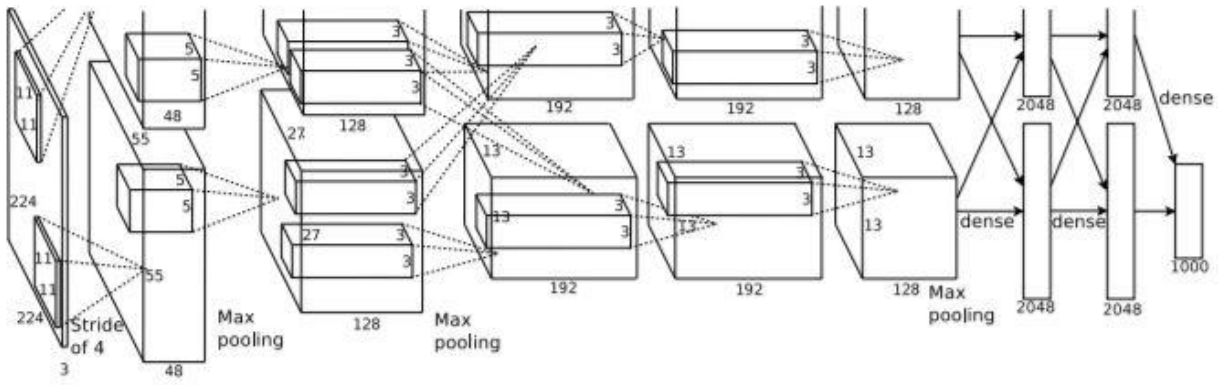
AlexNet



First Layer:
The input for AlexNet is a 227x227x3 RGB image which passes through the first convolutional layer with 96 feature maps or filters having size 11×11 and a stride of 4. The image dimensions changes to 55x55x96.
Then the AlexNet applies maximum pooling layer or sub-sampling layer with a filter size 3×3 and a stride of two. The resulting image dimensions will be reduced to 27x27x96.
Second Layer:
Next, there is a second convolutional layer with 256 feature maps having size 5×5 and a stride of 1.
Then there is again a maximum pooling layer with filter size 3×3 and a stride of 2. This layer is same as the second layer except it has 256 feature maps so the output will be reduced to 13x13x256.
Third, Fourth and Fifth Layers:
The third, fourth and fifth layers are convolutional layers with filter size 3×3 and a stride of one. The first two used 384 feature maps where the third used 256 filters.
The three convolutional layers are followed by a maximum pooling layer with filter size 3×3, a stride of 2 and have 256 feature maps.
Sixth Layer:
The convolutional layer output is flatten through a fully connected layer with 9216 feature maps each of size 1×1.
Seventh and Eighth Layers:
Next is again two fully connected layers with 4096 units.
Output Layer:
Finally, there is a softmax output layer ŷ with 1000 possible values.

Summary:

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	227x227x3	-	-	-
1	Convolution	96	55 x 55 x 96	11x11	4	relu
	Max Pooling	96	27 x 27 x 96	3x3	2	relu
2	Convolution	256	27 x 27 x 256	5x5	1	relu
	Max Pooling	256	13 x 13 x 256	3x3	2	relu
3	Convolution	384	13 x 13 x 384	3x3	1	relu
4	Convolution	384	13 x 13 x 384	3x3	1	relu
5	Convolution	256	13 x 13 x 256	3x3	1	relu
	Max Pooling	256	6 x 6 x 256	3x3	2	relu
6	FC	-	9216	-	-	relu
7	FC	-	4096	-	-	relu
8	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	Softmax

https://engmrk.com/alexnet-implementation-using-keras/