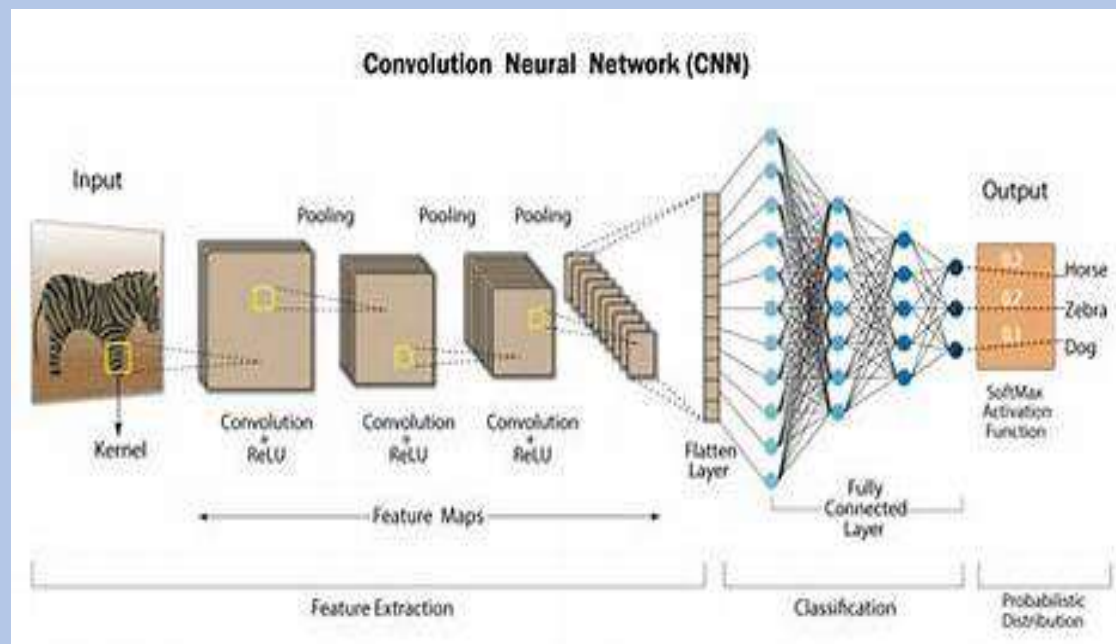


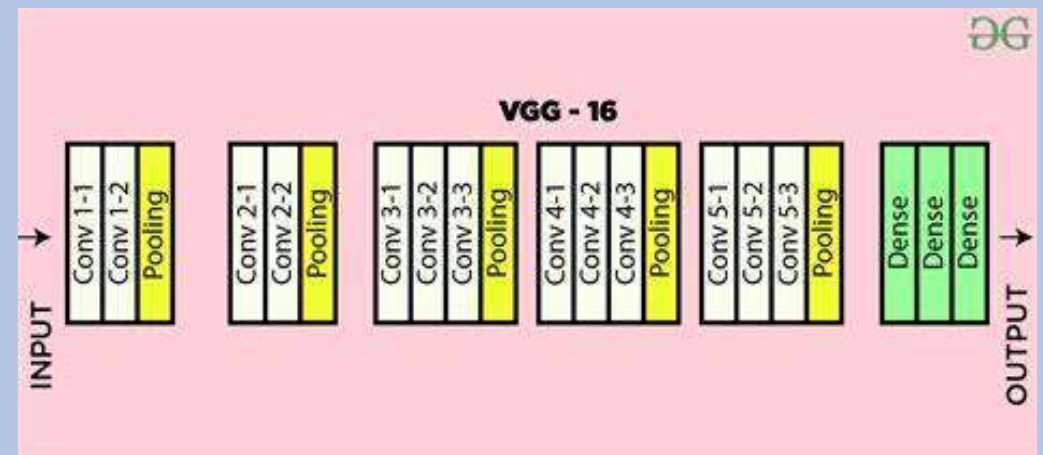
A STUDY OF LAYERS IN CNN



WE WOULD CONSIDER VGG-16 FOR REFERNECE TO STUDY ABOUT LAYERS

VGG-16

- Proposed by Visual Geometry Group (VGG) at University Of Oxford
- Consists of 16 layers- i) 13 Convolutional Layers
- ii) 3 Fully connected layers
- Five MaxPooling layers are also present



```
graph TD; A[ARCHITECTURE] --> B[FEATURE EXTRACTOR]; A --> C[CLASSIFIER]
```

ARCHITECTURE

FEATURE
EXTRACTOR

CLASSIFIER

STRUCTURE OF VGG-16

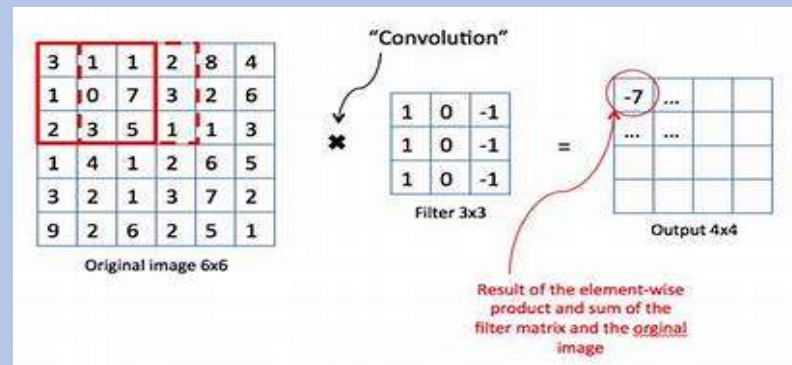
- The input taken is a color image of shape: $224 \times 224 \times 3$ where 3 denotes that the input images are colored (RGB) in nature
- The input layers are processed by taking filters and performing convolution function across it to produce an output
- We can have a quick understanding of the convolutional operation from-
<https://deeplizard.com/resource/pavq7noze2->

- **STRIDE**

Determines the number of pixels for which the kernels shift over the input at a time. When the stride length is decreased, more features can be learned by the network

- **PADDING**

Helps to preserve the same dimensions as the input. This helps to increase the dimensions of the feature map(generated as a result of convolution operation on input and filter layer, and applying activation function)



- There are several padding techniques- Same padding, Valid padding, Zero padding
- Benefits of padding (using padding is not compulsory)-
 - i. Preserve Spatial Dimensions
 - ii. Control Output Size- 'same' padding ensures the output has the same dimensions as the input, while 'valid' padding produces a smaller output

- **OUTPUT CALCULATION :** $O = [(n-f+2p)/s] + 1$ where
 O=Output size
 n= Input size
 f= filter size
 p= padding
 s= Stride

- FILTERS AND THEIR POWER TO DETERMINES EDGES

Prewitt Filters-The Prewitt operator is comprised of two filters that help to detect vertical and horizontal edges. The horizontal (x-direction) filter helps to detect edges in the image which cut perpendicularly through the horizontal axis and vice versa for the vertical (y-direction) filter

Sobel Filters-Sobel operator is also made up of a vertical and horizontal edge detection filter. Detected edges are quite similar to results obtained using Prewitt filters but with a distinction of higher edge pixel intensity

Robinson Compass Masks-The Robinson Compass masks are edge detection filters which are made up of 8 different filters accounting for the 8 geographical compass directions as shown in the image above

Feature Maps: When a kernel convolves with an input image, it scans through it systematically. At each position, the dot product of the kernel and the local region of the image is computed. This produces a single value in the output feature map, which represents the presence of a particular feature at that location.

Edge Detection: Kernels are excellent at detecting edges in images. A kernel designed for edge detection will highlight the transitions between regions of differing intensities, such as from light to dark. This feature is a building block for recognizing more complex patterns

Hierarchical Layers: CNNs typically consist of multiple convolutional layers stacked on top of each other. Each layer uses different kernels, and these layers progressively learn higher-level features.

Examples of Hierarchical Learning:

The first convolutional layer may detect simple features like edges and textures.

The second layer builds upon these to detect more complex features like shapes or parts of objects.

Subsequent layers recognize even more complex objects or patterns.

The below article shows how layers are detected layer by layer

[Visualize Features of a Convolutional Neural Network - MATLAB & Simulink
\(mathworks.com\)](https://www.mathworks.com/help/deeplearning/visualize-features-of-a-convolutional-neural-network.html)