## **Understanding all the layers of AlexNet**

Let’s now look at the entire AlexNet

1. Before moving into the AlexNet architecture, let us understand why we have decided to use a Convolutional layer instead of a Fully-connected layer
   1. In the convolutional layer, the number of parameters is (11x11x3) x 96 = **34,848**
   2. However, in a FC layer, the number of parameters would be
      1. Input:(227x227x3) x Output:(55x55x96) = **4.49 x 1010**
   3. Thus, because of sparse-connectivity and weight sharing, we are able to achieve a similar degree of complexity with a significantly smaller number of parameters with a Convolutional layer.
2. Now, let us **break down the entire AlexNet architecture**
3. Let’s look at each of the layers in depth
4. **Input Layer**: 227x227x3 (colour images of 227x227 Width x Height)
   1. WIn = 227
   2. HIn = 227
   3. DIn = 3
5. **Convolutional Layer 1**: Input is 227x227x3
   1. Filter Size (**F**) = 11 (11x11x3)
   2. No. of Filters (**K**) = 96
   3. Stride (**S**) = 4
   4. Padding (**P**) = 0
   5. **Parameters** = (11x11x3) x 96 = 34,848
   6. **W1** = 55
   7. **H1** = 55
   8. **D1** = K = 96
   9. **ReLU** Non-linearity function is applied to every 2D area in the output volume.
6. **Max-Pooling Layer 1**: Input is 55x55x96
   1. Filter Size (**F**) = 3 (i.e. 3x3x96)
   2. Stride (**S**) = 4
   3. **Parameters** = 0 (no parameters in max pooling)
   4. **W1m** = 27
   5. **H1m** = 27
   6. **D1m** = 96
7. **Convolutional Layer 2**: Input is 27x27x96
   1. Filter Size (**F**) = 5 (5x5x96)
   2. No. of Filters (**K**) = 256
   3. Stride (**S**) = 1
   4. Padding (**P**) = 0
   5. **Parameters** = (5x5x96) x 256 = 614,400
   6. **W2** = 23
   7. **H2** = 23
   8. **D2** = K = 256
   9. **ReLU** Non-linearity function is applied.
8. **Max-Pooling Layer 2**: input is 23x23x256
   1. Filter Size (**F**) = 3 (3x3x256)
   2. Stride (**S**) = 3
   3. **Parameters** = 0
   4. **W2m** = 11
   5. **H2m** = 11
   6. **D2m** = 256
9. **Convolutional Layer 3**: input is 11x11x256
   1. Filter Size (**F**) = 3 (3x3x256)
   2. No. of Filters (**K**) = 384
   3. Stride (**S**) = 1
   4. Padding (**P**) = 0
   5. **Parameters** = (3x3x256) x 384 = 884,736
   6. **W3** = 9
   7. **H3** = 9
   8. **D3** = K = 384
   9. **ReLU** Non-linearity function is applied.
10. **Convolutional Layer 4**: input is 9x9x384
    1. Filter Size (**F**) = 3 (3x3x384)
    2. No. of Filters (**K**) = 384
    3. Stride (**S**) = 1
    4. Padding (**P**) = 0
    5. **Parameters** = (3x3x384) x 384 = 1,327,104
    6. **W4** = 7
    7. **H4** = 7
    8. **D4** = K = 384
    9. **ReLU** Non-linearity function is applied.
11. **Convolutional Layer 5**: input is 7x7x384
    1. Filter Size (**F**) = 3 (3x3x384)
    2. No. of Filters (**K**) = 256
    3. Stride (**S**) = 1
    4. Padding (**P**) = 0
    5. **Parameters** = (3x3x384) x 256 = 884,736
    6. **W5** = 5
    7. **H5** = 5
    8. **D5** = K = 256
    9. **ReLU** Non-linearity function is applied.
12. **Max-Pooling Layer 3**: input is 5x5x256
    1. Filter Size (**F**) = 3 (3x3x256)
    2. Stride (**S**) = 2
    3. **Parameters** = 0
    4. **W5m** = 2
    5. **H5m** = 2
    6. **D5m** = 256
13. **Fully Connected Layer 1**: input is 2x2x256 = 1024
    1. Number of Neurons = 4096
    2. Parameters = (2x2x256) x 4096 = 4,194,304
14. **Fully Connected Layer 2**: input = 4096
    1. Number of Neurons = 4096
    2. Parameters = 4096 x 4096 = 16,777,216
15. **Fully Connected Layer 3**: input = 4096
    1. Number of Neurons/output-classes = 1000
    2. Parameters = 4096 x 1000 = 4,096,000
16. Totally, there are around 27.55 Million parameters, out of which roughly 25 Million parameters were in the last 3 Fully-connected layers.
17. When counting the total number of layers, we do not include the max-pooling layers as they do not carry weights. Thus, we say that **AlexNet has 8 layers**