**PadhAI: From Convolution Operation to Neural Network** One Fourth Labs

**Understanding the input/output dimensions**

Let’s look at the input and output dimensions for a Convolutional Operation 1. As we have seen before, a CNN can be compared to a normal Neural Network, the difference being that CNNs take the RGB pixel values as inputs and output calculation is done with a localised neighborhood of inputs. 2. Consider the following diagram of a CNN. Let us dissect the first convolutional operation in

depth.

a. From the above diagram, we are analysing the convolutional operation on the grey input layer. b. The input dimensions are as follows

i. WI = 227 ii. HI = 227 iii. DI = 3 c. The filter is of scale F = 11, i.e 11x11x3, where 3 is the same depth as DI d. We apply 96 Filter operations, so therefore K = 96 e. We do not take any padding (P=0) and we choose a stride length of S = 4 f. Thus, going by the above information, the output volume can be calculated as follows

i. ii. *W O* = *HO* = *W I* − *F* + 2*P H I* − *F S* + 2*P S*

+ 1 = 55 + 1 = 55 iii. *DO* = *K* = 96 g. Thus, the output of the convolutional layer has the dimensions 55x55x96