Types of Convolution

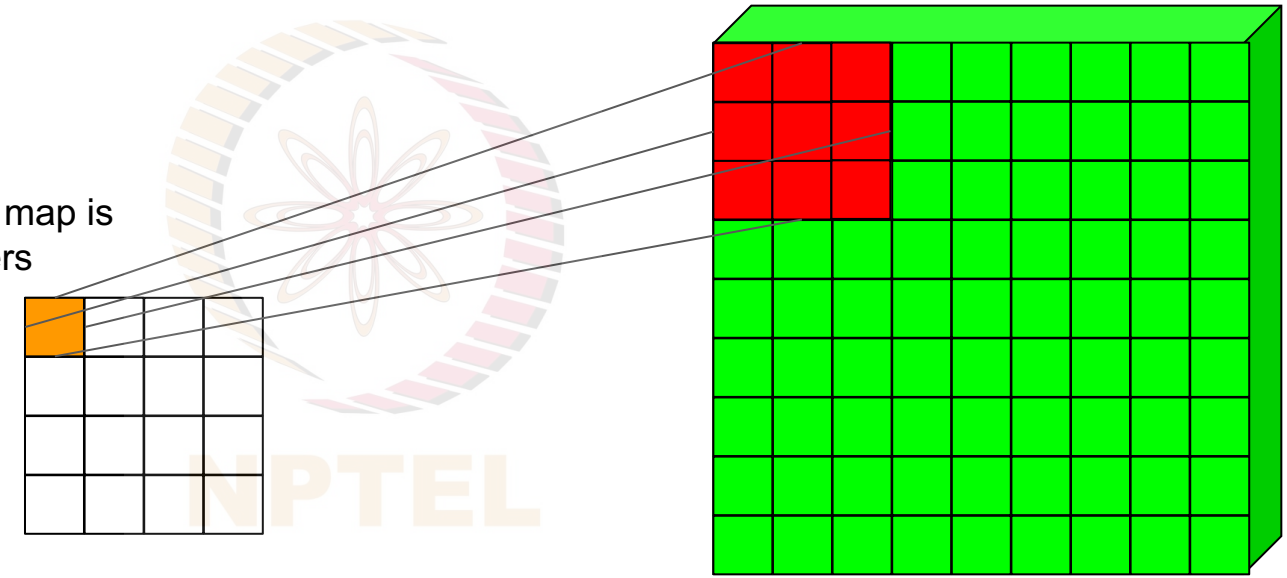
1. Convolution
2. Dilated Convolution
3. Transpose Convolution

Naive Convolution

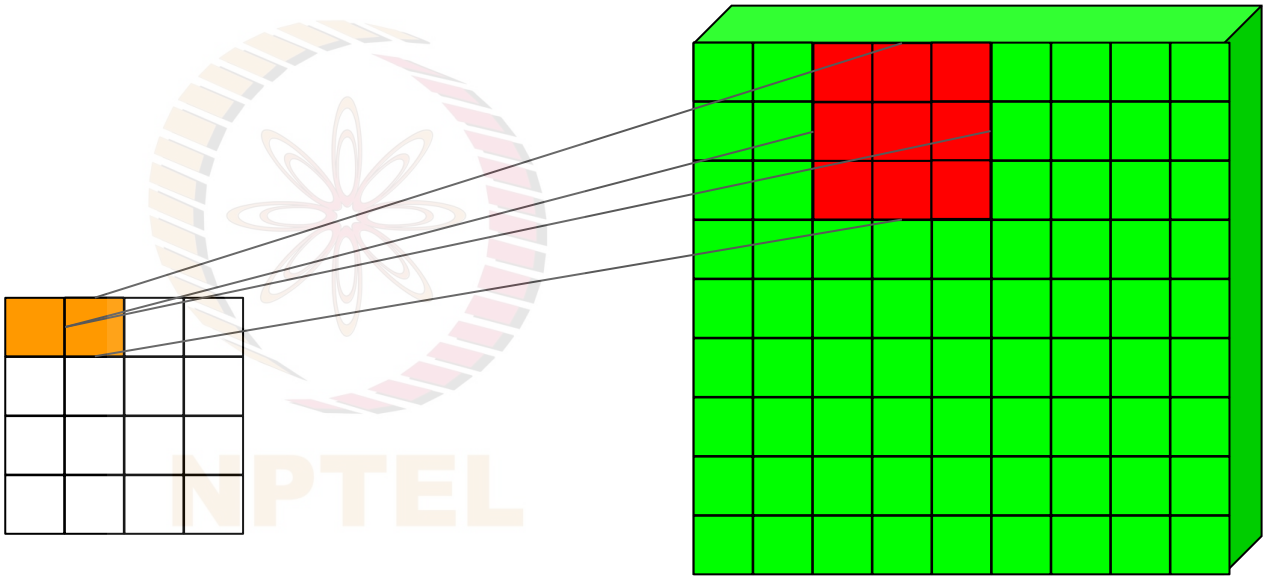
1. Building block of convolution neural network
2. Parameters
   a. Kernel size (Kw, Kh)
   b. Stride (Sw, Sh)
   c. Padding (P)
3. Dimension of Output feature map is governed by above parameters

$$O_w = \frac{I_w - K_w + 2P}{S_w} + 1$$

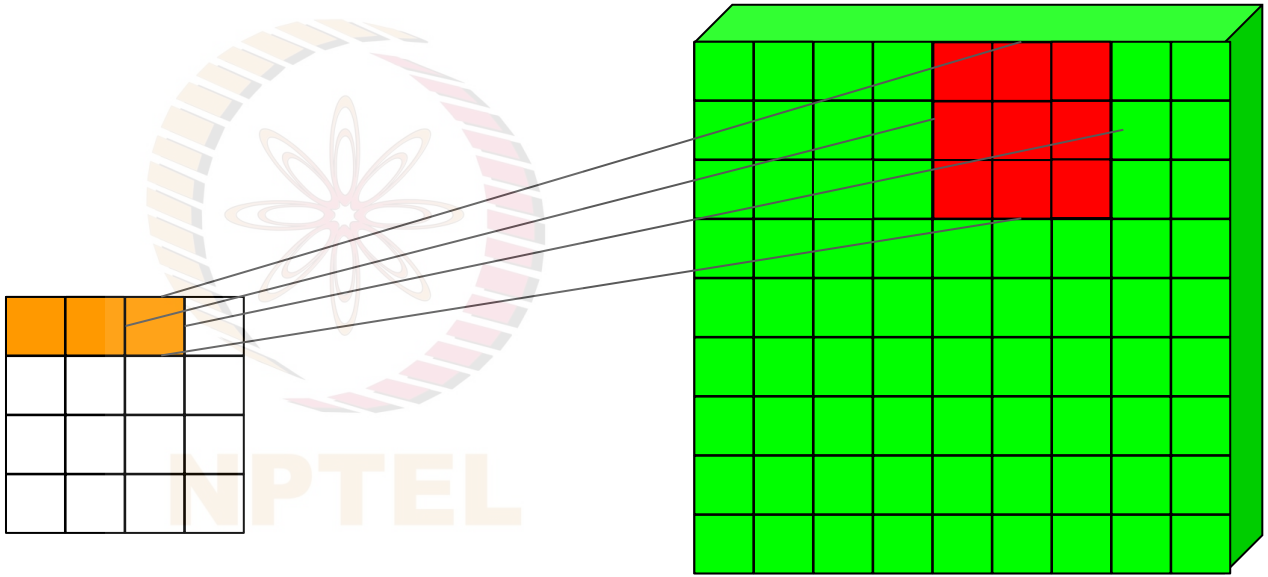
$$O_h = \frac{I_h - K_h + 2P}{S_h} + 1$$



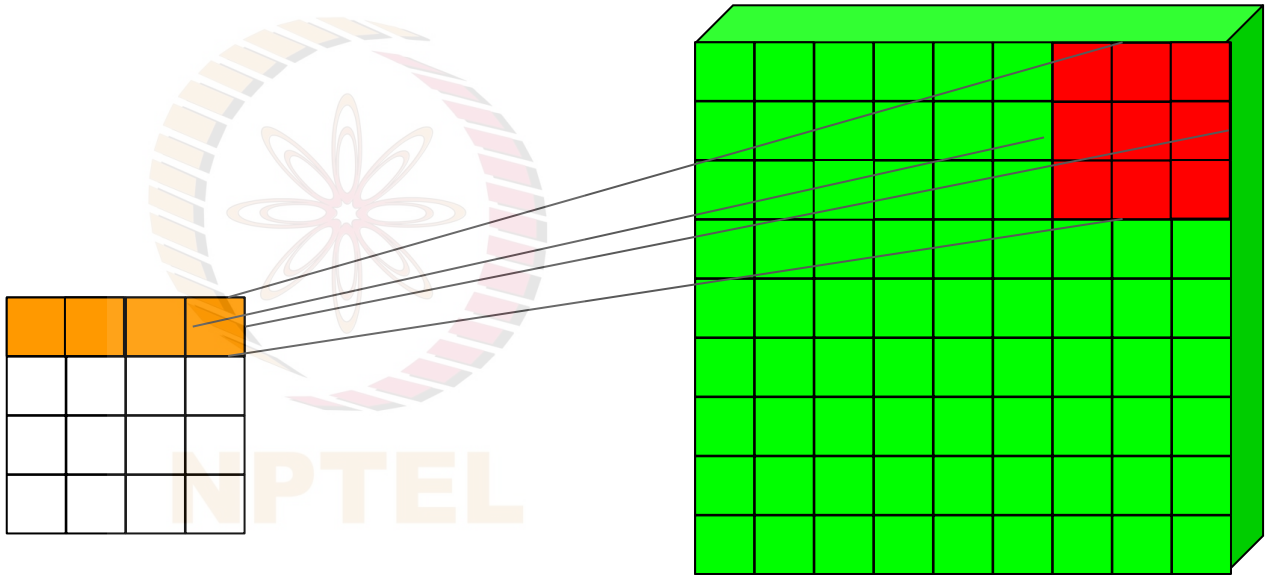
3x3 convolution, Stride=2, Padding=0



3x3 convolution, Stride=1, Padding=0



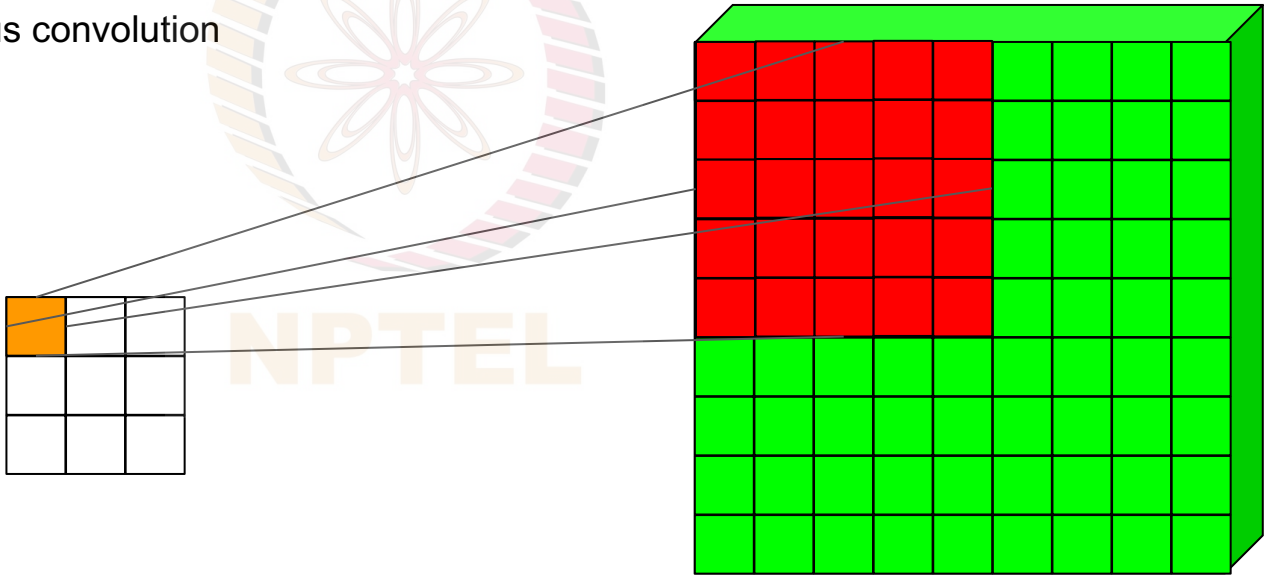
3x3 convolution, Stride=1, Padding=0



3x3 convolution, Stride=1, Padding=0

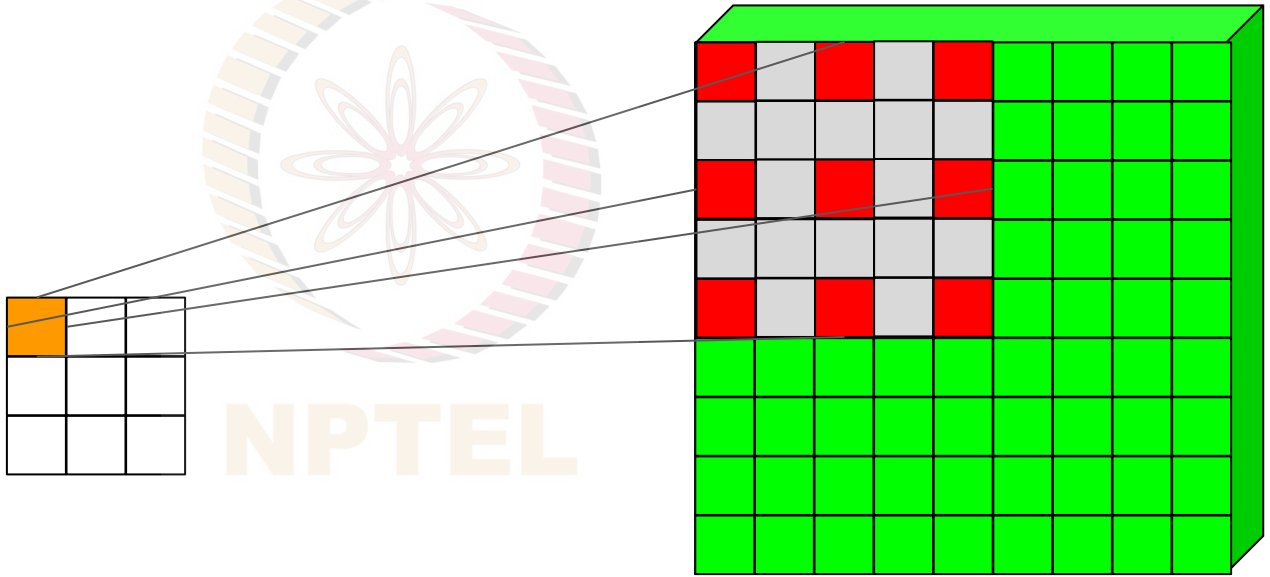
Dilated Convolution

- Technique used to increase the receptive field in a convolution
- Additional parameter - “Dilation Rate”
- Dilation rate varied to resultant feature map is visualize larger areas
- 3 x 3 convolution with Dilation rate set to 2 visualize same area as naive 5x5 convolution, whilst having only 9 parameters.
- Also known as Atrous convolution



Naive 5X5 Convolution

In Dilated Convolution, Kernel is filed with zero appropriately to see larger receptive field



3x3 Convolution, Dilation rate = 2, Receptive field is similar to 5x5 naive convolution

Transposed Convolution

- Convolution operation that aid in increasing the size of the output feature map
- Used in Encoder-Decoder Networks to increase the spatial dimension of the feature map
- In transposed convolution reconstructs the original spatial resolution and performs a convolution.
- The input image is appropriately padded before convolution operation.