Convolution Neural Network

Praveen Kumar Chandaliyal

Malaviya National Institute of Technology

2016rcp9511@mnit.ac.in

Under The Supervision of : Dr. Neeta Nain

November 1, 2018

Hyper Parameter Calculate Formal's

1. Convolutional Layer Size Calculate Formula

$$K = \frac{(W - F + 2P)}{S} + 1 \tag{1}$$

2. Pooling Layer Size Calulate Formula

$$K = \frac{(W - F)}{S} + 1 \tag{2}$$

- W: Input volume size
- F: Filter Size
- S: Stride (Sampling in the Case of Pooling)
- P: Padding

CNN Architecture

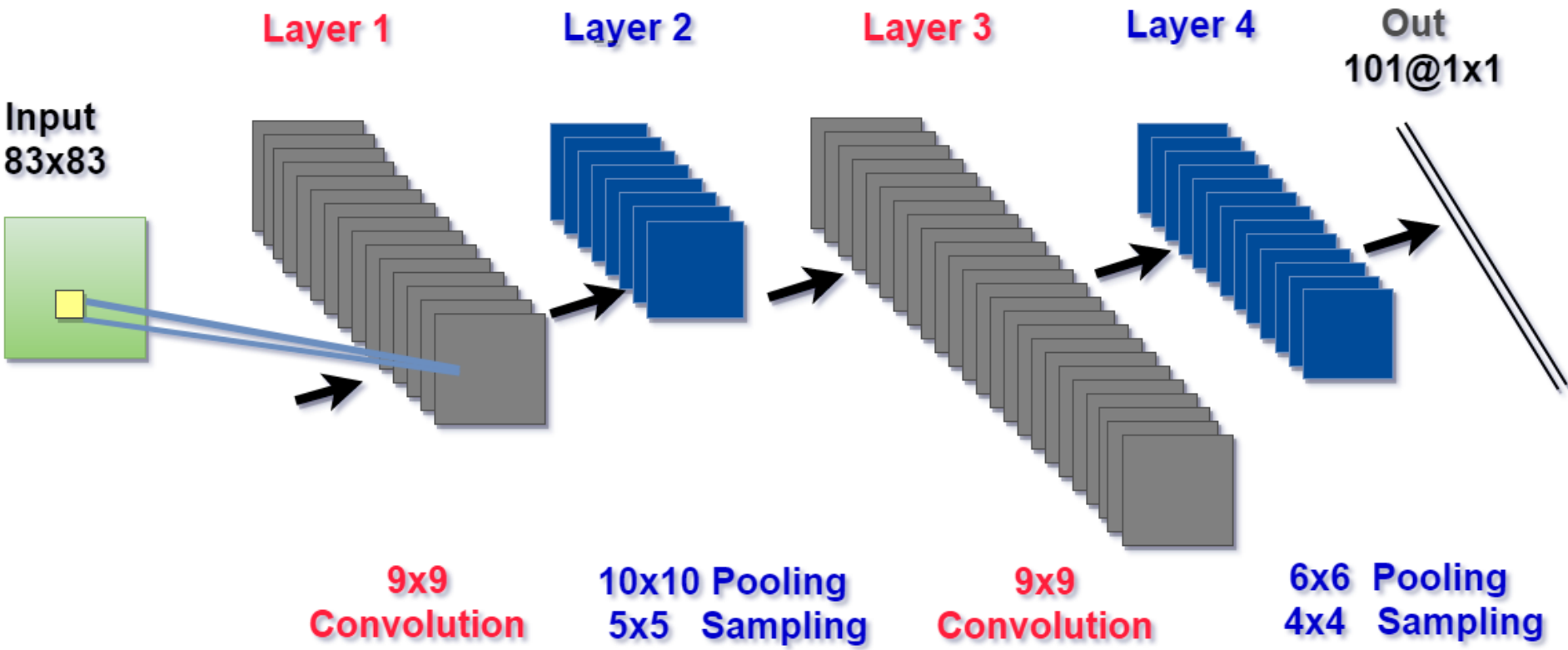

Figure 1: CNN

CNN Architecture

Input (W=83, F=9, P=0, S=1) => Layer 1 Size : 64 @ 75 x 75

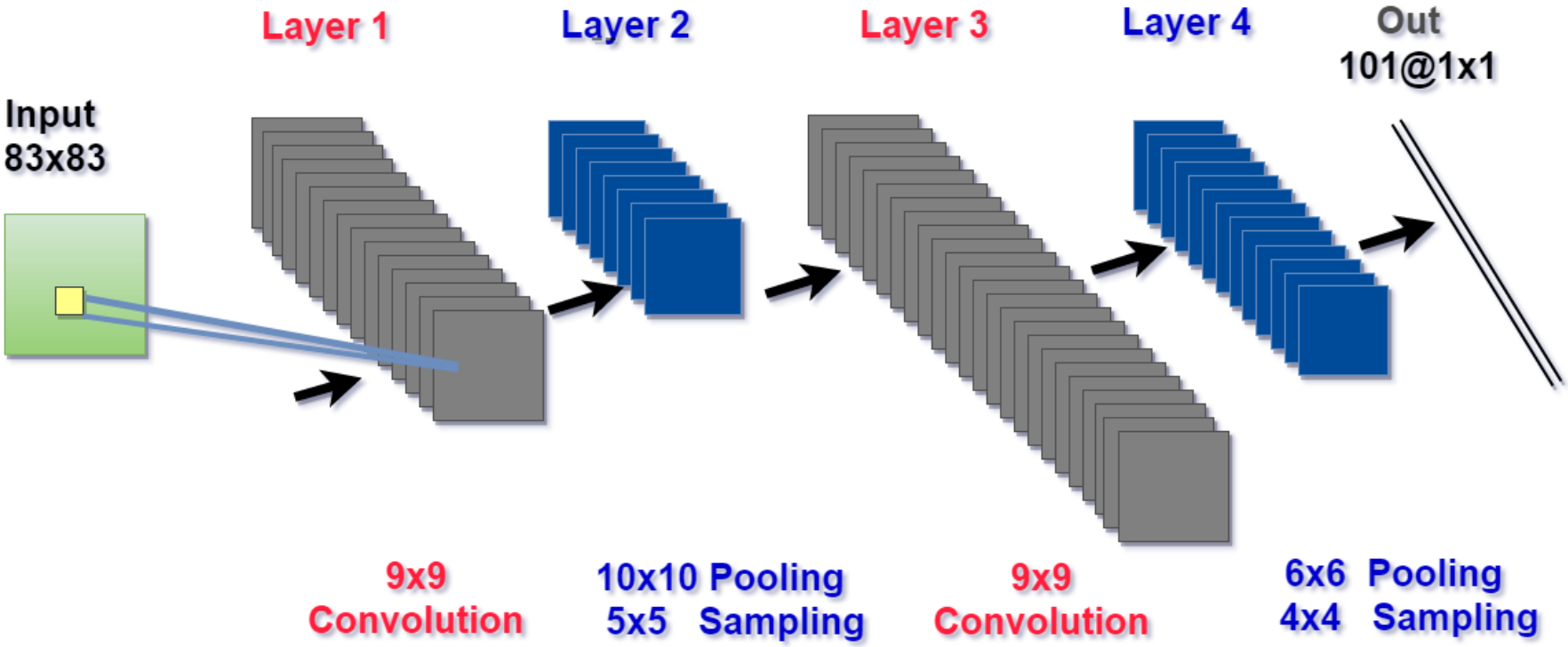

Figure 2: CNN

CNN Architecture

Input (W=75, F=10, P=0, S=5) => Layer 2: 64 @ 14 x 14

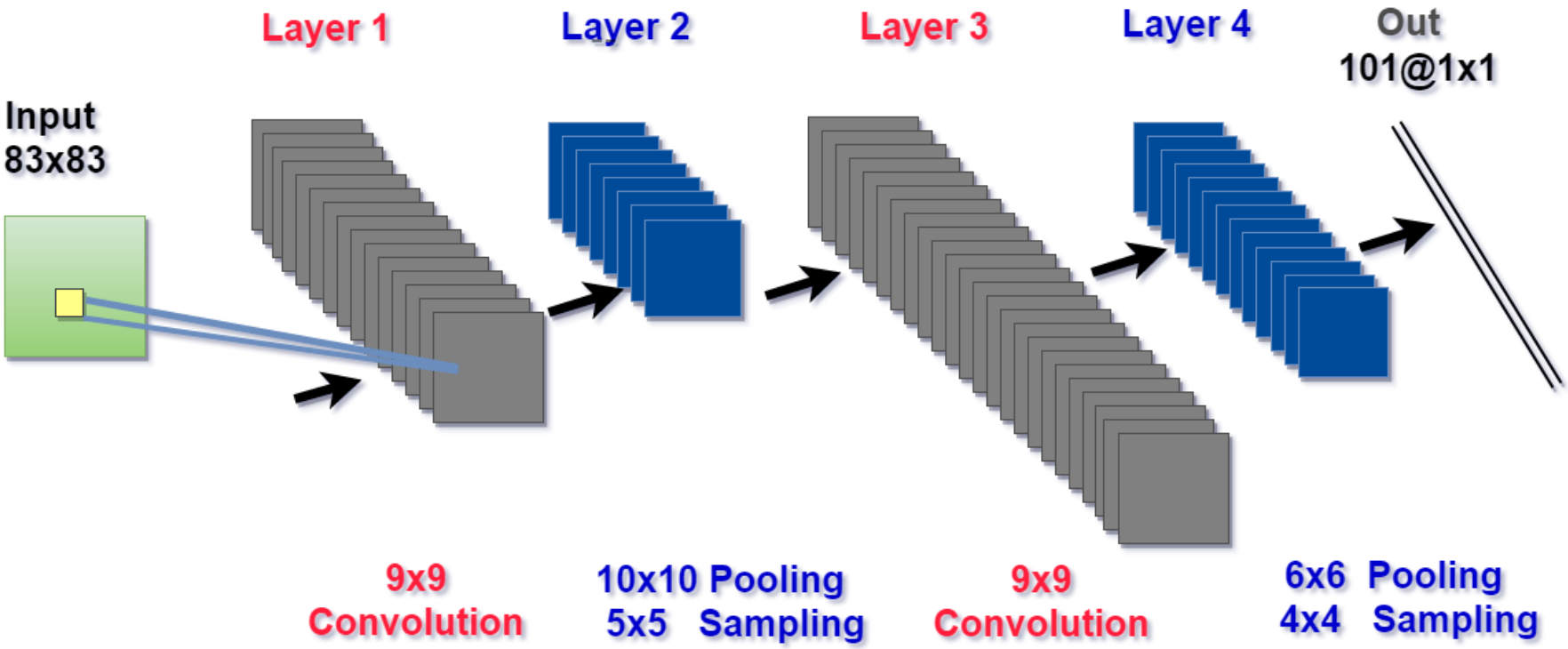

Figure 3: CNN

CNN Architecture

Input (W=14, F=9, P=0, S=1) => Layer 3: 256 @ 6 x 6

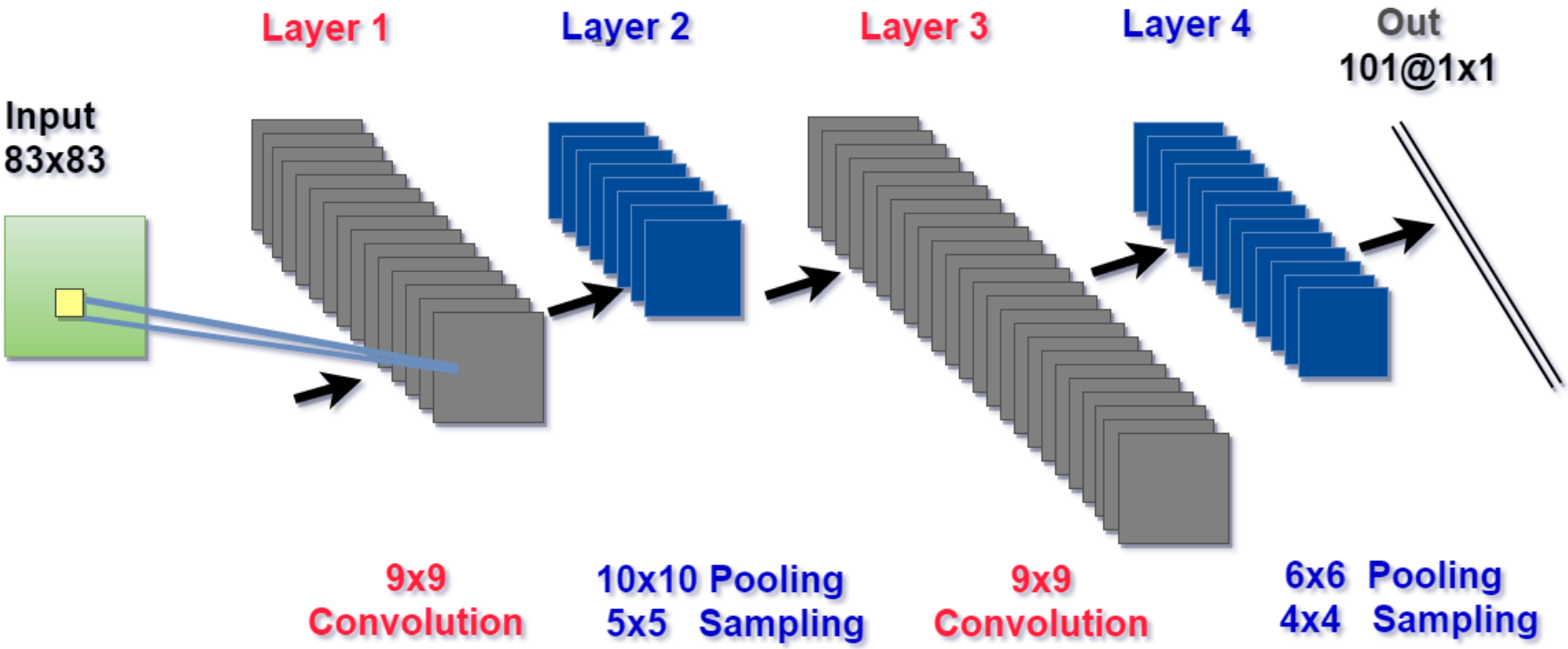

Figure 4: CNN

CNN Architecture

Input (W=6 F=6 P=0 S=4) => Layer 4: 256 @ 1 x 1

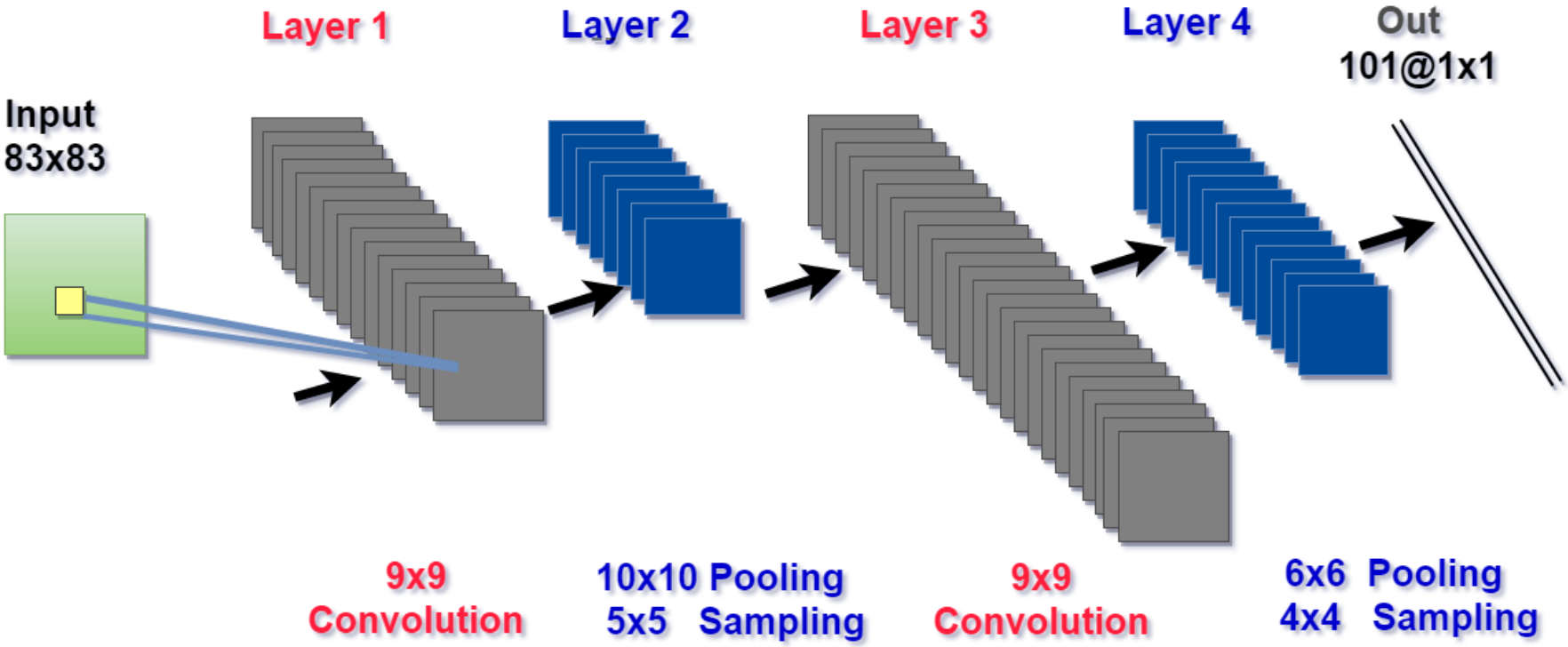

Figure 5: CNN

Learnable Parameters For CNN

1. Input Layer: No parameter learning
2. Convolutional Layers:

$$\text{Number of parameter} = (n * m * l + 1) * k \tag{3}$$

n and m is filter size
l is feature maps as input
k is feature maps as outputs

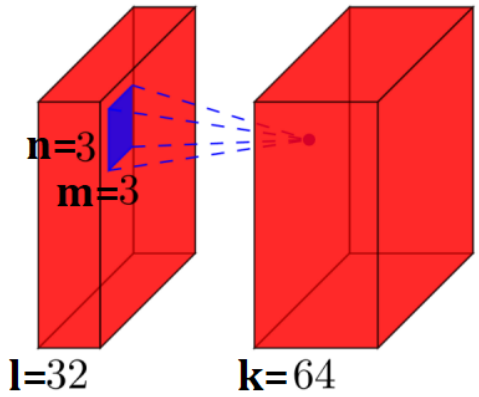

Figure 6: Convolution Layer

Learnable Parameters For CNN

3. Pooling layers: No Parameter learning
4. Fully-connected layers:

$$\text{Number of parameter} = (n+1) * m \tag{4}$$

n is the number inputs
m is the numbers outputs

5. Output layer:

$$\text{Number of parameter} = (n+1) * m \tag{5}$$

n is the number inputs
m is the numbers outputs

Architecture of LexNet-5

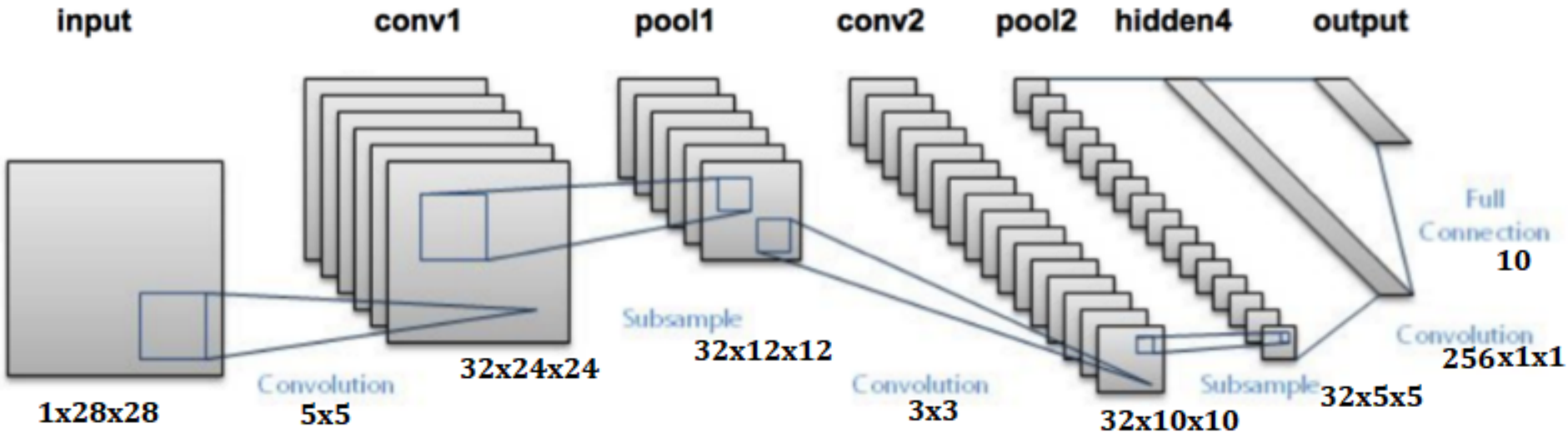

Figure 7: LexNet-5

Learnable Parameters For LexNet-5

Name	Size	Formula	Parameters
0 Input	1x28x28	-	0
1 Convolution (28-(5-1))=24	32x24x24	$(n \times m \times l+1) \times k$	(5x5x1+1)x32=832
2 Maxpool	32x12x12	-	0
3 Convolution (12-(3-1))=10	32x10x10	$(n \times m \times l+1) \times k$	(3x3x32+1)x32=9248
4 Maxpool	32x5x5	-	0
5 Fully Connected	256	$(n+1) \times m$	(32x5x5+1)x256=205056
6 Output	10	$(n+1) \times m$	(256+1)x10=2570

Table 1: Learnable Parameters For LexNet-5

AlexNet Architecture

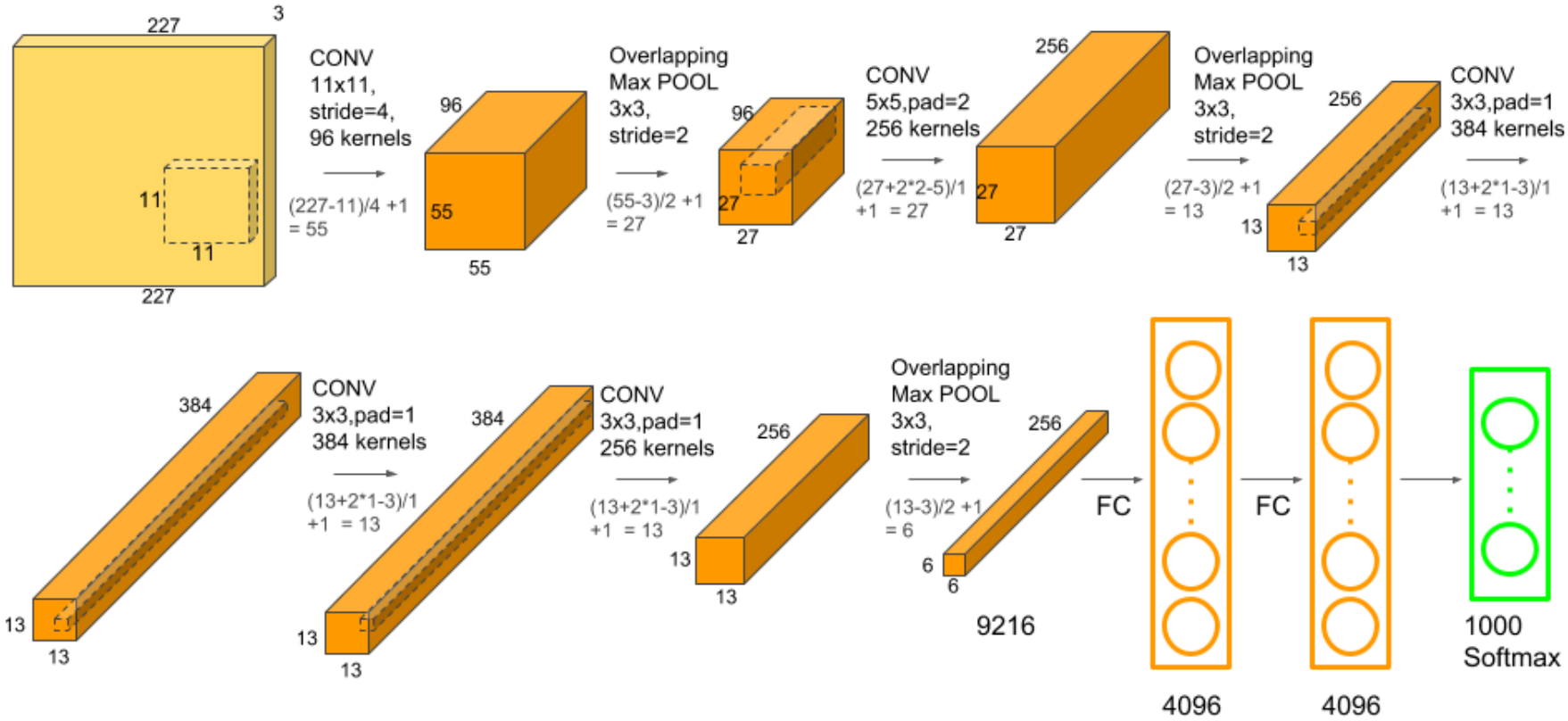

Figure 8: AlexNet Architecture

source : www.learnopencv.com/understanding-alexnet

AlexNet Architecture

1. AlexNet consists of 5 Convolutional Layers and 3 Fully Connected Layers.
2. The first two Convolutional layers are followed by the Overlapping Max Pooling layers.
3. The third, fourth and fifth convolutional layers are connected directly the output of which goes into a series of two fully connected layers.
4. The second fully connected layer feeds into a softmax classifier with 1000 class labels.

Size / Operation	Filter	Depth	Stride	Padding	Number of Parameter
3* 227 * 227					
Conv1	11 * 11	96	4		(11*11*3 + 1) * 96=34944
96 * 55 * 55					
Max Pooling	3 * 3		2		
96 * 27 * 27					
Conv2	5 * 5	256	1	2	(5 * 5 * 96 + 1) * 256=614656
256 * 27 * 27					
Max Pooling	3 * 3		2		
256 * 13 * 13					
Conv3	3 * 3	384	1	1	(3 * 3 * 256 + 1) * 384=885120
384 * 13 * 13					
Conv4	3 * 3	384	1	1	(3 * 3 * 384 + 1) * 384=1327488
384 * 13 * 13					
Conv5	3 * 3	256	1	1	(3 * 3 * 384 + 1) * 256=884992
256 * 13 * 13					
Max Pooling	3 * 3		2		
256 * 6 * 6					
FC6					256 * 6 * 6 * 4096=37748736
4096					
FC7					4096 * 4096=16777216
4096					
FC8					4096 * 1000=4096000
1000 classes					
Overall					62369152=62.3 million
Conv VS FC	Conv:3.7million (6%) , FC: 58.6 million (94%)				

Figure 9: Learnable Parameters For AlexNet

VGG Architecture

Figure 10: Architecture of VGG

Feature Extractor and Fine-tune

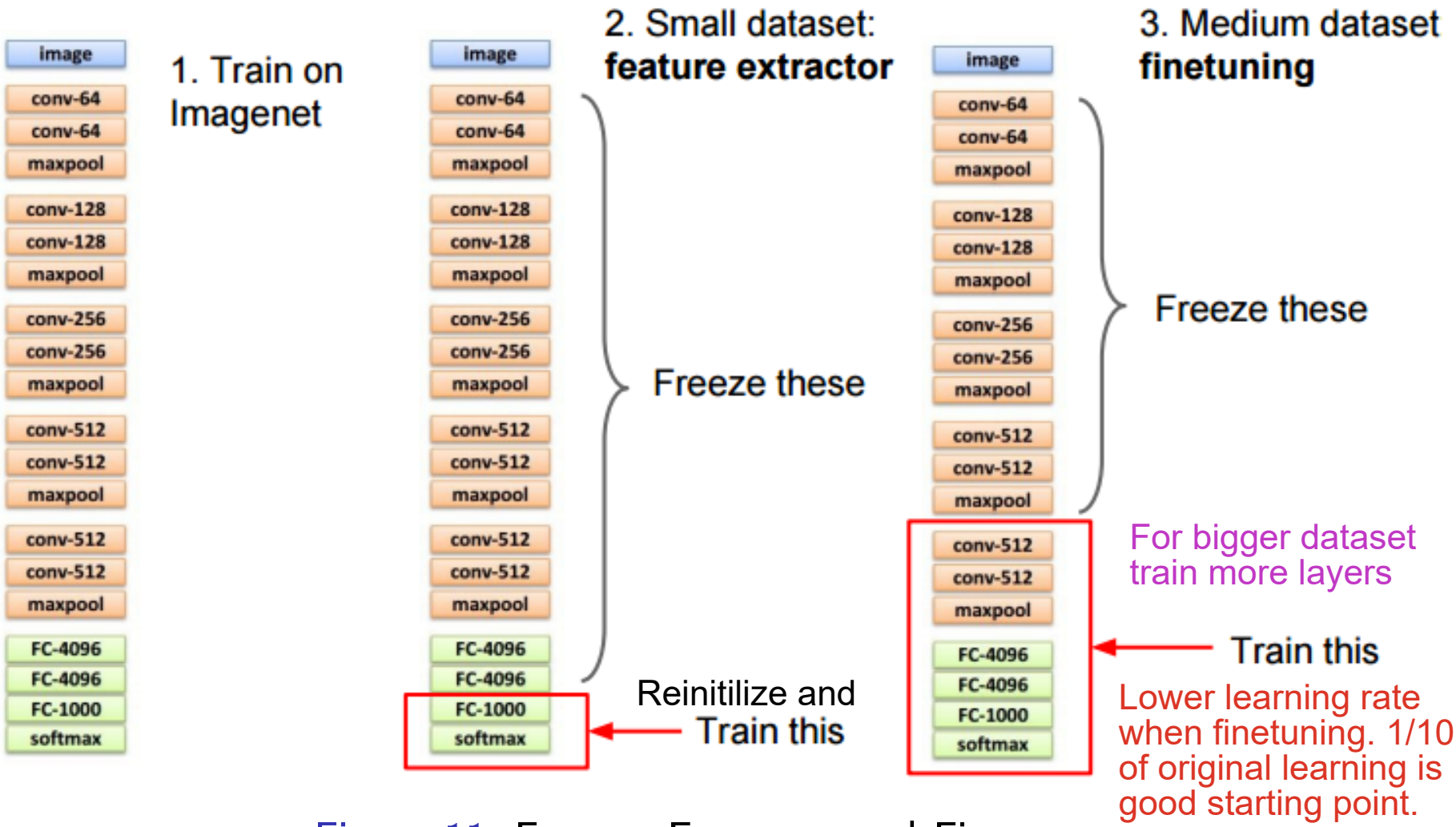

Figure 11: Feature Extractor and Fine-tune

1. www.jefkine.com
2. www.learnopencv.com