

Answer 1

- (a) There are 3 filters of dimensions $5 \times 5 \times 3$ with a valid padding. To get the learnable parameter for convolution layer, we need to multiply the number of filters = 3 with the number of parameters which is $5 \times 5 \times 3 = 75$, so there are 225 learning parameters in this convolutional layer.
- (b) To replicate the behavior of this convolutional layer using a fully connected layer, the parameters we need is the same as above which is 225 parameters as it is replicating the same behavior of the convolutional layer mentioned above.

Answer 2

Firstly, I would use convolution on the image shown in figure 1(a) with the filters mentioned in answer 1. Then, I would use summation and add the two outputs I got from each image. Here, each filter is used for convolution. This would help us to find the highest match of feature to the image. Finally, I would use argmax to find the index that corresponds to the kernel where the highest point is where it is matched with the desired pattern.

Answer 3

Convolution is translation invariant for 1D convolution which can be achieved with the help of pooling layers. In max pooling, the highest activation input is taken. Even if the image is translated to a different location, the exact same operations will also take place in the new location too. With the help of filters, we look at each piece of image and it is done for each filter. Hence, convolution is translation invariant.

Answer 4 (b)

Attached to github

<https://github.com/aryan-s1/Deep-Learning-Intro-/tree/main/HW3>