Module 5

A Convolution Operation

What is Convolution?

Image processing utilizes convolution, a mathematical operation where a matrix (or kernel) traverses the image and performs a dot product with the overlapping region. A convolution operation involves the following steps:

- Define a small matrix (filter).
- This kernel moves across the input image.
- At each location, the convolution operation computes the dot product of the kernel and the portion of the image it overlaps.
- The result of each dot product forms a new matrix, that represents transformed features of the original image.

The primary goal of using convolution in image processing is to extract important features from the image and discard the rest. This results in a condensed representation of an image.

How Convolution Works in CNNs

Convolution Neural Networks (CNNs) is a deep learning architecture that uses multiple convolutional layers combined with multiple Neural Network Layers.

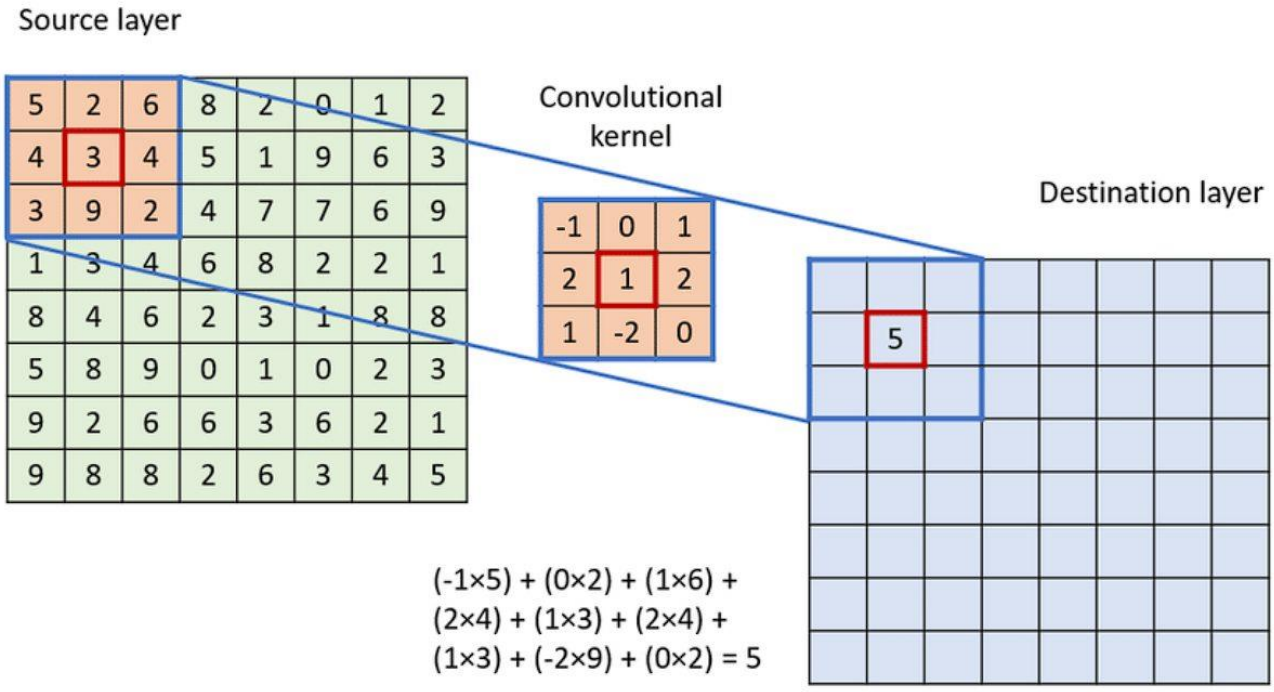
Each layer applies different filters (kernels) and captures various aspects of the image. With increasing layers, the features extracted become dense. The initial layers extract edges and texture, and the final layers extract parts of an image, for example, a head, eyes, or a tail.

Here is how convolution works in CNNs:

- Layers: Lower layers capture basic features, while deeper layers identify more complex patterns like parts of objects or entire objects.
- Learning Process: CNNs learn the filters during training. The network adjusts the filters to minimize the loss between the predicted and actual outcomes, thus optimizing the feature extraction process.

Module 5

- Pooling Layers: After the convolution operations, pooling takes place, which reduces the spatial size of the representation. A pooling layer in CNN downsamples the spatial dimensions of the input feature maps and reduces their size while preserving important information.
- Activation Functions: Neural networks use activation functions, like ReLU (Rectified Linear Unit), at the end to introduce non-linearities. This helps the model learn more complex patterns.



To apply the convolution:

- Overlay the Kernel on the Image: Start from the top-left corner of the image and place the kernel so that its center aligns with the current image pixel.
- Element-wise Multiplication: Multiply each element of the kernel with the corresponding element of the image it covers.

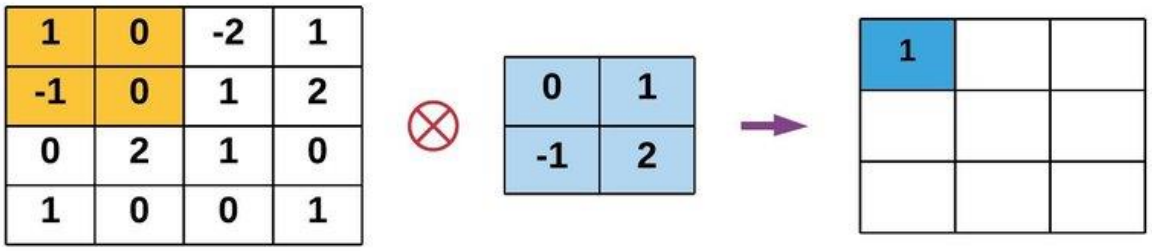
Module 5

- Summation: Sum up all the products obtained from the element-wise multiplication. This sum forms a single pixel in the output feature map.
- Continue the Process: Slide the kernel over to the next pixel and repeat the process across the entire image.

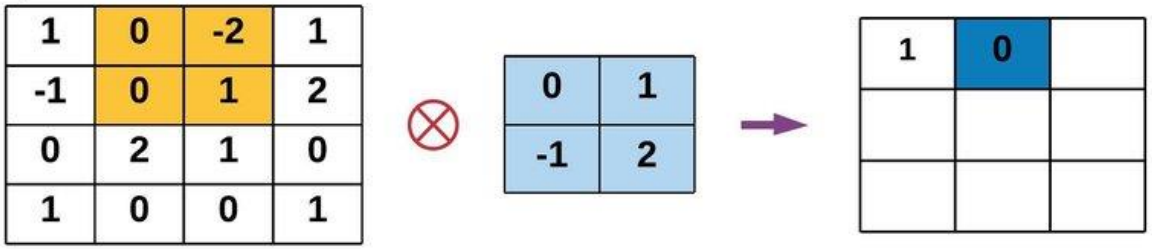
Example of Convolution Operation

Module 5

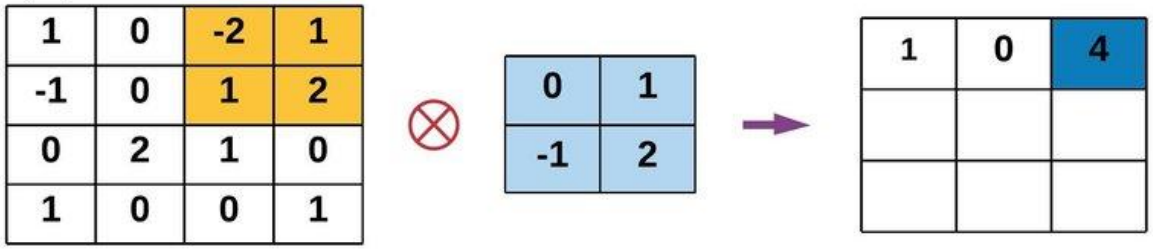
Step-1



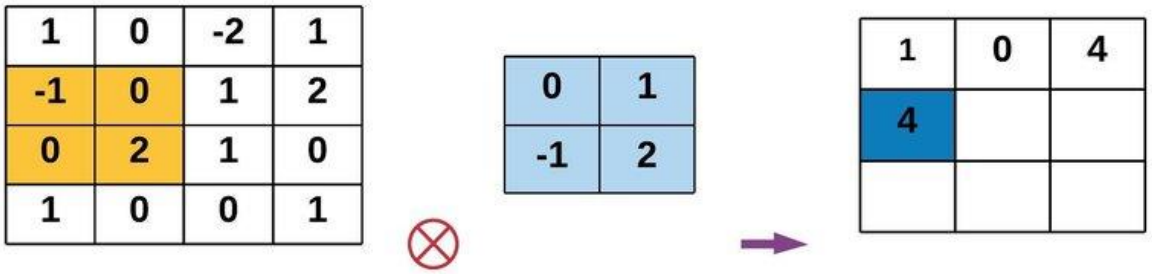
Step-2



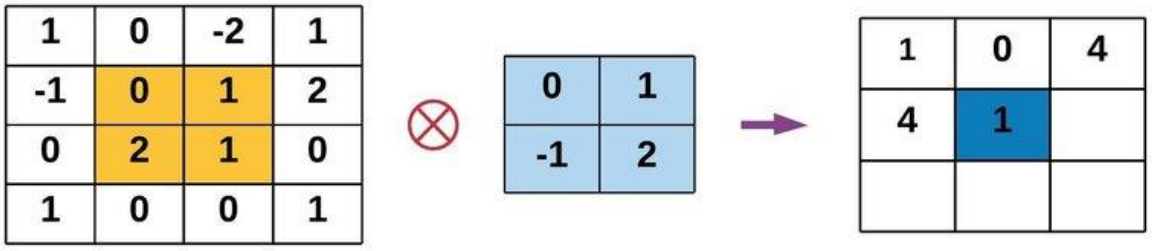
Step-3



Step-4



Step-5



Module 5

Convolution Operation

Key Terms in Convolution Operation

- **Kernel Size:** The convolution operation uses a filter, also known as a kernel, which is typically a square matrix. Common kernel sizes are 3×3, 5×5, or even larger. Larger kernels analyze more context within an image but come at the cost of reduced spatial resolution and increased computational demands.
- **Stride:** Stride is the number of pixels by which the kernel moves as it slides over the image. A stride of 1 means the kernel moves one pixel at a time, leading to a high-resolution output of the convolution. Increasing the stride reduces the output dimensions, which can help decrease computational cost and control overfitting but at the loss of some image detail.
- **Padding:** Padding involves adding an appropriate number of rows and columns (typically of zeros) to the input image borders. This ensures that the convolution kernel fits perfectly at the borders, allowing the output image to retain the same size as the input image, which is crucial for deep networks to allow the stacking of multiple layers.