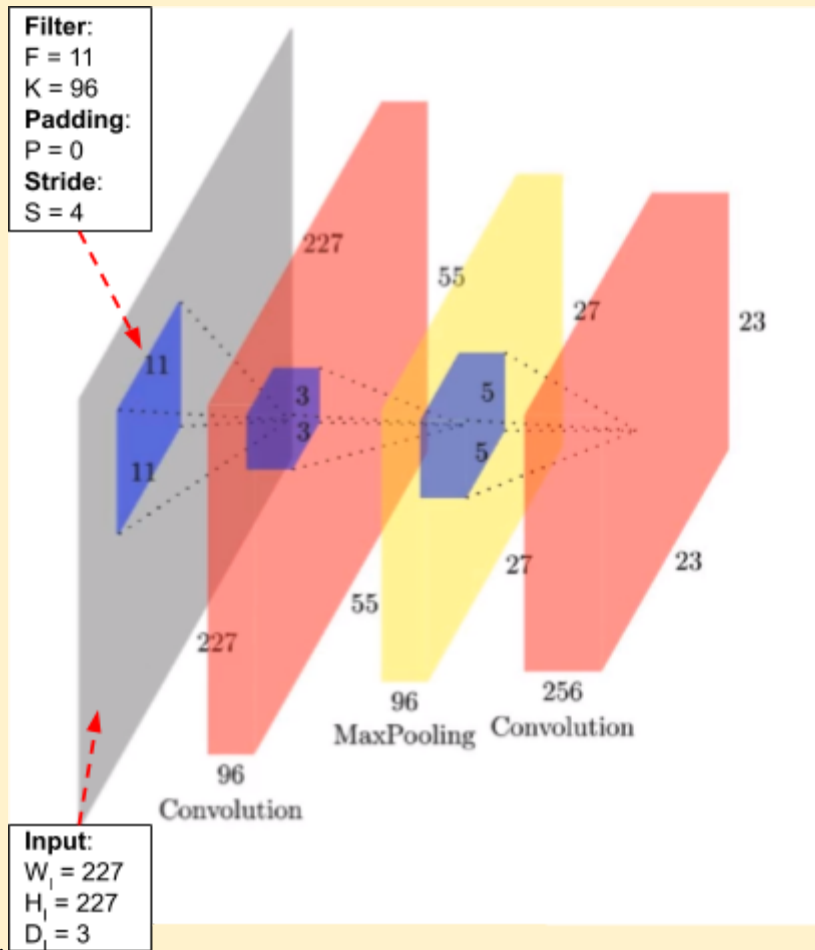


## One Fourth Labs

### 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.  $W_I = 227$
  - ii.  $H_I = 227$
  - iii.  $D_I = 3$
- c. The filter is of scale  $F = 11$ , i.e  $11 \times 11 \times 3$ , where 3 is the same depth as  $D_I$
- 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.  $W_O = \frac{W_I - F + 2P}{S} + 1 = 55$
  - ii.  $H_O = \frac{H_I - F + 2P}{S} + 1 = 55$
  - iii.  $D_O = K = 96$
- g. Thus, the output of the convolutional layer has the dimensions  $55 \times 55 \times 96$