CNN – Introduction & Operations

Have you ever wondered how....

→ Facebook recognize your face?

☺ → Ms. Smiley

→ Tesla Cars detect objects?

🍎 → Apple (Rotten)

It's all possible, thanks to

Convolutional Neural Networks (CNN)

- Himanshu Goswami

Just like how a child learns to recognize objects, we need to show an algorithm, millions of pictures before it can generalize the inputs & make predictions for image it has never seen before.

Computer's see in a different way than we do. Their world consist of only numbers (no colors!). Every image can be represented as 2-dimensional array of numbers, known as pixels.

How Humans see Dog

How Computer see Dog

But the fact that they perceive image differently, doesn't mean we can't train them to recognize patterns as we do.

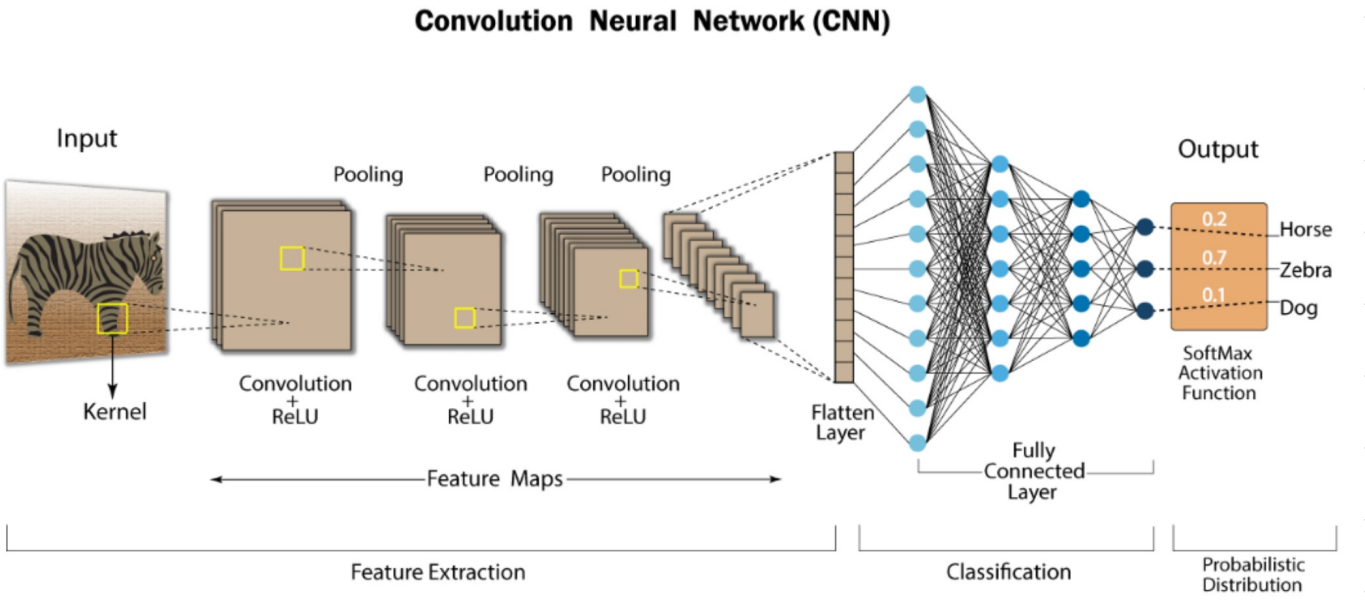
To teach an algorithm on how to recognize objects in images, we use a specific type of Artificial Neural Network : Convolutional Neural Network.

CNN take a biological inspiration from the Visual Cortex present in our Brain.

- Himanshu Goswami

Operations in CNN

Let's take a brief look at this pic... (Don't worry if it looks overwhelming)



Now we will see each component / layer step-by-step & everything will make sense ☺

① Filter a.k.a. Kernel / Feature Detector

Purpose of filter is to detect edges from the pic. A filter is a matrix of values, called weights, that are trained to detect specific features.

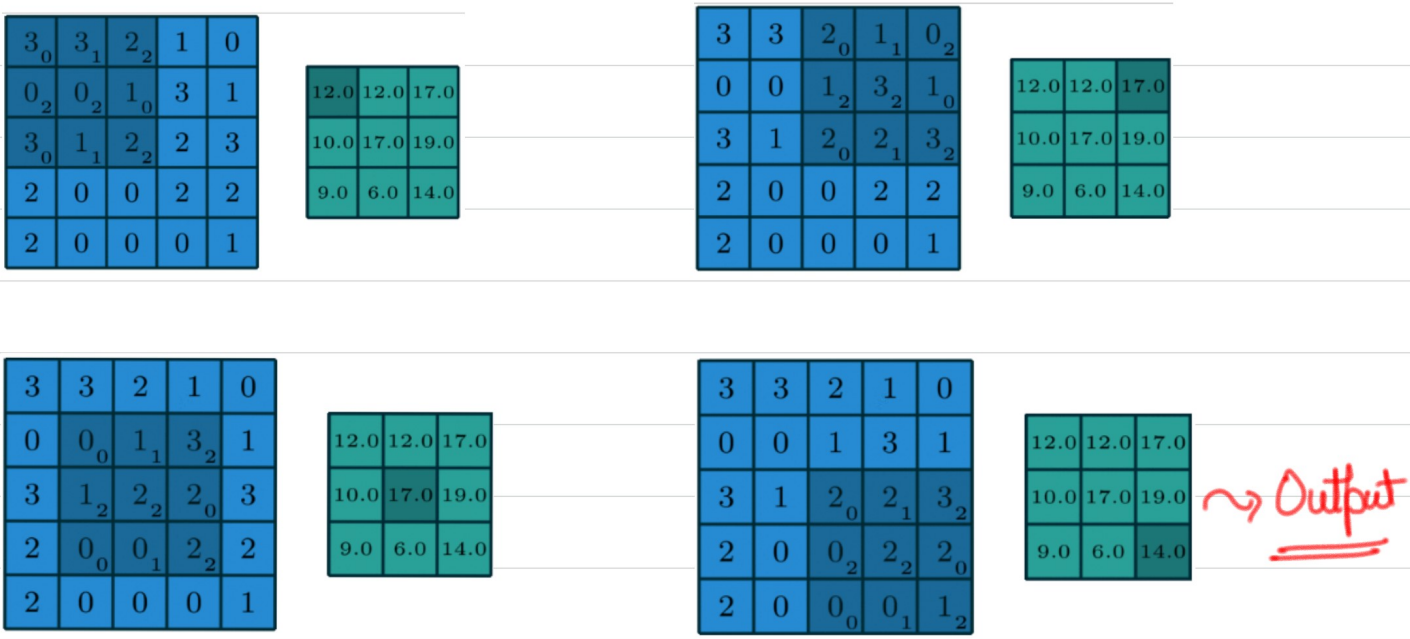
The filter moves over each part of image & carries out Convolutional operation.

- Himanshu Goswami

Let say our image & Kernel are :-

$$\begin{bmatrix} 3 & 3 & 2 & 1 & 0 \\ 0 & 0 & 1 & 3 & 1 \\ 3 & 1 & 2 & 2 & 3 \\ 2 & 0 & 0 & 2 & 2 \\ 2 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 \\ 2 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

Element Wise Product & Sum = Output



But there occurs a problem of losing information / pixel. Since we typically use small kernels, for any convolution, we might lose a few pixels.

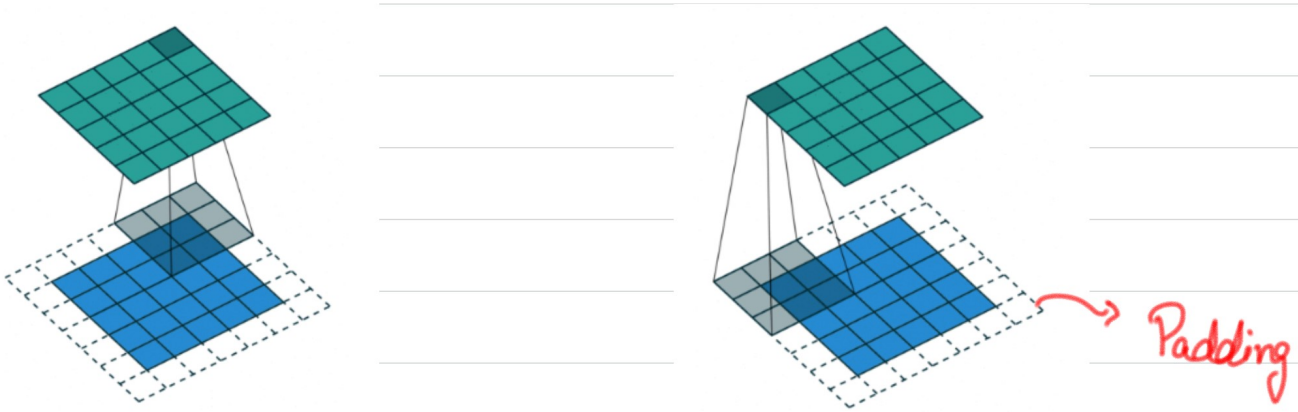
↳ This Problem adds up when we apply many convolutional layers.

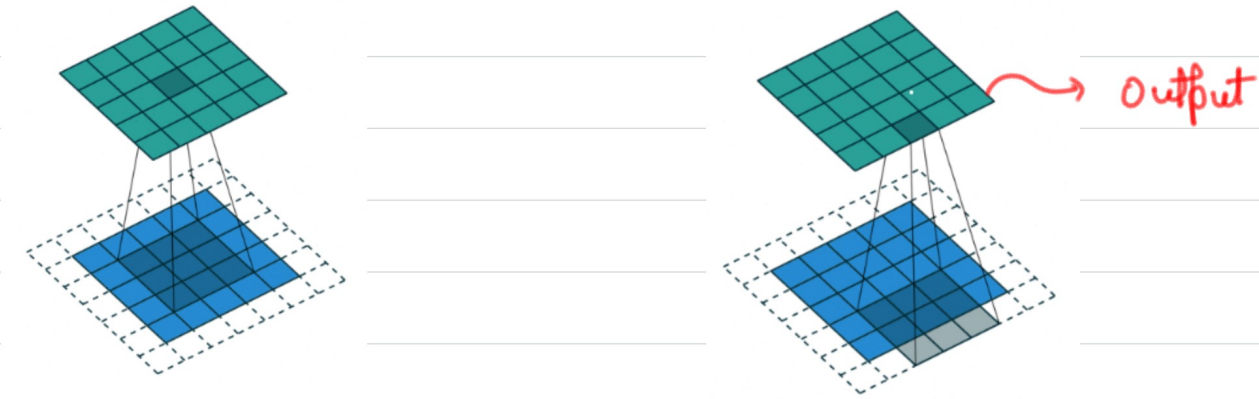
↳ This is where comes a component named PADDING which solves this issue.

- Himanshu Goswami

2 PADDING

Padding adds another boundary in image for the filter to prevent information loss.





This time it didn't reduce the dimension of output. It's exactly the same like Input because it didn't lose any information / pixel this time.

Those "Fake" pixels hold zero value that prevents it from altering real image.

- Himanshu Goswami

Till here it's known as CONVOLUTION!

Now since our filter is a weight which isn't fixed & need to be updated for better result, we need to apply Activation Function on our convolution.

(3) Activation Function Mapping

We apply ReLU activation function to convolution so that during Backward Propogation our filter can update.

— Our First Layer : Convolutional Layer — Completes here

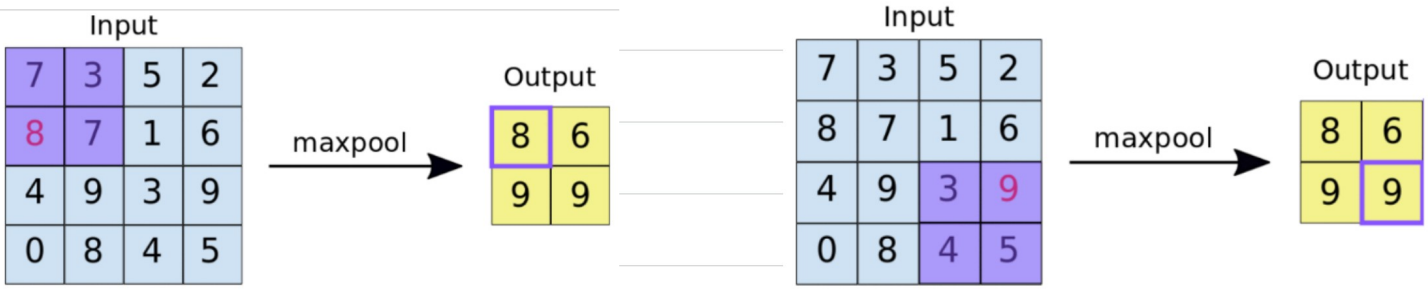
Now on top of Convolutional layer, we apply POOLING technique through which we try to extract information from the layer.

(4) Pooling

- Max Pooling
- Min Pooling
- Mean Pooling

- Himanshu Goswami

On the basis of use case we pick Pooling technique. Generally we pick Max - Pooling which extracts most information from the convolution.



This is how only max information gets picked.

Since there are number of filters in each convolution layer, Pooling is applied over each layer.

↳ There is information loss here.

We repeat Convolution layer + ReLU + Pooling till we get desired result.

↳ As we saw earlier, pooling leads to information loss, so...

↳ Deeper the Network, more we will apply Pooling & More information we lose.

↳ Smaller Output Layer

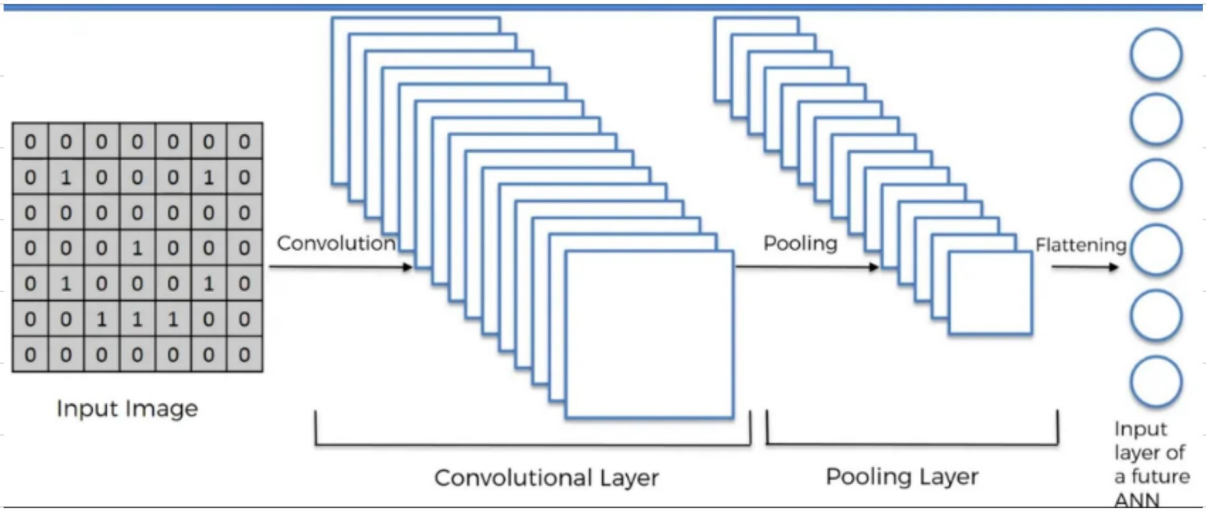
- Himanshu Goswami

Now whatever the smaller size output we are left with, we will FLATTEN it into 1D array

↳ It will become Input Layer of our ANN.

(5) Flattening

We will flatten the entire layer of filters to be used as input for ANN.



As our input layer for ANN is defined, we can set the number of layers of our Neural Network.

↳ These layers are called Fully Connected layer.

↳ We get Classification result as output.

- Himanshu Goswami

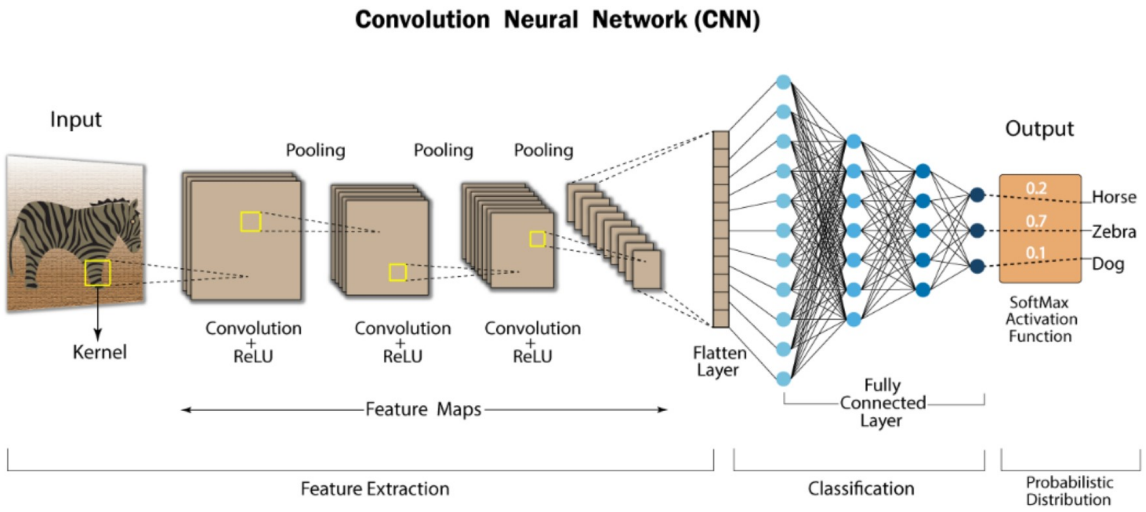
This completes FORWARD PROPAGATION.

↳ Any error will result in BACKWARD PROPAGATION

↳ Update Filters, Weights, Bias.

↳ Repeat until EPOCHS we defined.

Coming Back to the pic, which I hope doesn't look overwhelming anymore....



Thanks for reading ^^

Let's connect...

- Himanshu Goswami