



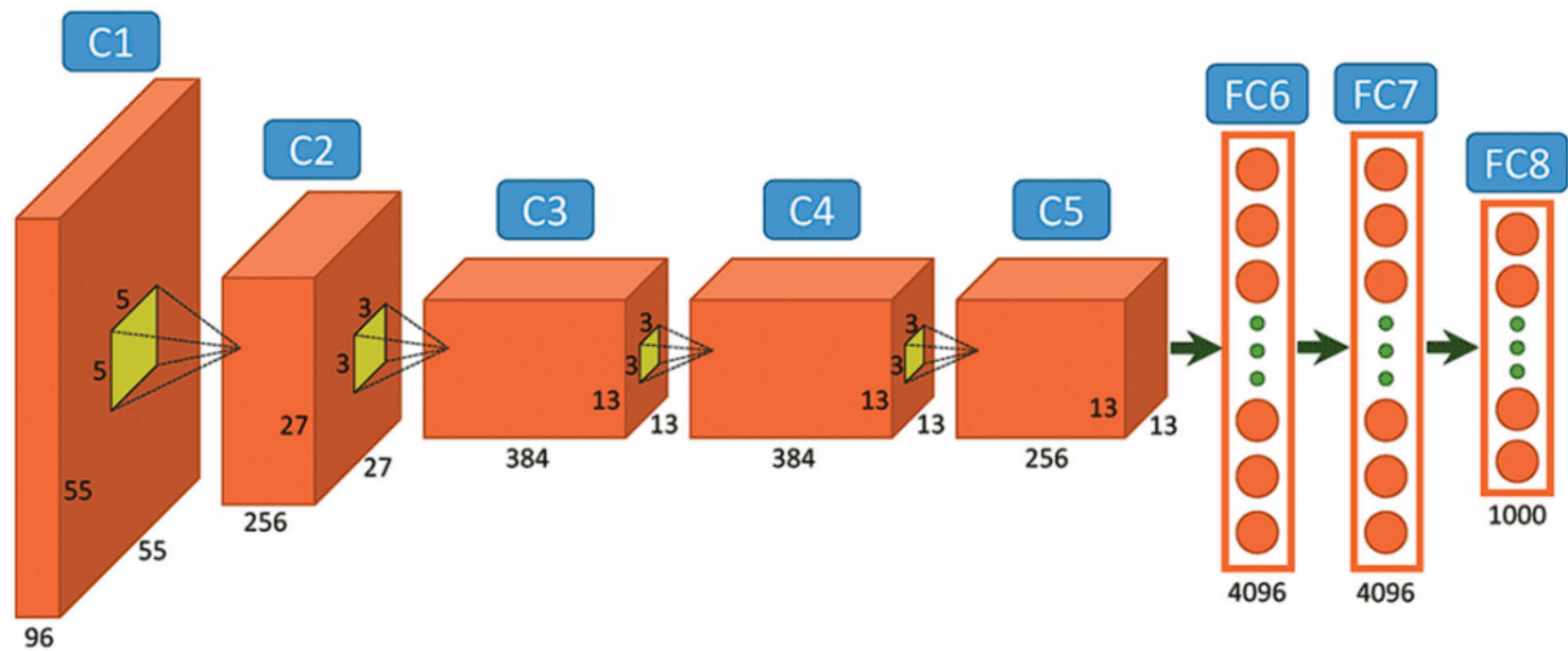
ALEXNET

THE ARCHITECTURE

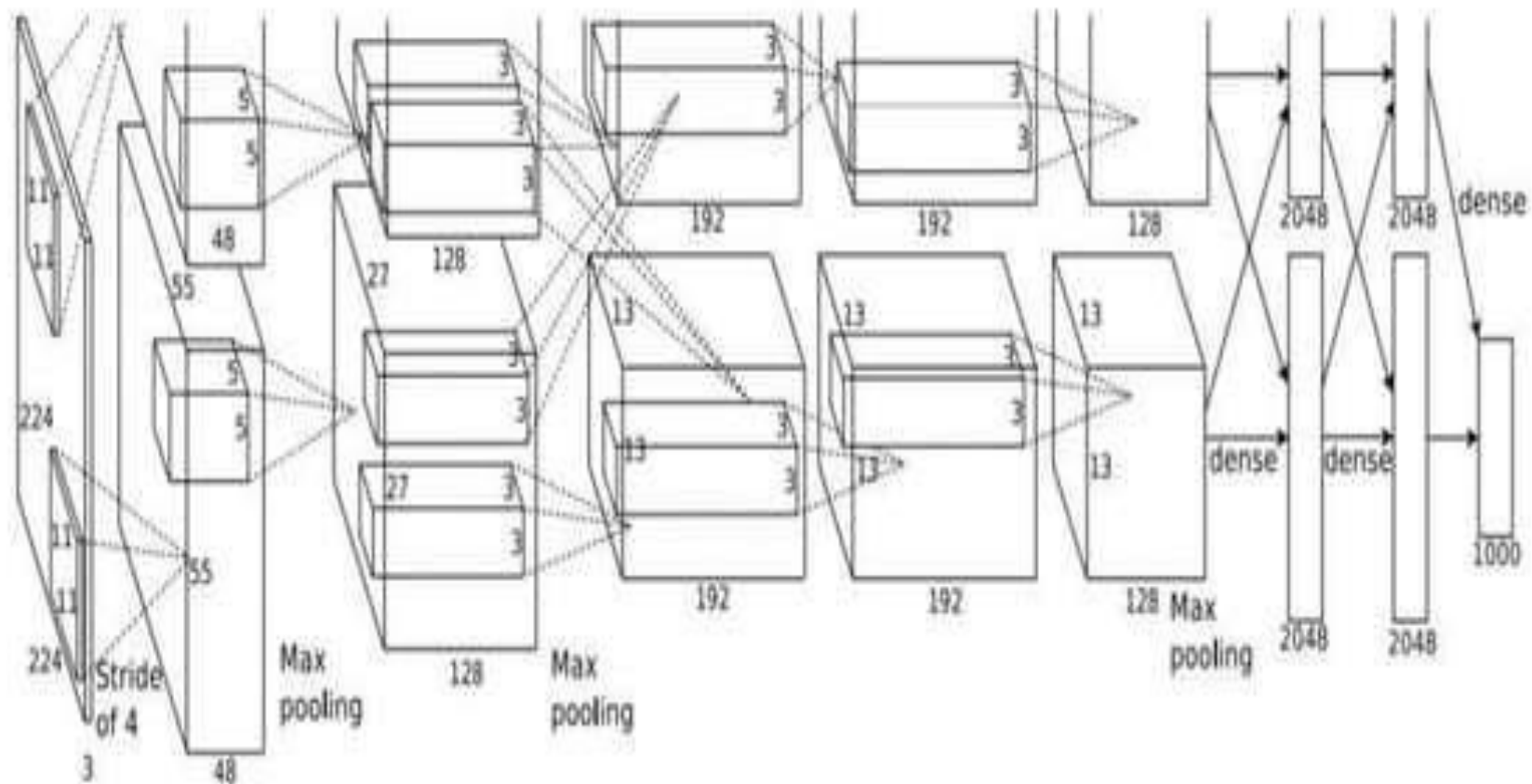
- The AlexNet architecture was proposed by Alex Krizhevsky, Llya Sutskever, and Geoffrey E. Hilton and research paper written based on this architecture is "[imageNet Classification With Deep Convolution Neural Networks](#)"
- The architecture has:
 1. Five Convolution Layers
 2. Three Fully Connected Layers (Final Layer is softmax)

CONTD...

- There are 60 Million Parameters in AlexNet so it is called as Deep Neural network
- It has more filters per layer as compared to LeNet. Convolution layers are more in AlexNet
- Dropout feature is added to ensure overfitting problem
- Data Augmentation(Mirroring image, to increase the training volume) is also done



THE ARCHITECTURE



Input Layer

AlexNet takes images of the Input size of $227 \times 227 \times 3$ RGB Pixels.

Convolutional Layers

- First Layer:** The first layer uses 96 kernels of size 11×11 with a stride of 4, activates them with the ReLU activation function, and then performs a Max Pooling operation.
- Second Layer:** The second layer takes the output of the first layer as the input, with 256 kernels of size $5 \times 5 \times 48$.
- Third Layer:** 384 kernels of size $3 \times 3 \times 256$. No pooling or normalization operations are performed on the third, fourth, and fifth layers.
- Fourth Layer:** 384 kernels of size $3 \times 3 \times 192$.
- Fifth Layer:** 256 kernels of size $3 \times 3 \times 192$.

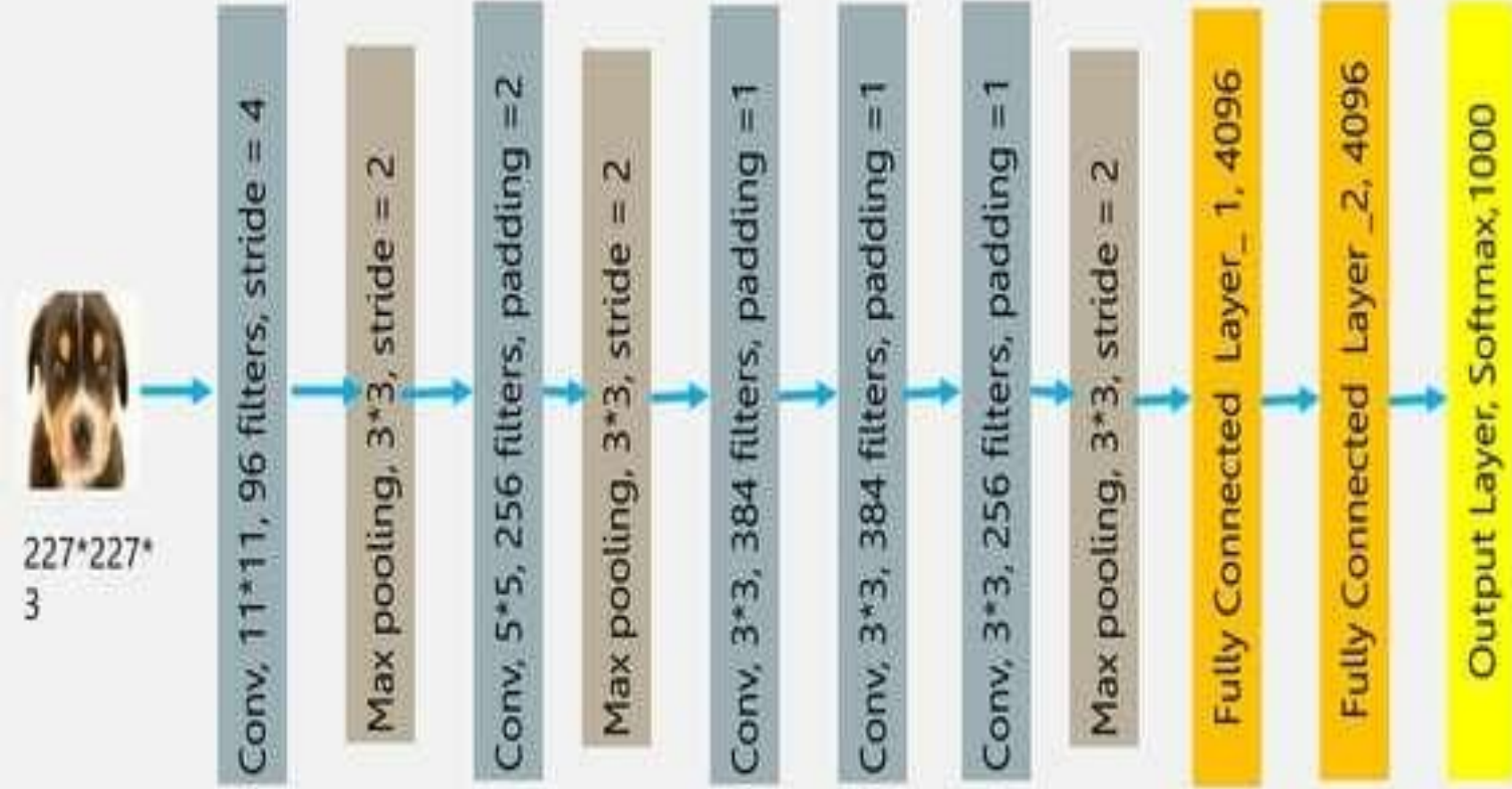
Fully Connected Layers

The fully connected layers have 4096 neurons each.

Output Layer

The output layer is a SoftMax layer that outputs probabilities of the 1000 class labels.

THE ARCHITECTURE



DIMENSIONS AND DETAILS

	Layer	Filters	size	Kernel Size	Stride	Activation
Input	Image	1	227*227*3	-		
1	Convolution	96	55*55*96	11*11	4	relu
	Max Pooling	96	27*27*256	3*3	2	
2	Convolution	256	27*27*256	5*5	1	relu
	Max Pooling	256	13*13*256	3*3	2	
3	Convolution	384	13*13*384	3*3	1	relu
4	Convolution	384	13*13*384	3*3	1	relu
5	Convolution	256	13*13*256	3*3	1	relu
	Max Pooling	256	6*6*256	3*3	2	
6	Fully Connected	-	9216 (Neurons)	-	-	relu
7	Fully Connected	-	4096 (Neurons)	-	-	relu

DIMENSIONS AND DETAILS

Without Padding Formula

$$\left(\frac{n-f}{s} + 1, \frac{n-f}{s} + 1 \right)$$
$$\left(\frac{227-11}{4} + 1, \frac{227-11}{4} + 1 \right) = 55*55*96$$

With Padding Formula

$$\left(\frac{n+2p-f}{s} + 1, \frac{n+2p-f}{s} + 1 \right)$$
$$\left(\frac{27+2(2)-5}{1} + 1, \frac{27+2(2)-5}{1} + 1 \right) = 27*27*256$$

THE ARCHITECTURE

