Q. Explain input shape, output shape, filter, padding, stride and tensor ..... 10M MU-Last Exam

1. Input Shape:

- The input shape refers to the dimensions of the input data that is fed into a neural network layer.
- In CNNs, the input shape represents the dimensions of the input image or feature map, typically expressed as (height, width, channels).
- For example, an input shape of (228, 228, 3) indicates an input image with a height and width of 228 pixels and three color channels (e.g., RGB).

2. Output Shape:

- The output shape refers to the dimensions of the output data produced by a neural network layer after applying transformations.
- In CNNs, the output shape depends on factors such as the filter size, padding, stride, and the architecture of the layer.
- For example, the output shape of a convolutional layer depends on the number of filters used and the size of the input feature map.

3. Filter:

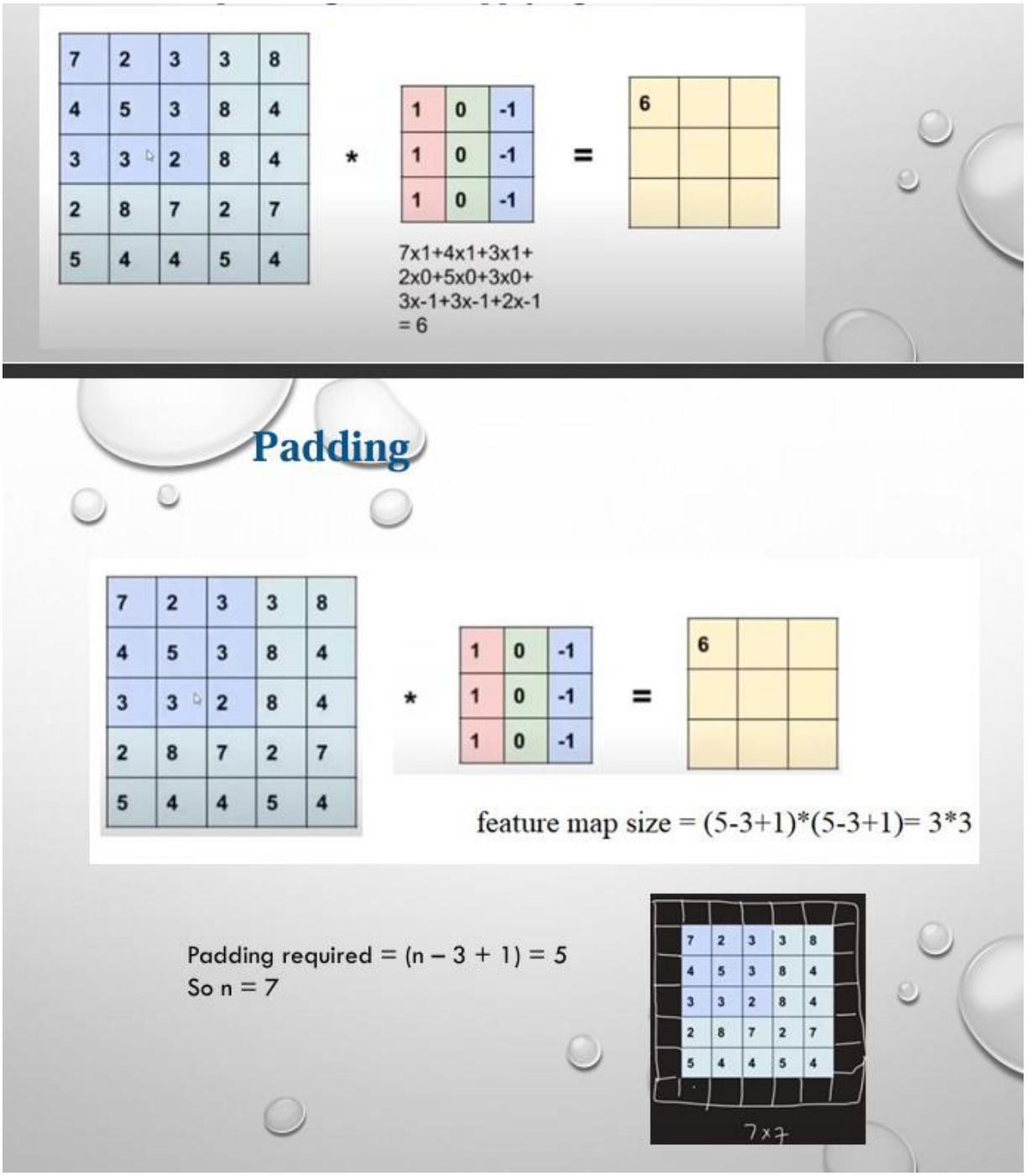
- A filter, also known as a kernel, is a small matrix of weights that is applied to the input data during convolutional operations.
- Filters are used to extract features from input data by performing element-wise multiplications and summations.
- In CNNs, filters slide across the input data to compute convolutions and generate feature maps.

4. Padding:

Padding is a technique used in convolutional neural networks (CNNs) to manage the size of feature maps as they pass through convolutional layers.

- When a filter is applied to an input image, the output feature map is typically smaller than the input image due to the border pixels being ignored.

- Padding involves adding additional pixels (usually zeros) around the border of the input image before applying the filter.



The main reason for using padding is to preserve spatial information and ensure that the output feature map has the same spatial dimensions as the input image.

- This is important because it allows the network to capture information from the entire input image, including pixels near the border.

Without padding, the size of the feature maps would shrink with each convolutional layer, leading to a loss of spatial information.

- This reduction in size can also cause issues at the boundaries of the image, where important features may be located.
- Padding helps mitigate these issues by ensuring that the convolutional operation is applied uniformly across the entire input image, thereby preserving spatial information and improving the performance of the network.

Feature map size after padding = (n +2p – f + 1) by (n + 2p – f + 1)

5. Stride:

- Strides refer to the step size with which the convolutional filter slides (moves) across the input image or feature map during the convolution operation.
- The stride determines the amount by which the filter shifts over the input image at each step. For example, if the stride is set to 1, the filter moves one pixel at a time. If the stride is set to 2, the filter moves two pixels at a time, and so on.
- The main purpose of using strides is to control the spatial dimensions of the output feature map produced by the convolutional layer. By adjusting the stride size, we can control the amount of overlap between adjacent receptive fields (regions of the input image that are covered by the filter).
- Using larger stride values reduces the spatial dimensions of the output feature map, leading to a decrease in computational complexity and memory usage. Conversely, smaller stride values result in a larger output feature map with more spatial information.
- Strides allow us to adjust the spatial resolution of feature maps in CNNs, providing flexibility in balancing computational efficiency with spatial information preservation.

6. Tensor:

- A tensor is a multi-dimensional array used to represent data in neural networks.
- In CNNs, tensors are used to store and manipulate input data, weights, biases, and intermediate feature maps.
- Tensors have a rank, shape, and data type, and they are fundamental to the operations performed in deep learning models.