ADITYA AMIN

ASSIGN : CV - 01

1. What exactly is a feature?

In the context of computer vision (CV), a feature refers to a specific, measurable attribute or characteristic of an image that can be used to represent and distinguish it from other images. Features are fundamental building blocks in many CV tasks, such as object detection, image recognition, and image matching.

Point features: These represent distinctive points in an image, such as corners or keypoints, that are invariant to translation, rotation, and scale changes. Examples include the Scale-Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF).

1. For a top edge detector, write out the convolutional kernel matrix.

One of the commonly used edge detectors is the Sobel operator. The Sobel operator consists of two kernels: one for detecting horizontal edges and the other for detecting vertical edges.

Horizontal Edge Detection Kernel:

-1 -2 -1

0 0 0

1 2 1

Vertical Edge Detection Kernel:

-1 0 1

-2 0 2

-1 0 1

3. Describe the mathematical operation that a 3x3 kernel performs on a single pixel in an image.

A 3x3 kernel performs a mathematical operation known as convolution on a single pixel in an image. The convolution operation involves multiplying each element of the kernel with the corresponding pixel value in the image, and then summing up these multiplications.

Let's denote the 3x3 kernel as K and the image as I. Suppose we want to perform the convolution operation on a pixel located at position (x, y) in the image. The convolution operation can be defined as follows:

Convolution at pixel (x, y) = Summation of (K(i, j) \* I(x+i-1, y+j-1)) for i = -1 to 1 and j = -1 to 1

4. What is the significance of a convolutional kernel added to a 3x3 matrix of zeroes?

the significance of adding a convolutional kernel to a 3x3 matrix of zeroes lies in preserving spatial dimensions, handling edges effectively, and providing a neutral starting point for further training or modification of the kernel.

5. What exactly is padding?

In the context of image processing and convolutional neural networks (CNNs), padding refers to the technique of adding extra pixels or values around the borders of an image. It is used to adjust the spatial dimensions of the image or feature maps before performing certain operations, such as convolution or pooling.

6. What is the concept of stride?

In the context of image processing and convolutional neural networks (CNNs), stride refers to the step size or the number of pixels by which a convolutional kernel moves across an image or feature map during the convolution operation. It determines the amount of spatial overlap between successive kernel positions.

7. What are the shapes of PyTorch's 2D convolution's input and weight parameters?

Input Tensor Shape: The input tensor shape for a 2D convolution in PyTorch is typically represented as [batch\_size, channels, height, width].

Weight Tensor Shape: The weight tensor shape for a 2D convolution layer in PyTorch depends on the number of input channels and output channels of the layer. It is represented as [output\_channels, input\_channels, kernel\_height, kernel\_width].

8. What exactly is a channel?

In the context of image processing and convolutional neural networks (CNNs), a channel refers to an individual component or layer of information within an image or feature map. Channels represent specific aspects or characteristics of an image, such as color information, texture, or specific visual patterns.

9.Explain relationship between matrix multiplication and a convolution?

Matrix multiplication and convolution are mathematically related operations, and the relationship between them becomes apparent when considering the convolution operation in the context of image processing and convolutional neural networks (CNNs).