Padding

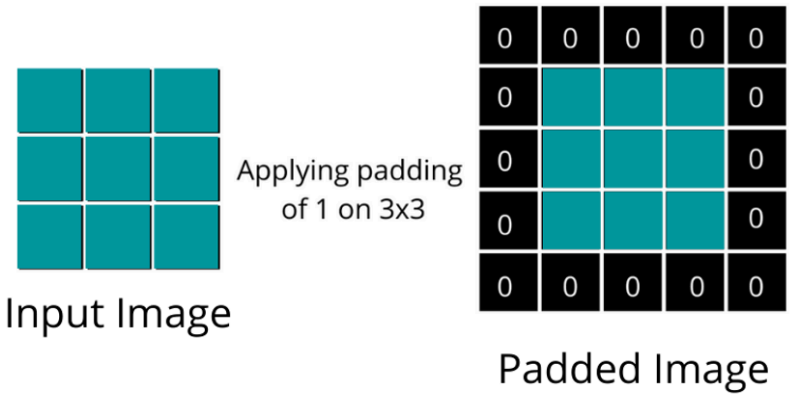
Padding in Convolutional Neural Networks (CNNs) is a technique used to control the spatial dimensions (height and width) of the output feature maps. When applying convolution operations, the spatial dimensions of the input tensor often decrease, which can lead to a significant reduction in the size of the feature maps as the network depth increases. Padding helps to address this issue.

Valid Padding (No Padding):

- In this mode, no padding is applied. The convolution operation is performed only where the kernel (filter) completely overlaps with the input.
- The spatial dimensions of the output feature map are smaller than the input.

Same Padding (Zero Padding):

- In this mode, padding is applied to the input so that the output feature map has the same spatial dimensions as the input.
- Typically, zeros are added around the border of the input tensor.



Purpose of Padding

1. **Preserve Spatial Dimensions**: Padding allows the output feature maps to maintain the same height and width as the input, which is often desirable in many architectures, such as in image segmentation or when stacking multiple convolutional layers.
2. **Maintain Information at the Borders**: Without padding, information at the borders of the input tensor is lost after successive convolution operations. Padding ensures that border information is preserved and considered during convolution.
3. **Control Output Size**: By adjusting the amount of padding, you can control the size of the output feature maps, which can be crucial for designing architectures with specific requirements for feature map sizes.

Strides

In Convolutional Neural Networks (CNNs), strides refer to the step size with which the convolution filter moves across the input image or feature map. The stride determines how much the filter shifts at each step as it convolves around the input.

Key Concepts of Strides

1. **Stride Definition**:
    - **Stride**: The number of pixels by which the filter moves after each operation.
    - Typically denoted as a single integer or a tuple of two integers (stride height and stride width).
2. **Effect on Output Dimensions**:
    - Larger strides result in a smaller output dimension since the filter skips more pixels.
    - Smaller strides (e.g., stride of 1) result in larger output dimensions as the filter moves more gradually across the input.