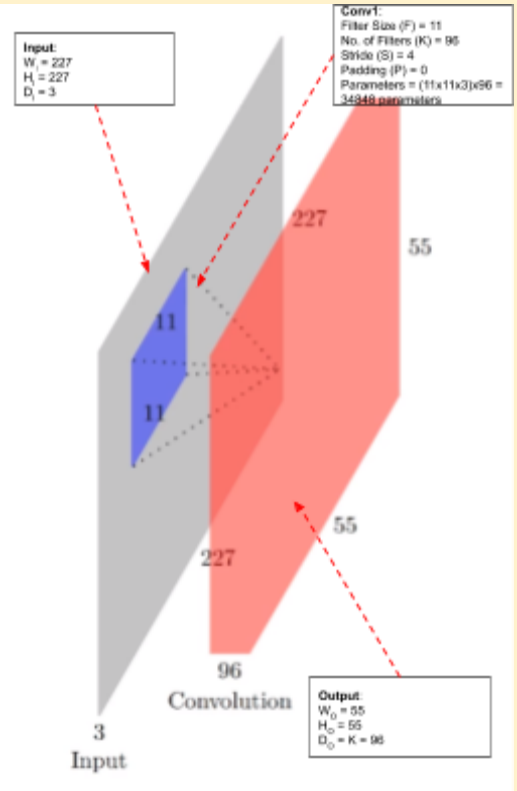
PadhAI: CNN Architectures

One Fourth Labs

Understanding the first layer of AlexNet

Let's break down the first layer of the AlexNet architecture

1. AlexNet was the winning architecture for the 2012 Imagenet Challenge
2. Let us look at the convolutional layer



3. The details are as follows
   a. **Input images**: 227x227x3 (colour images of 227x227 Width x Height)
      i. $W_I = 227$
      ii. $H_I = 227$
      iii. $D_I = 3$
   b. **Filter/Conv1** layer:
      i. Filter Size (F) = 11 (i.e. FxFxD_I or 11x11x3)
      ii. No. of Filters (K) = 96
      iii. Stride (S) = 4
      iv. Padding (P) = 0
      v. Parameters = (11x11x3) x 96 = 34,848
      vi. These values were determined through extensive experimentation
   c. **Output**:
      i. $W_O = \frac{W_I - F + 2P}{S} + 1 = 55$
      ii. $H_O = \frac{H_I - F + 2P}{S} + 1 = 55$
      iii. $D_O = K = 96$

4. This was a standard architecture and can be used for a variety of tasks.