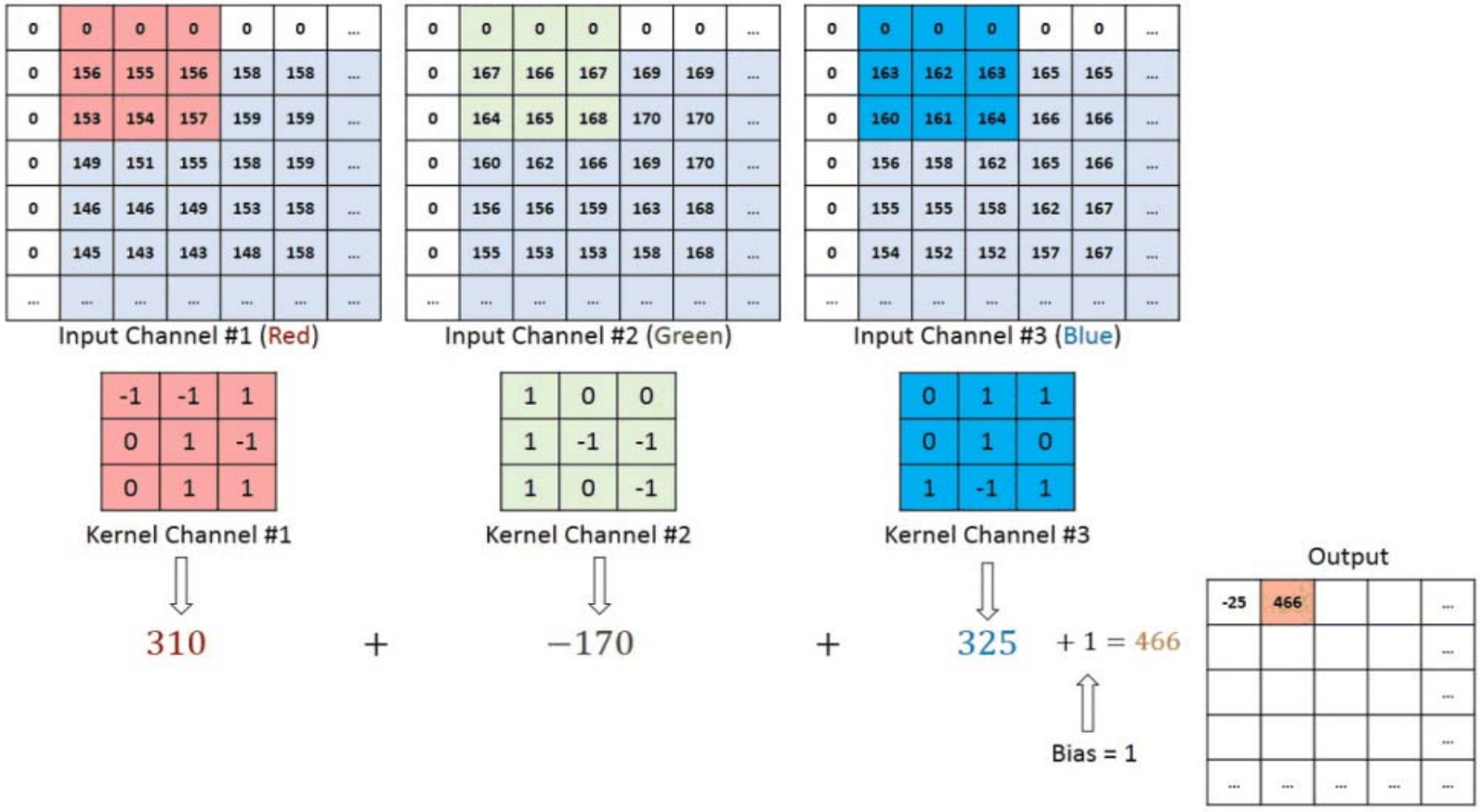
Convolutional Layer

Convolution operation on a MxNx3 image matrix with a 3x3x3 Kernel



Convolutional Layer

Convolution operation on a MxNx3 image matrix with a 3x3x3 Kernel



Problem Associated with the Convolutional Layer

There are two main problems associated with convolutional layer:

- Reducing the dimension of image: Here we use only on convolutional layer for convolution but if we pass the image through the network and get convolved by more filters then the size of the image will be reduced and may be useless.

- The edges features role is very less comparing to the other features.

Zero Padding

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

- Zero padding is a technique that allows us to preserve the original input size.
- In the zero padding we add a border of pixel all with values zero around the edges of the input images.
- When we use the zero padding then the dimension of the input size will be increase.
- After convolution, the dimension of the image will be same as the input image.

Let us take a example of input image size is 5X5. If we add zero around the edges of the input image then the input image size will be 7X7. Feature detector size is 3X3. Then the output image size will be same as input image size(5X5).

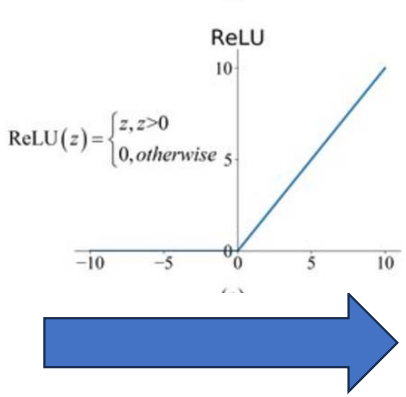
$$R = \frac{M - P}{Stride} + 1 = \frac{7 - 3}{1} + 1 = 5 \quad (1)$$

$$S = \frac{N - Q}{Stride} + 1 = \frac{7 - 3}{1} + 1 = 5 \quad (2)$$

Activation (non-linear) Function

- Replace all negative pixel values in the feature map by zero.
- The purpose of ReLu is to introduce non-linearity in CNN, since most of the real world data would be non-linear
- Other non-linear functions such as tanh (-1,1) or sigmoid (0,1), can also be used instead of ReLU (0,input).

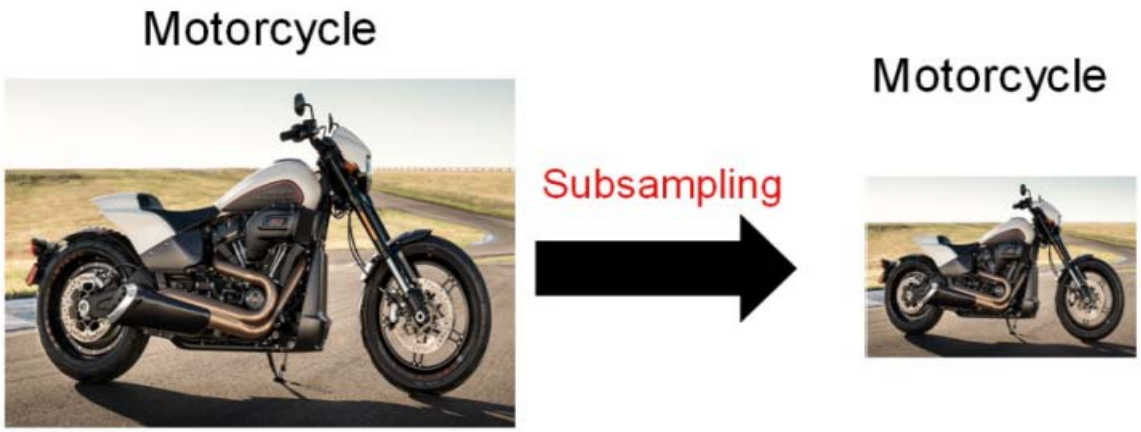
-249	-91	-37
250	-134	101
27	61	-153



0	0	0
250	0	101
27	61	0

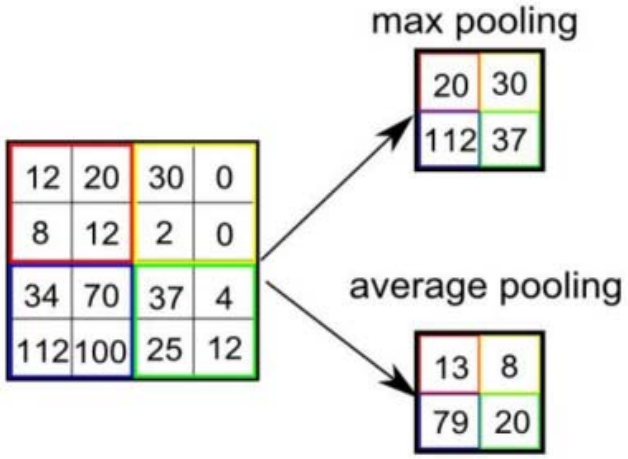
Pooling Layer

- Pooling layer function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network.
- Pooling layer operates on each feature map independently.



Subsampling pixels will not change the object.

Pooling Layer



Different Types of Pooling:

- Max Pooling
- mean Pooling
- Sum Pooling
- Min Pooling
- Median Pooling

The most common approach used in pooling is Max Pooling.

Pooling Layer

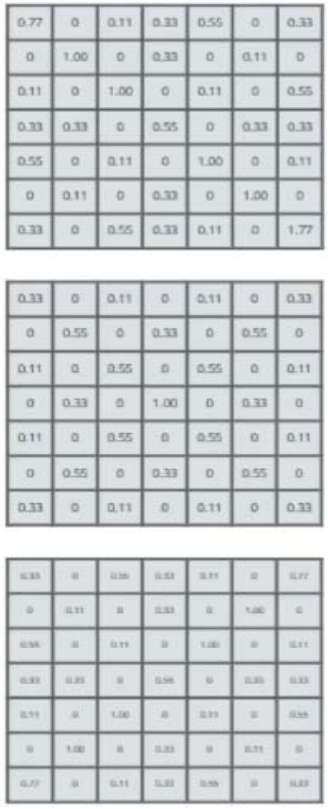
Output after passing through pooling layer.

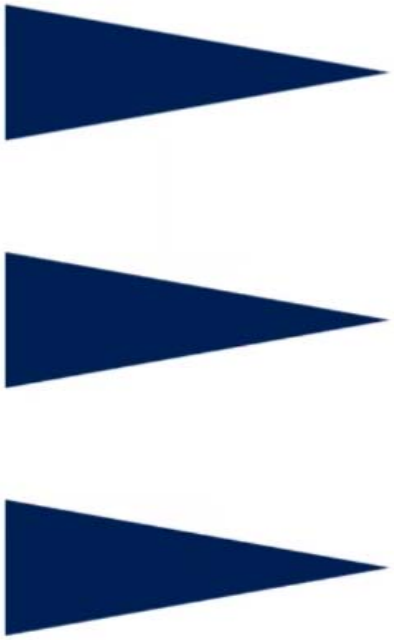
The basic role of the pooling layer is to shrink the size of our image matrix.

Here we have converted a 7x7 matrix to a 4x4 matrix.

Since we took 3 features in the beginning, we have 3 outputs after the pooling layer.

Now next we have to stack up all the layers.





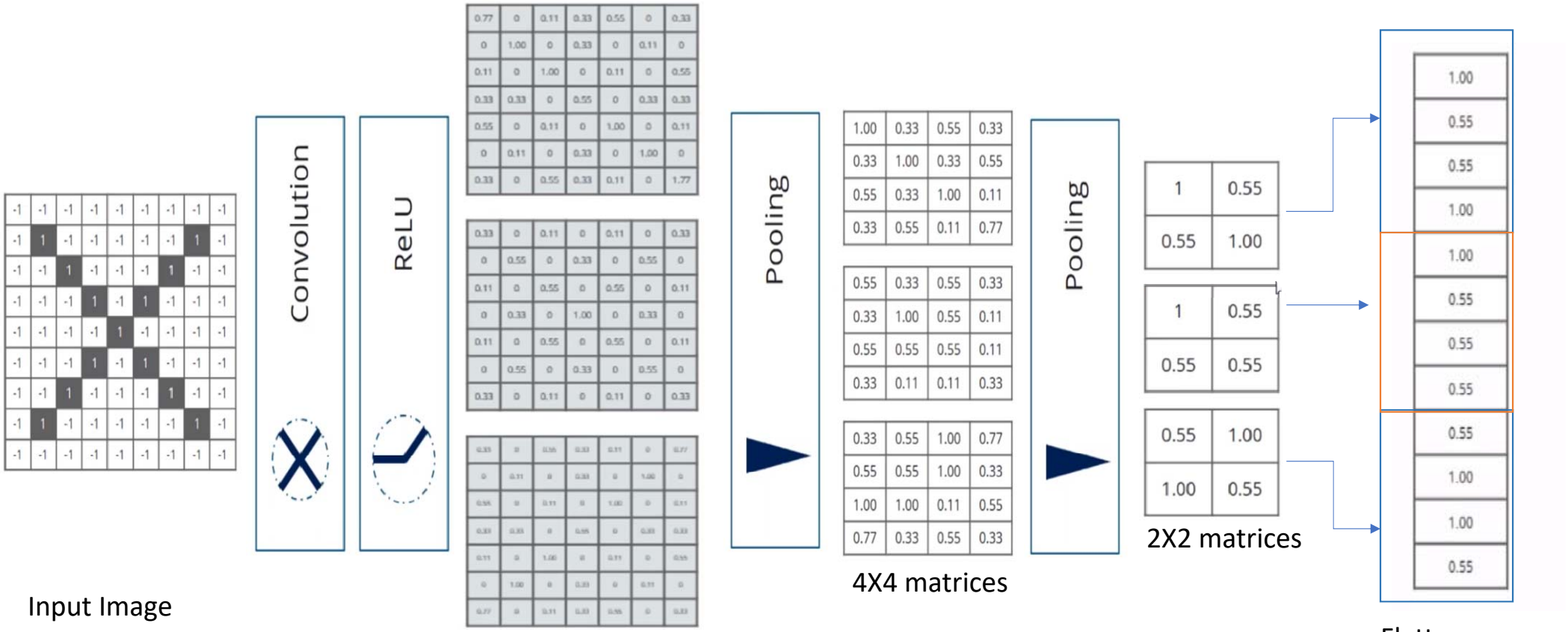
1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

0.55	0.33	0.55	0.33
0.33	1.00	0.55	0.11
0.55	0.55	0.55	0.11
0.33	0.11	0.11	0.33

0.33	0.55	1.00	0.77
0.55	0.55	1.00	0.33
1.00	1.00	0.11	0.55
0.77	0.33	0.55	0.33

Dhirubhai Ambani University
Formerly known as DA-IICT

Stacking up the layers



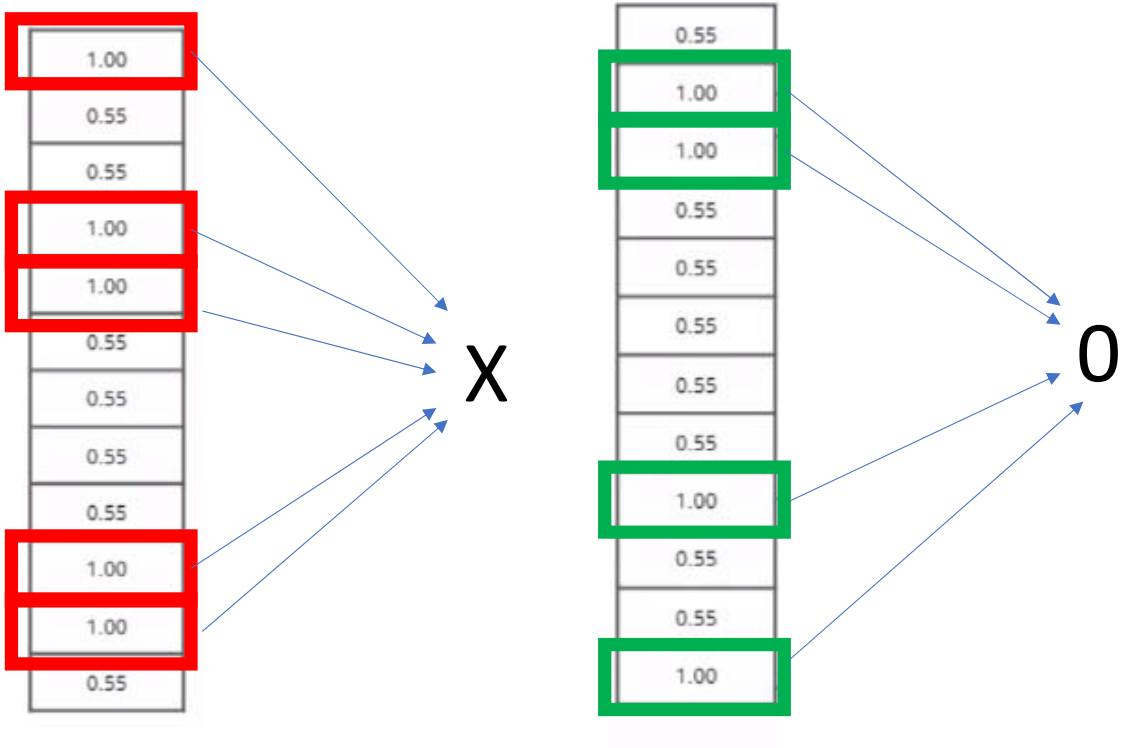
Output

When we feed in, ‘X’ and ‘O’. Then there will be some element that will be high.

Now if we have an input image with 1^{st}, 4^{th},5^{th},10^{th},11^{th} value high we can say that the image is X

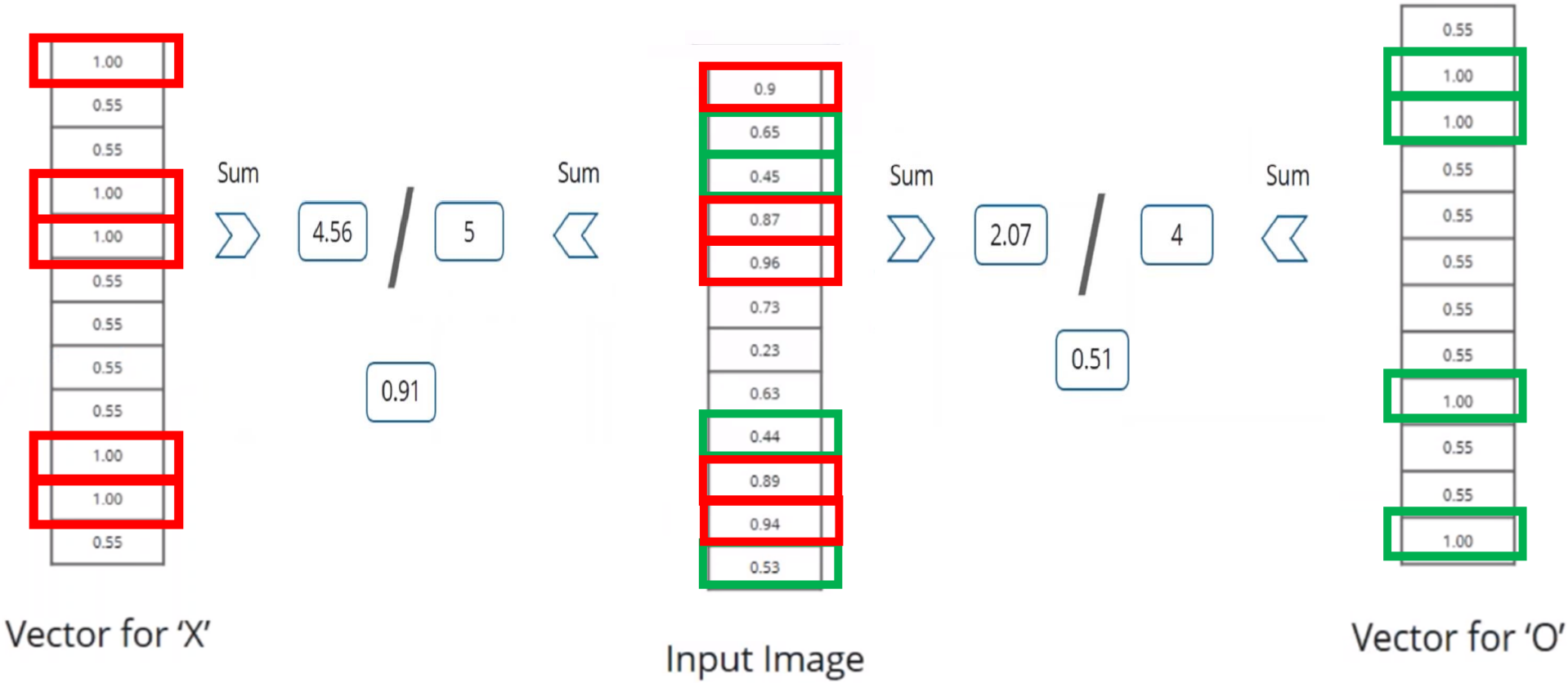
Similarly, if we have an image with 2^{nd},3^{rd},9^{th} and 12^{th} value high, we can say that it is O

This completes the training of our model. Now let’s see the results on some unseen images.



Compare the input vector with X and O

Let's compare with the list of 'X' and 'O'



Result

The new input image is classified as 'X'