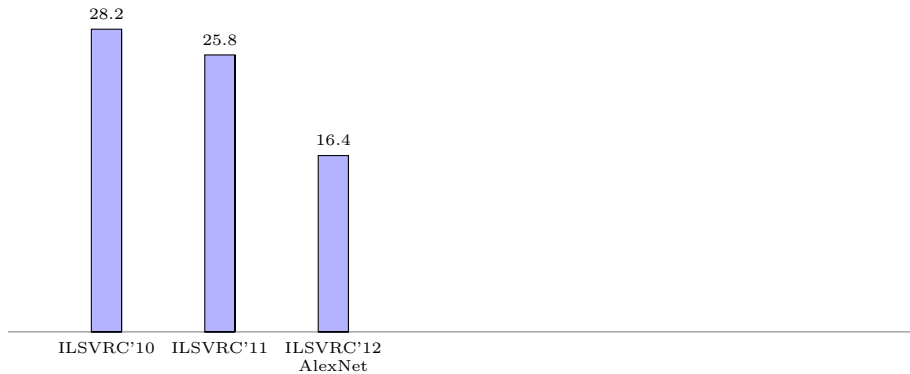
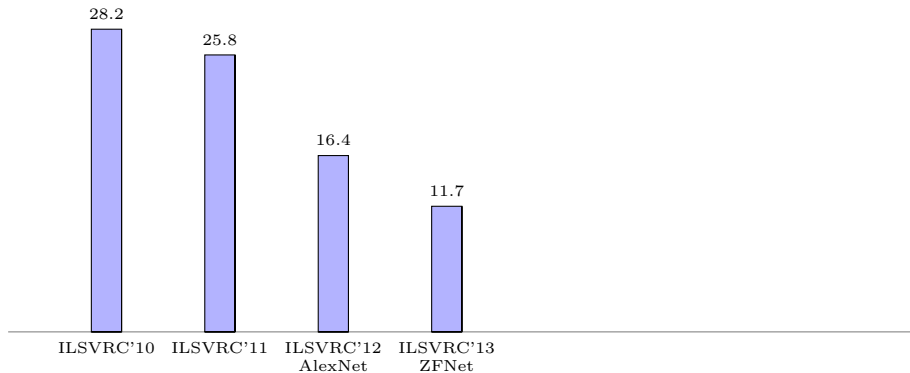
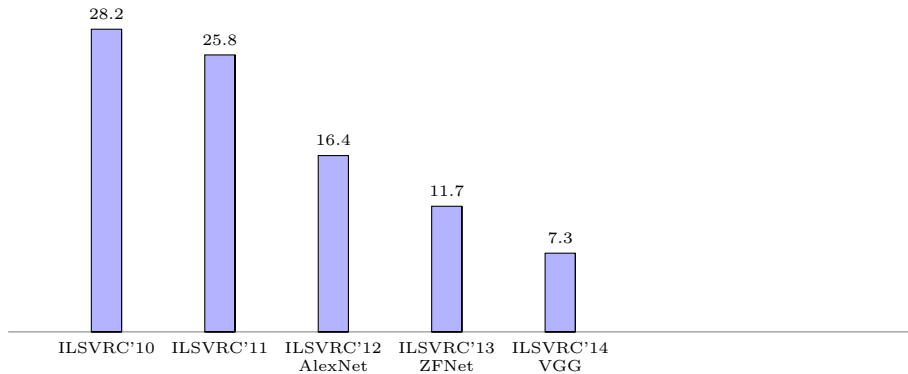
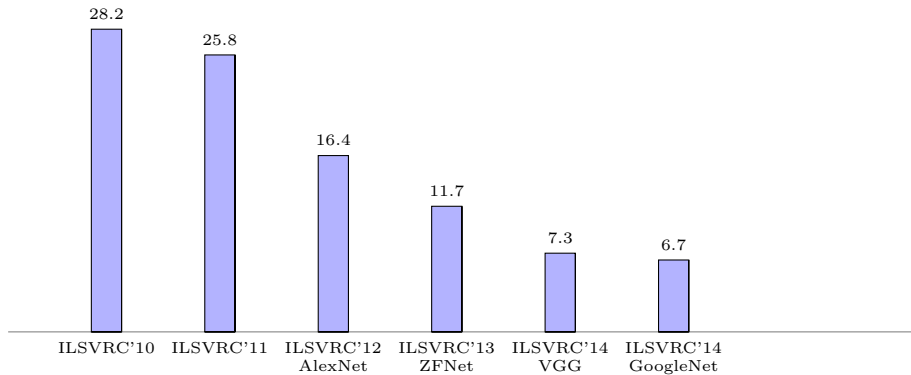
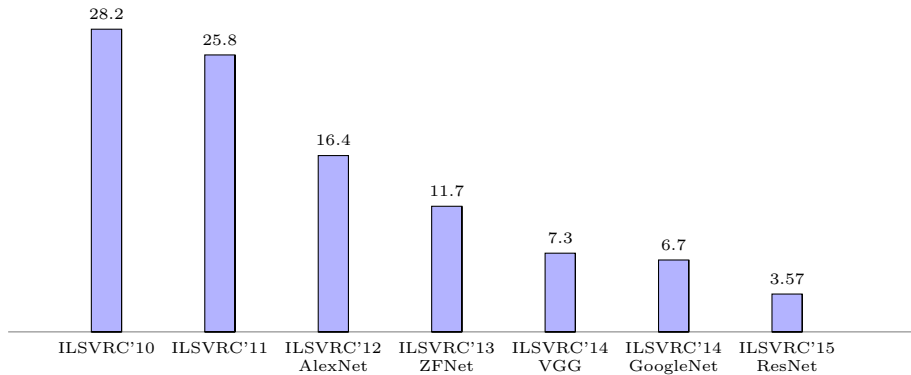


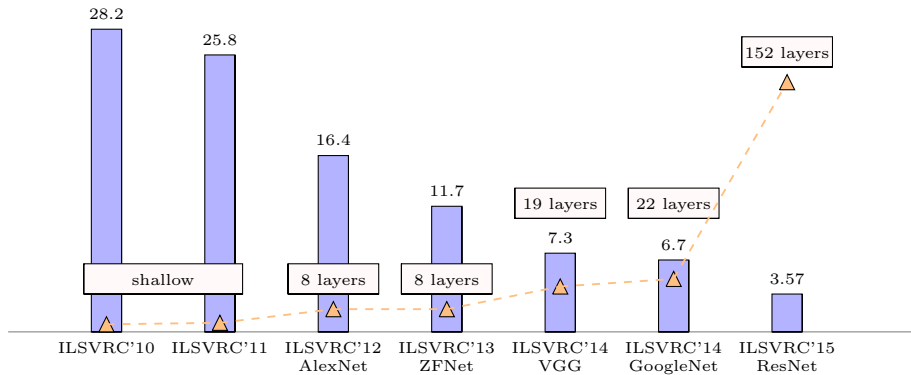


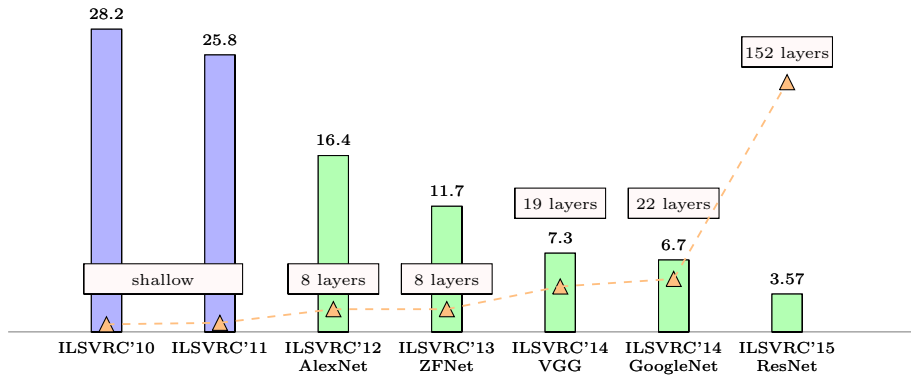






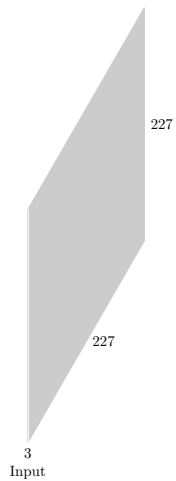


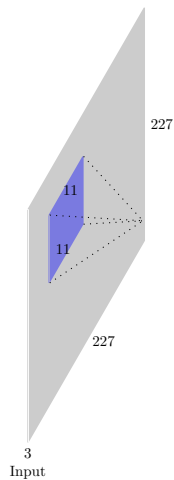




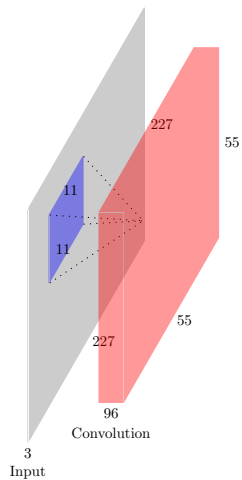
ImageNet Success Stories(roadmap for rest of the talk)

- AlexNet
- ZFNet
- VGGNet

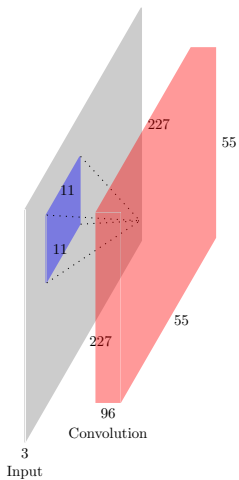




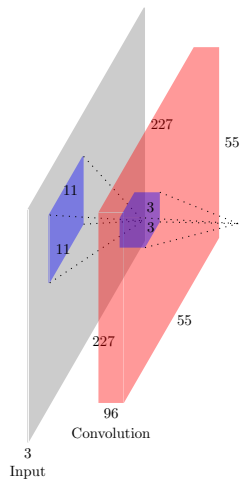
Input: $227 \times 227 \times 3$
Conv1: $K = 96, F = 11$
 $S = 4, P = 0$
Output: $W_2 = ?, H_2 = ?$
Parameters: ?



Input: $227 \times 227 \times 3$
 Conv1: $K = 96, F = 11$
 $S = 4, P = 0$
 Output: $W_2 = 55, H_2 = 55$
 Parameters: ?



Input: $227 \times 227 \times 3$
 Conv1: $K = 96, F = 11$
 $S = 4, P = 0$
 Output: $W_2 = 55, H_2 = 55$
 Parameters: $(11 \times 11 \times 3) \times 96 = 34K$

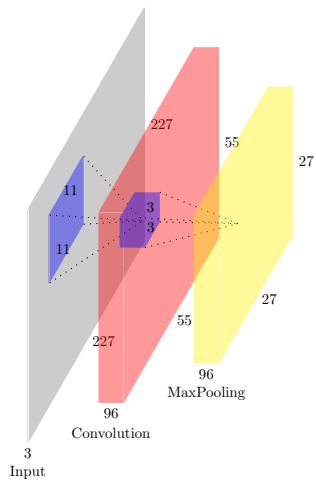


Max Pool Input: $55 \times 55 \times 96$

$F = 3, S = 2$

Output: $W_2 = ?, H_2 = ?$

Parameters: ?

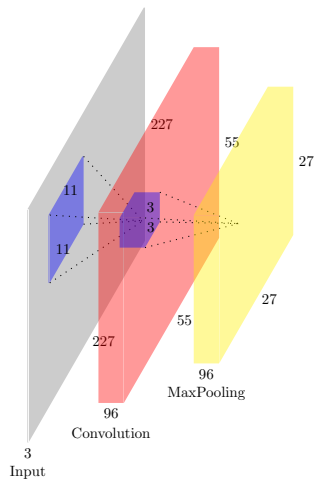


Max Pool Input: $55 \times 55 \times 96$

$F = 3, S = 2$

Output: $W_2 = 27, H_2 = 27$

Parameters: ?

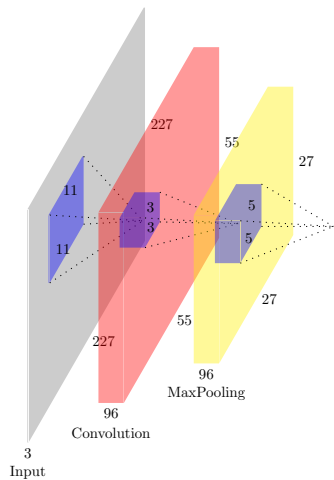


Max Pool Input: $55 \times 55 \times 96$

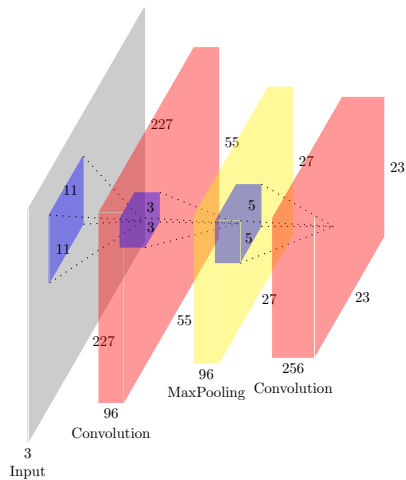
$F = 3, S = 2$

Output: $W_2 = 27, H_2 = 27$

Parameters: 0



Input: $27 \times 27 \times 96$
 Conv1: $K = 256, F = 5$
 $S = 1, P = 0$
 Output: $W_2 = ?, H_2 = ?$
 Parameters: ?



Input: $27 \times 27 \times 96$

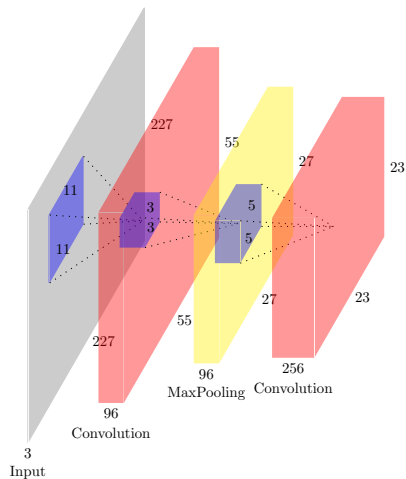
Conv1: $K = 256, F = 5$

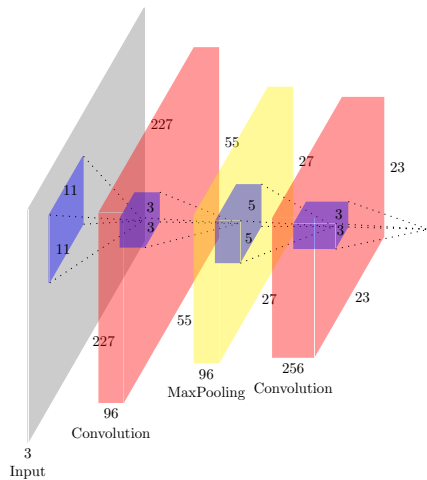
$S = 1, P = 0$

Output: $W_2 = 23, H_2 = 23$

Parameters: ?

Input: $27 \times 27 \times 96$
 Conv1: $K = 256, F = 5$
 $S = 1, P = 0$
 Output: $W_2 = 23, H_2 = 23$
 Parameters: $(5 \times 5 \times 96) \times 256 = 0.6M$



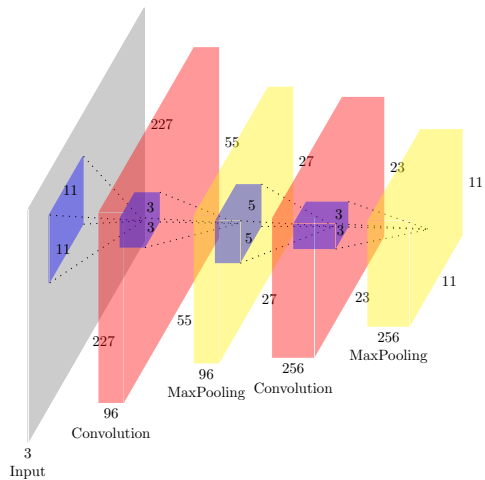


Max Pool Input: $23 \times 23 \times 256$

$F = 3, S = 2$

Output: $W_2 = ?$, $H_2 = ?$

Parameters: ?

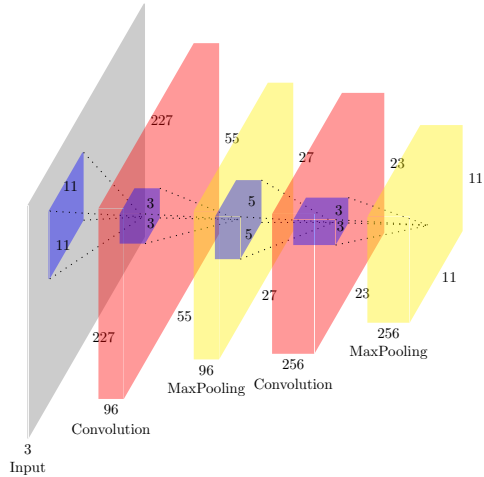


Max Pool Input: $23 \times 23 \times 256$

$F = 3, S = 2$

Output: $W_2 = 11, H_2 = 11$

Parameters: ?

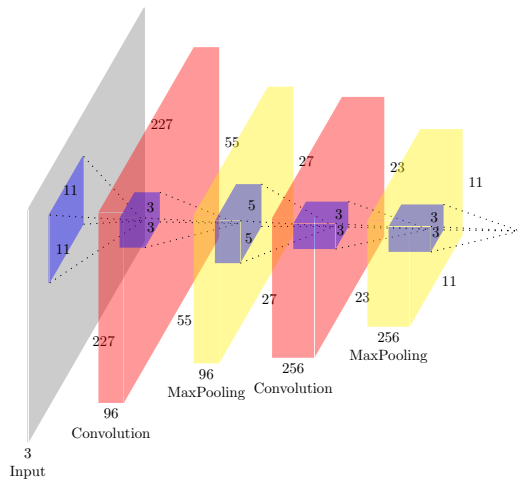


Max Pool Input: $23 \times 23 \times 256$

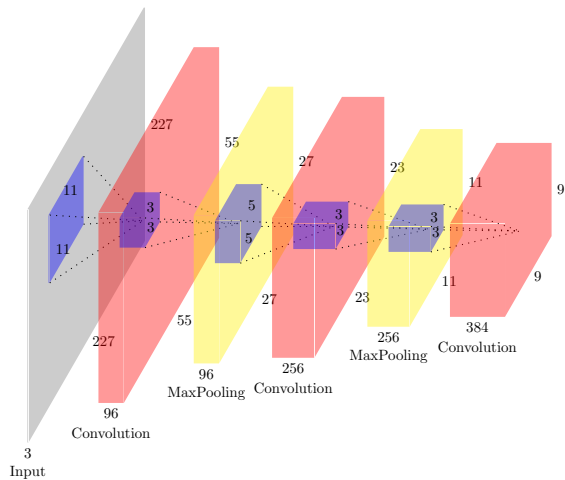
$F = 3, S = 2$

Output: $W_2 = 11, H_2 = 11$

Parameters: 0

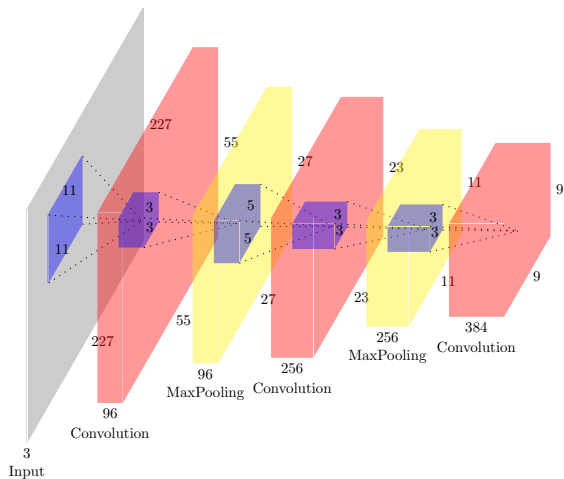


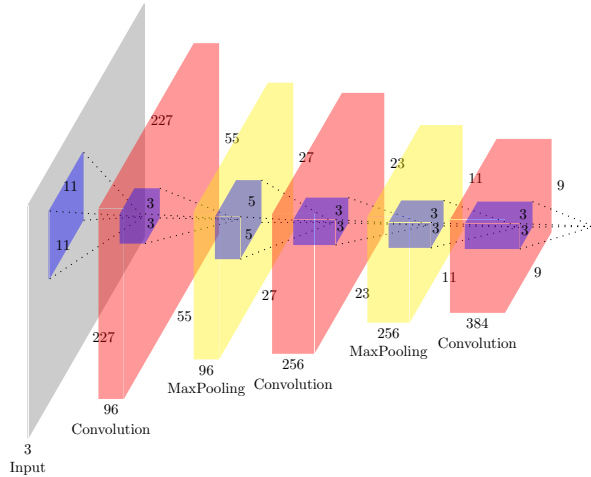
Input: $11 \times 11 \times 256$
 Conv1: $K = 384, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = ?, H_2 = ?$
 Parameters: ?



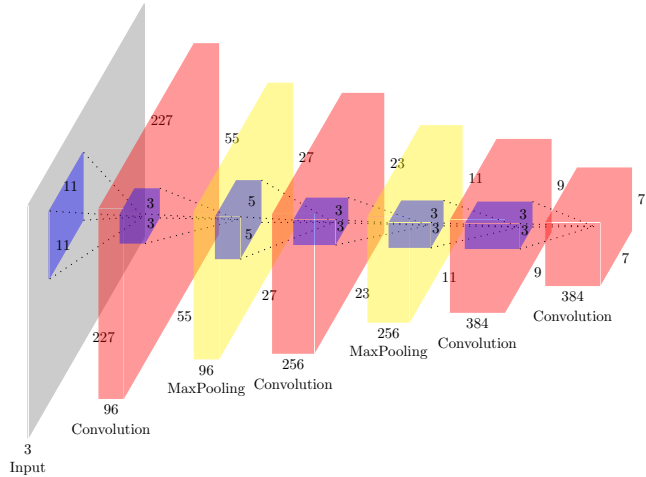
Input: $11 \times 11 \times 256$
 Conv1: $K = 384, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = 9, H_2 = 9$
 Parameters: ?

Input: $11 \times 11 \times 256$
 Conv1: $K = 384, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = 9, H_2 = 9$
 Parameters: $(3 \times 3 \times 256) \times 384 = 0.8M$



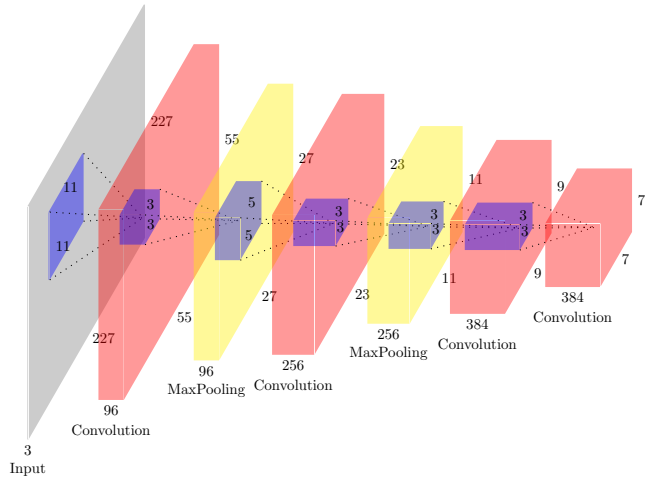


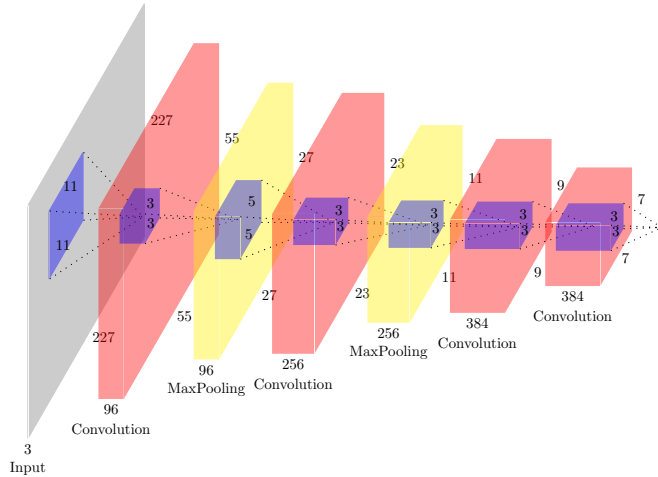
Input: $9 \times 9 \times 384$
 Conv1: $K = 384, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = ?, H_2 = ?$
 Parameters: ?



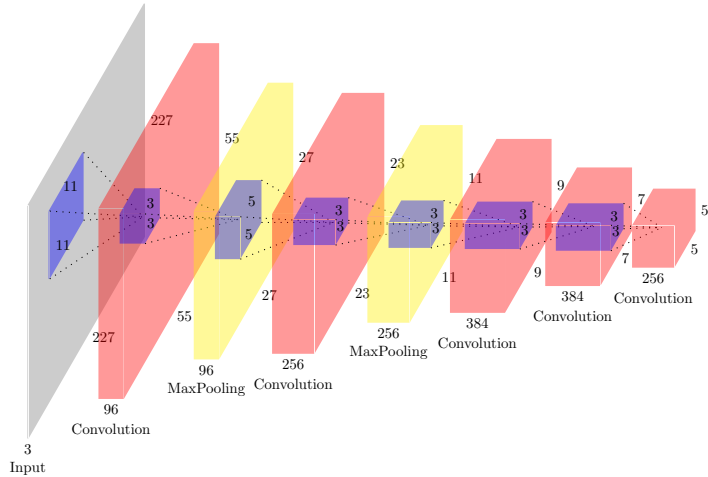
Input: $9 \times 9 \times 384$
 Conv1: $K = 384, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = 7, H_2 = 7$
 Parameters: ?

Input: $9 \times 9 \times 384$
 Conv1: $K = 384, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = 7, H_2 = 7$
 Parameters: $(3 \times 3 \times 384) \times 384 = 1.327M$



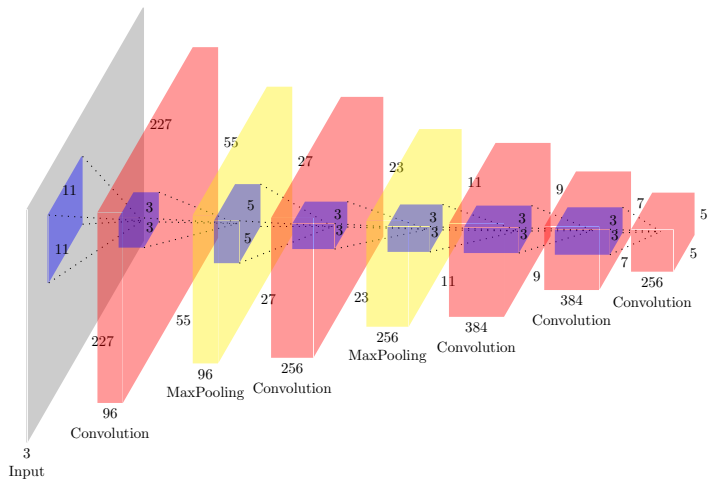


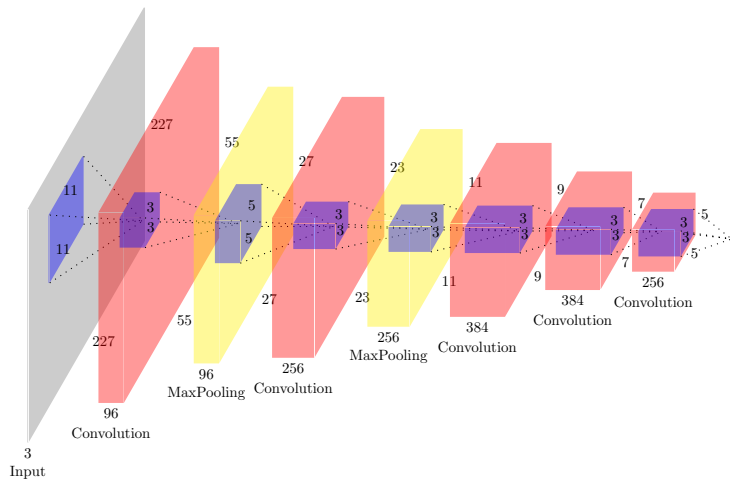
Input: $7 \times 7 \times 384$
 Conv1: $K = 256, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = ?, H_2 = ?$
 Parameters: ?



Input: $7 \times 7 \times 384$
 Conv1: $K = 256, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = 5, H_2 = 5$
 Parameters: ?

Input: $7 \times 7 \times 384$
 Conv1: $K = 256, F = 3$
 $S = 1, P = 0$
 Output: $W_2 = 5, H_2 = 5$
 Parameters: $(3 \times 3 \times 384) \times 256 = 0.8M$



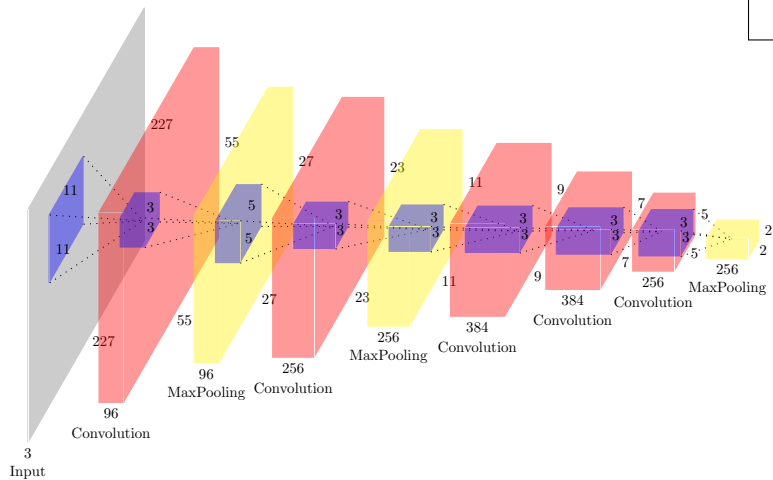


Max Pool Input: $5 \times 5 \times 256$

$F = 3, S = 2$

Output: $W_2 = ?$, $H_2 = ?$

Parameters: ?

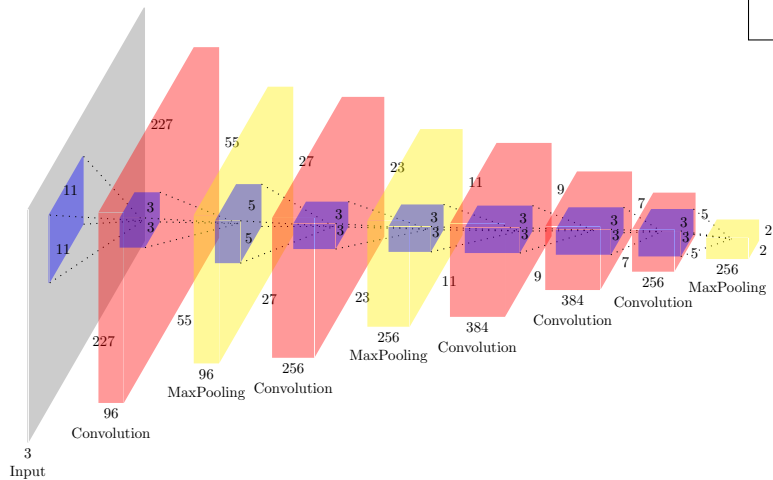


Max Pool Input: $5 \times 5 \times 256$

$F = 3, S = 2$

Output: $W_2 = 2, H_2 = 2$

Parameters: ?

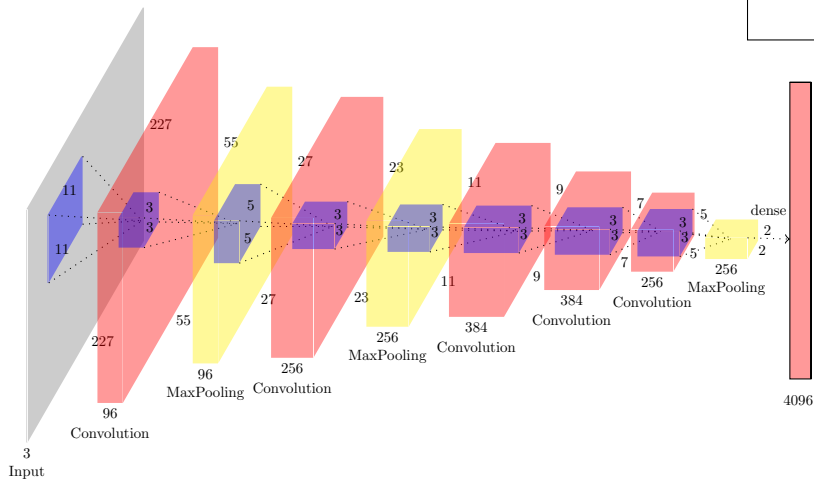


Max Pool Input: $5 \times 5 \times 256$

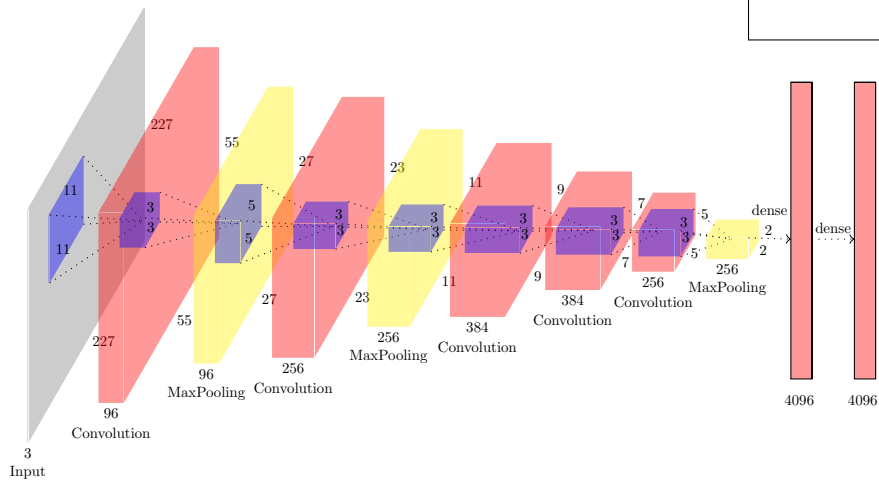
$F = 3, S = 2$

Output: $W_2 = 2, H_2 = 2$

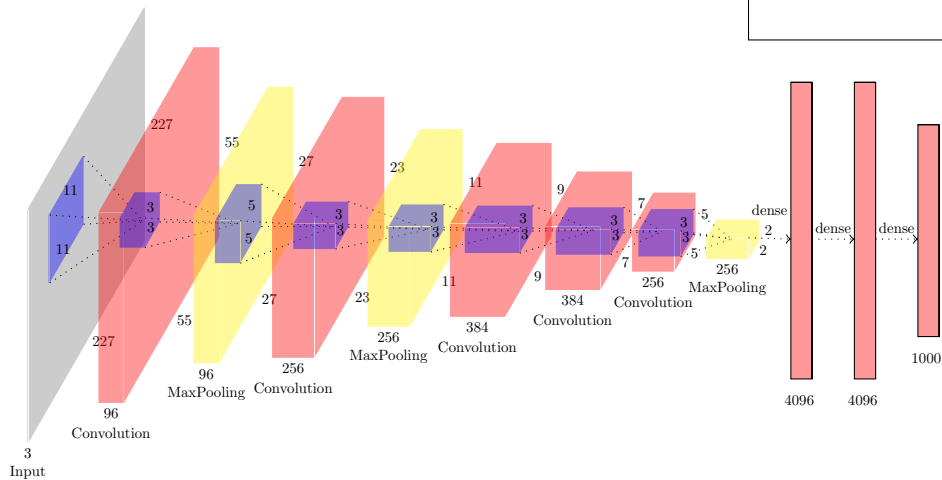
Parameters: 0



FC1
Parameters: $(2 \times 2 \times 256) \times 4096 = 4M$

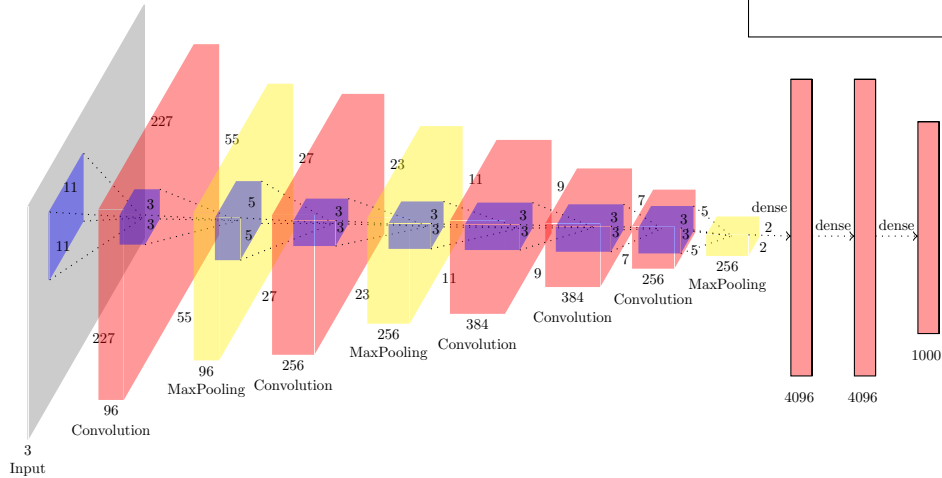


FC1
Parameters: $4096 \times 4096 = 16M$

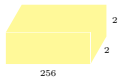


FC1
Parameters: $4096 \times 1000 = 4M$

Total Parameters: 27.55M

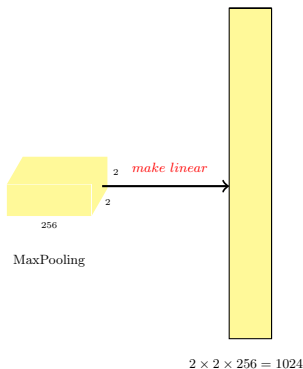


- Let us look at the connections in the fully connected layers in more detail

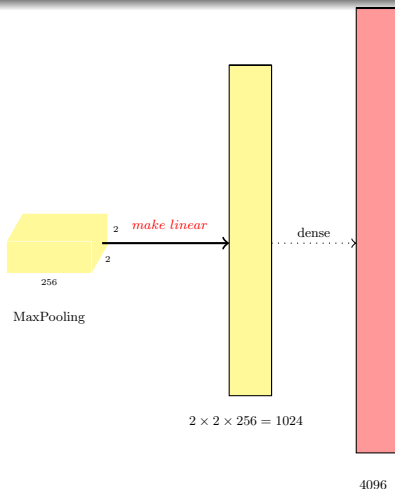


MaxPooling

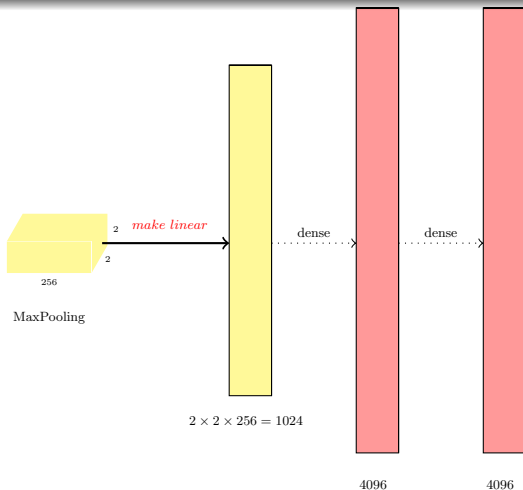
- Let us look at the connections in the fully connected layers in more detail
- We will first stretch out the last conv or maxpool layer to make it a 1d vector



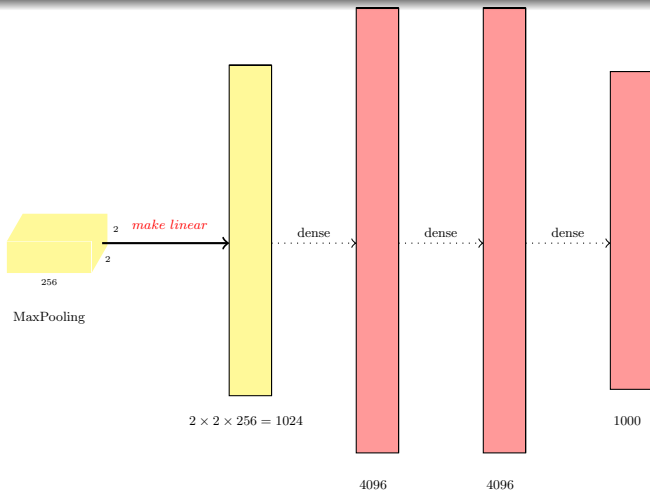
- Let us look at the connections in the fully connected layers in more detail
- We will first stretch out the last conv or maxpool layer to make it a 1d vector
- This 1d vector is then densely connected to other layers just as in a regular feedforward neural network



- Let us look at the connections in the fully connected layers in more detail
- We will first stretch out the last conv or maxpool layer to make it a 1d vector
- This 1d vector is then densely connected to other layers just as in a regular feedforward neural network

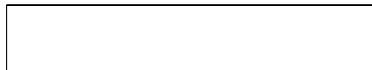
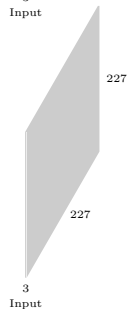
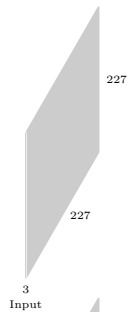


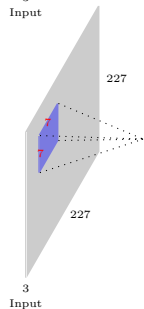
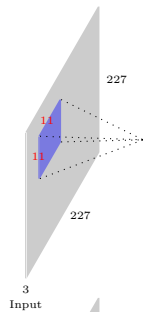
- Let us look at the connections in the fully connected layers in more detail
- We will first stretch out the last conv or maxpool layer to make it a 1d vector
- This 1d vector is then densely connected to other layers just as in a regular feedforward neural network



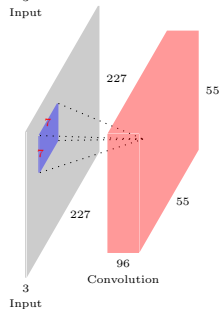
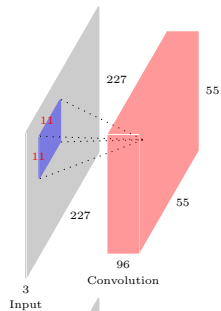
ImageNet Success Stories(roadmap for rest of the talk)

- AlexNet
- ZFNet
- VGGNet

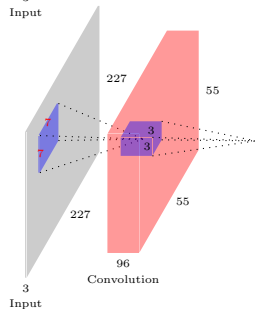
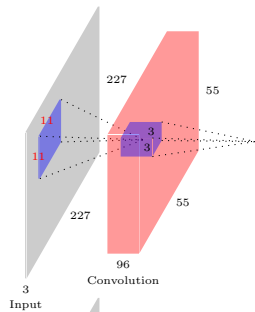




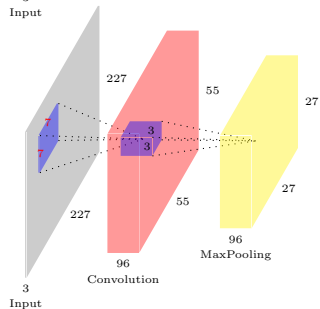
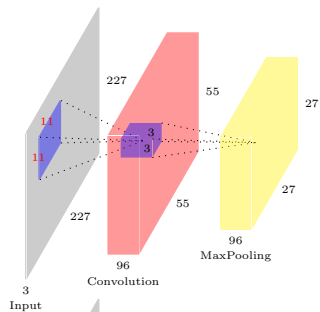
Layer1: $F = 11 \rightarrow 7$
 Difference in Parameters
 $((11^2 - 7^2) \times 3) \times 96 = 20.7K$



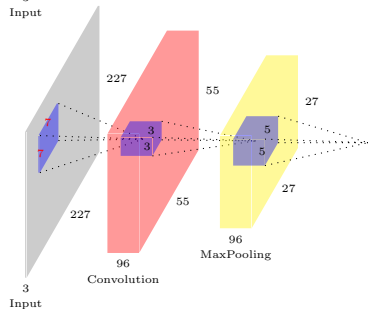
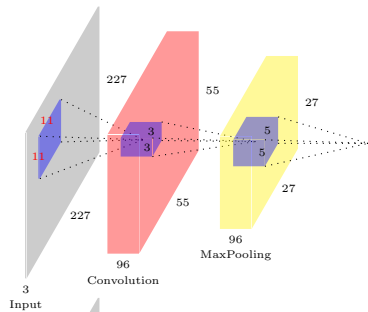
Layer1: $F = 11 \rightarrow 7$
 Difference in Parameters
 $((11^2 - 7^2) \times 3) \times 96 = 20.7K$



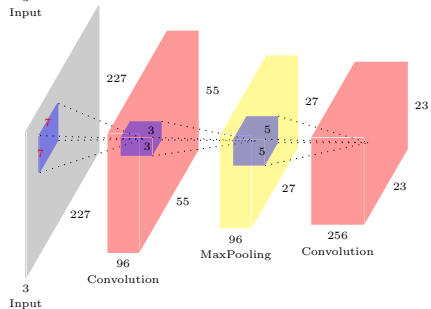
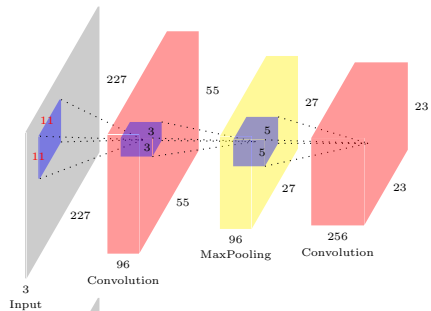
Layer2: No difference



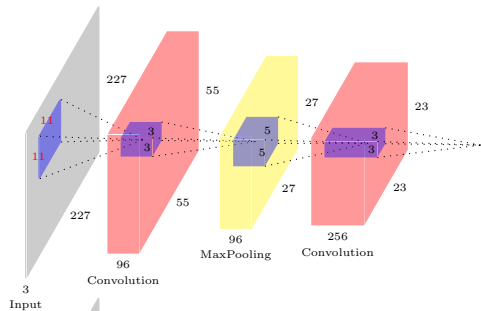
Layer2: No difference



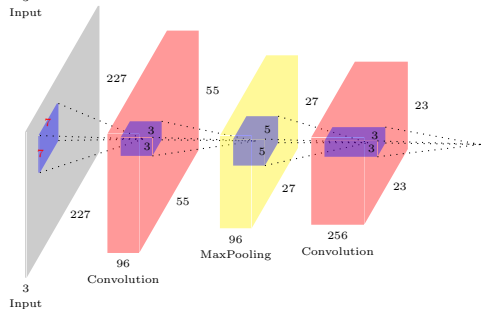
Layer3: No difference

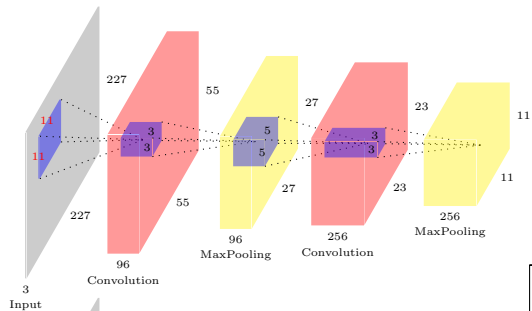


Layer3: No difference

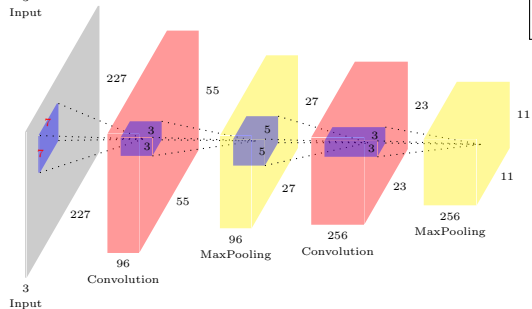


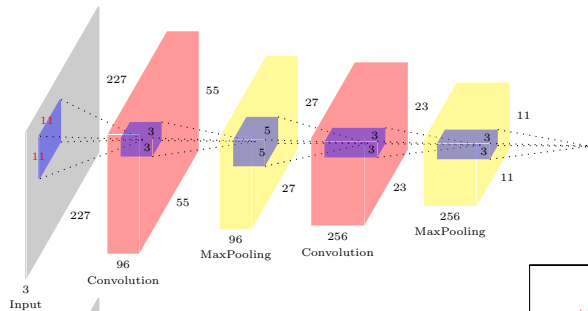
Layer4: No difference



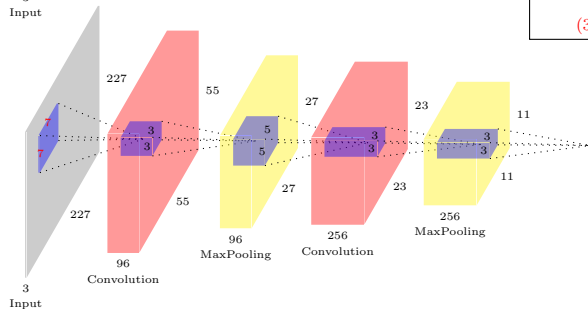


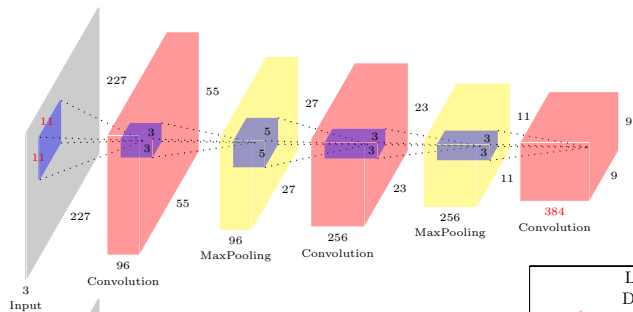
Layer4: No difference



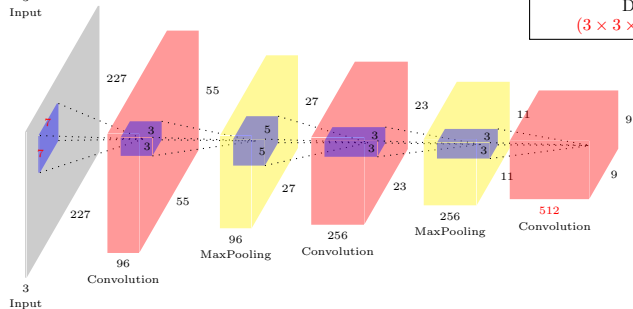


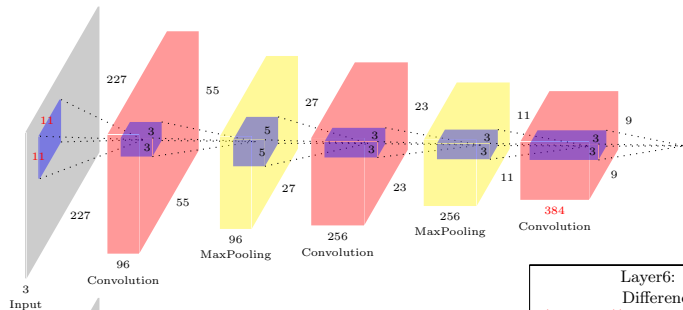
Layer5: $K = 384 \rightarrow 512$
 Difference in Parameters
 $(3 \times 3 \times 256) \times (512 - 384) = 0.29M$



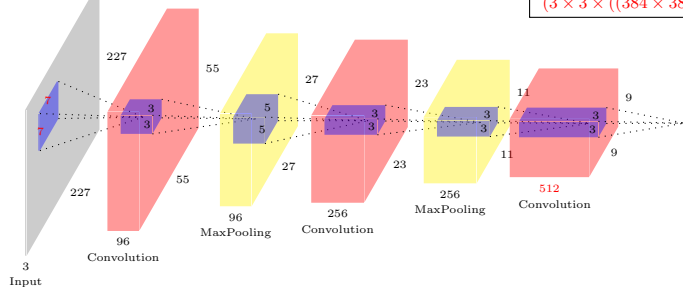


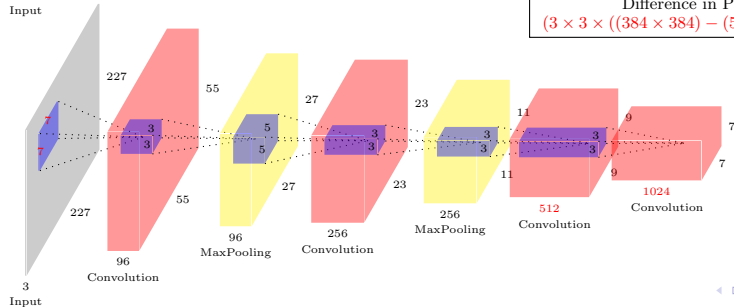
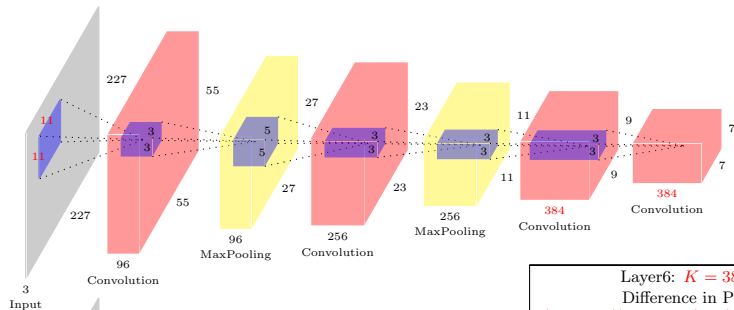
Layer5: $K = 384 \rightarrow 512$
 Difference in Parameters
 $(3 \times 3 \times 256) \times (512 - 384) = 0.29M$

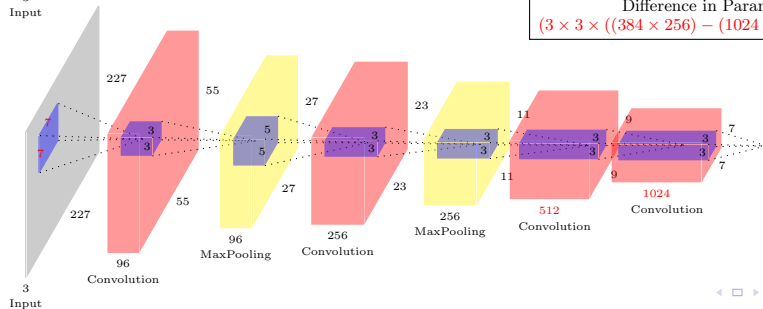
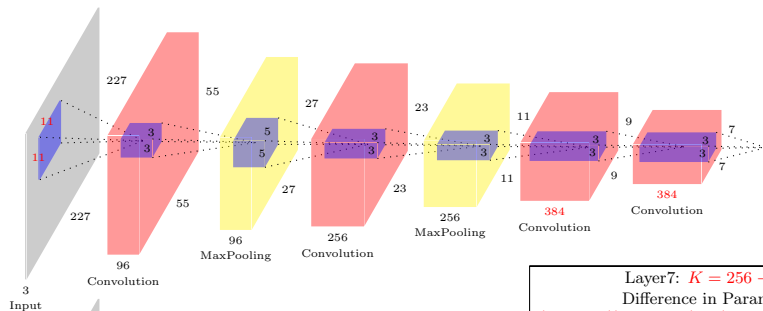


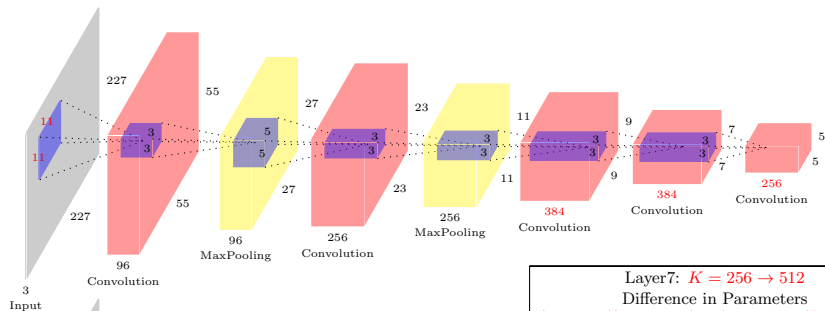


Layer6: $K = 384 \rightarrow 1024$
 Difference in Parameters
 $(3 \times 3 \times ((384 \times 384) - (512 \times 1024))) = 0.8M$

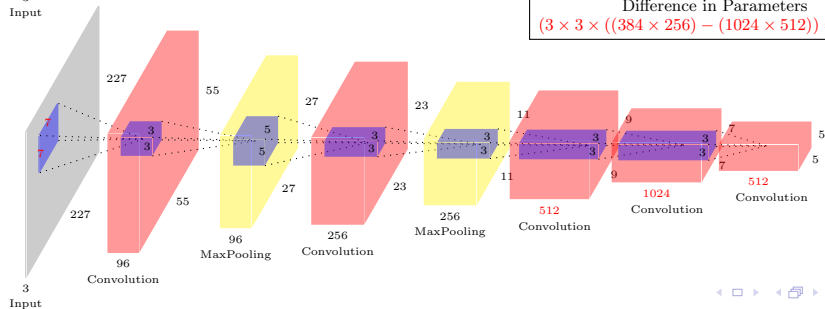


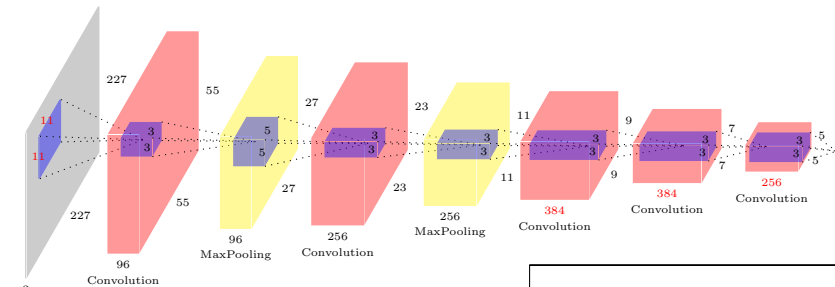




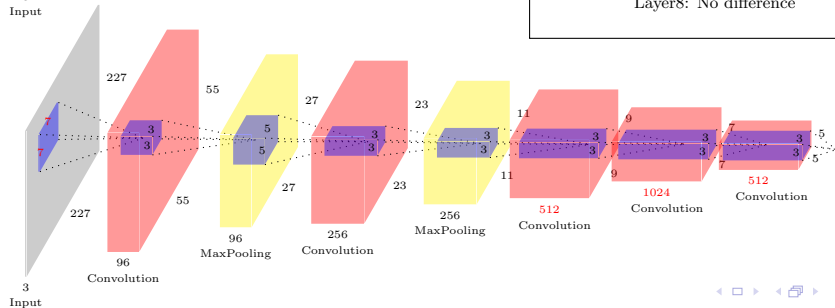


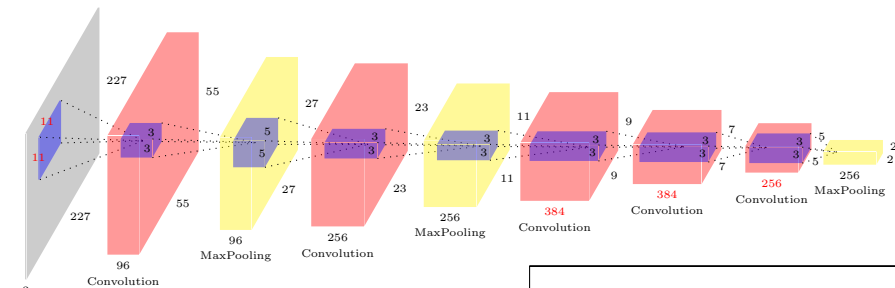
Layer7: $K = 256 \rightarrow 512$
 Difference in Parameters
 $(3 \times 3 \times ((384 \times 256) - (1024 \times 512))) = 0.36M$



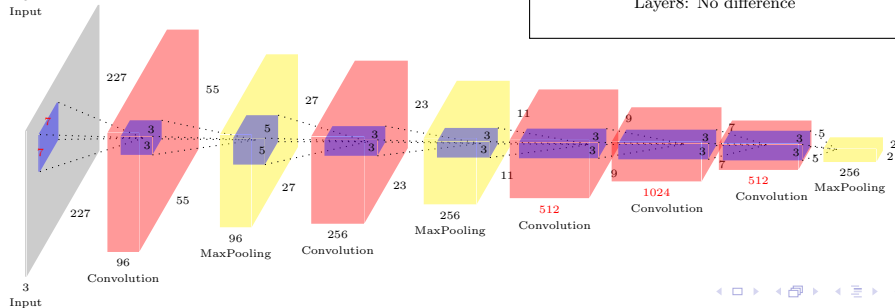


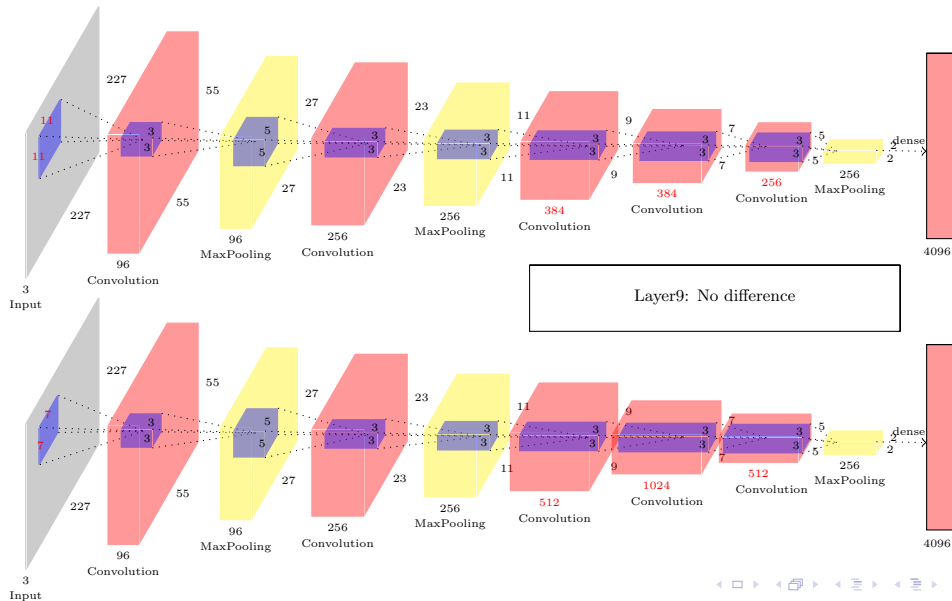
Layer8: No difference

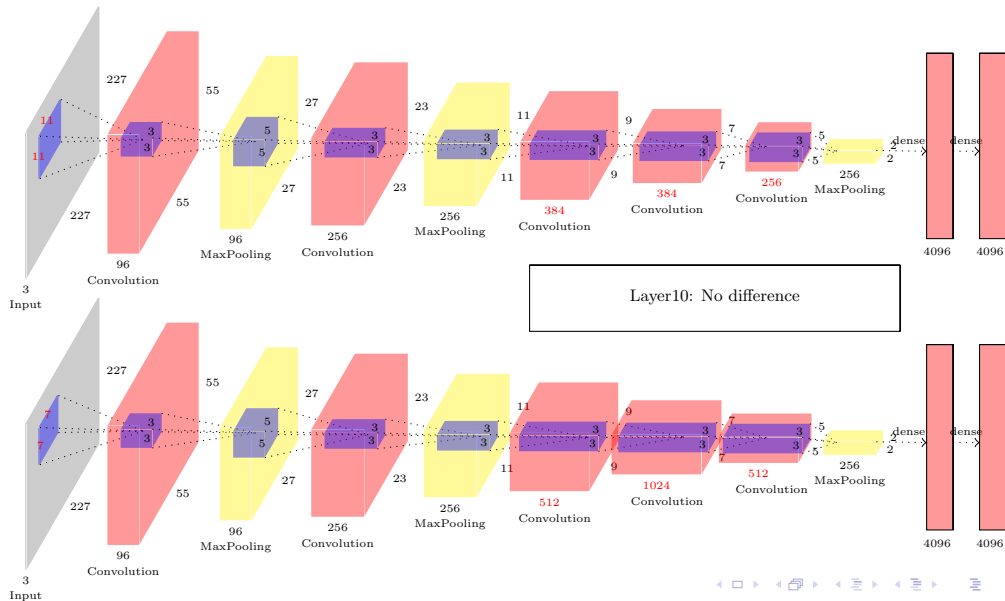


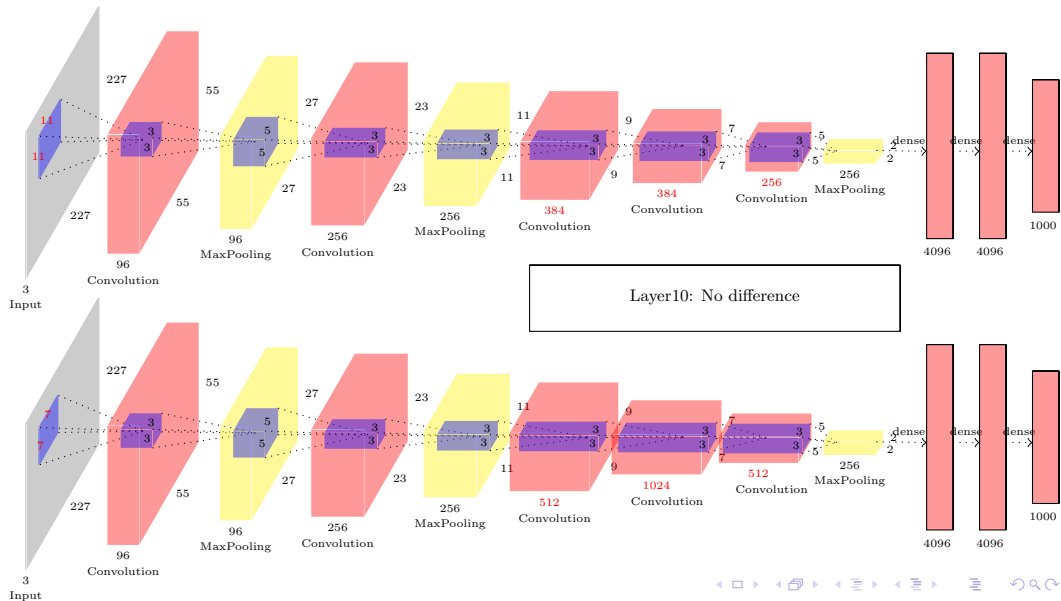


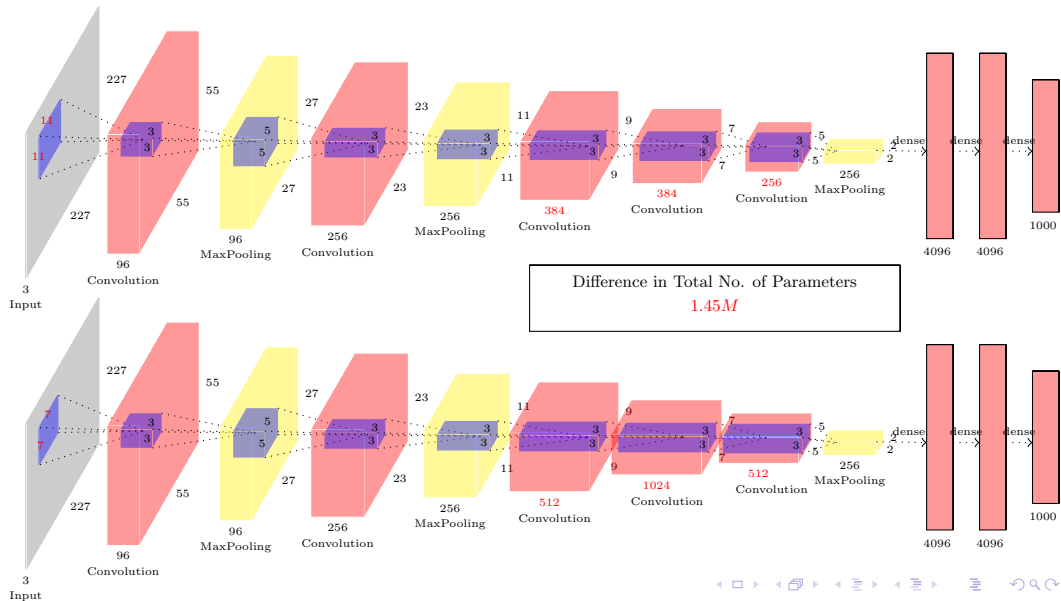
Layer8: No difference





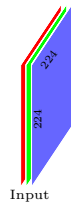


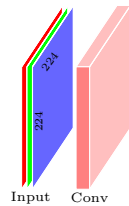


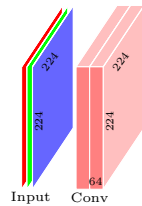


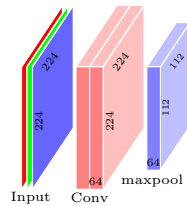
ImageNet Success Stories(roadmap for rest of the talk)

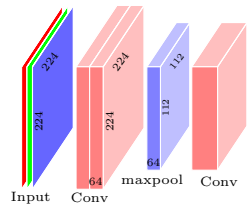
- AlexNet
- ZFNet
- VGGNet

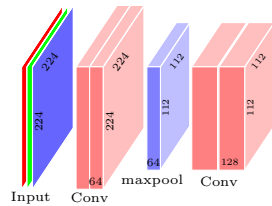


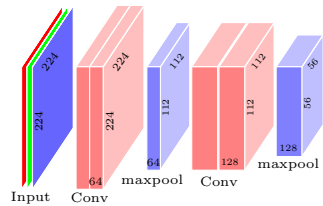


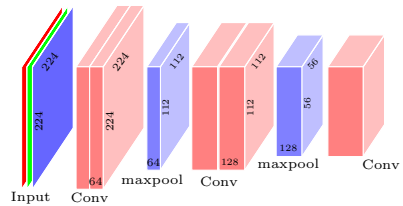


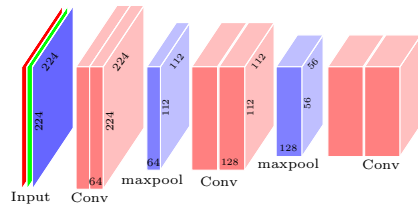


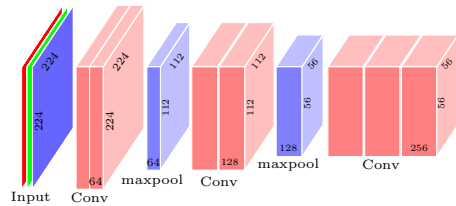


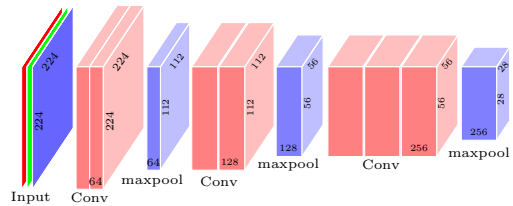


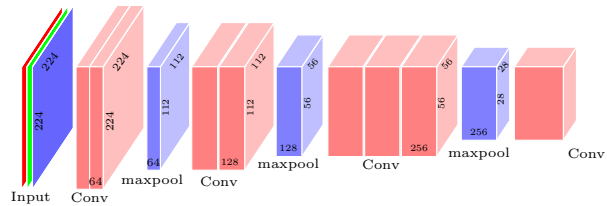


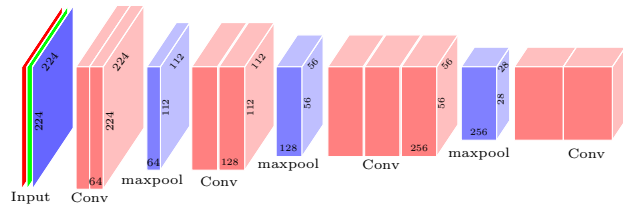


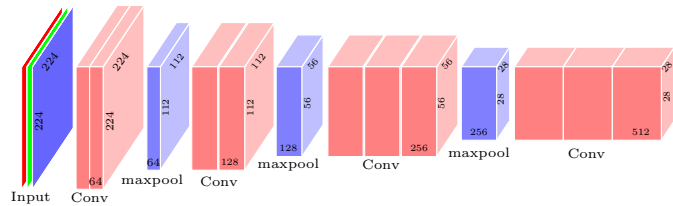


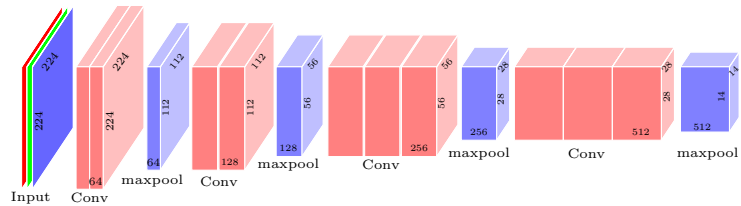


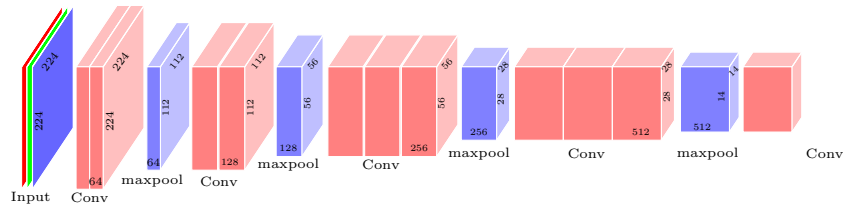


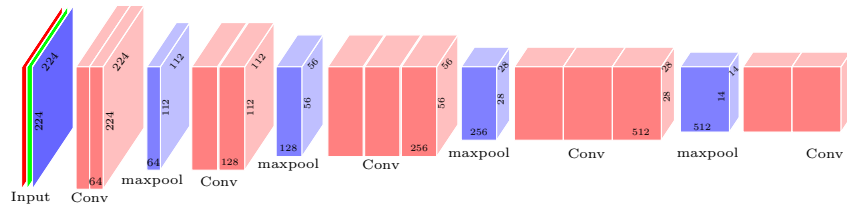


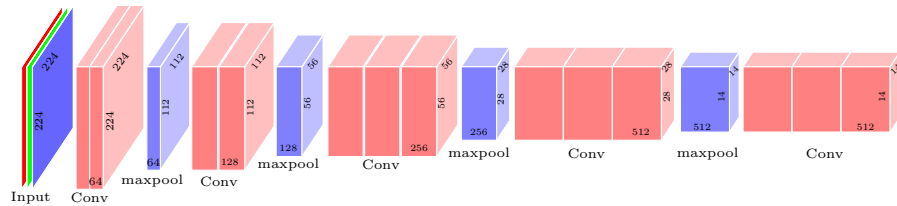


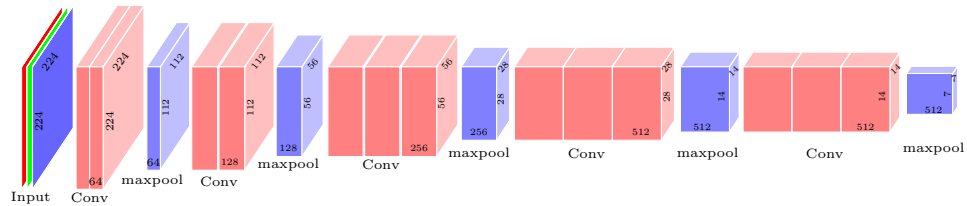


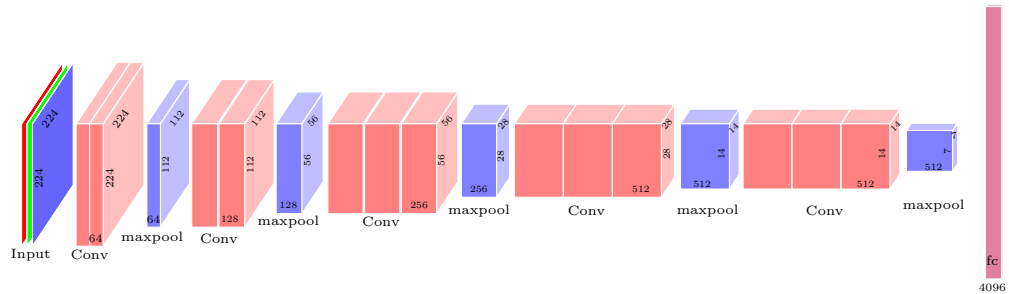


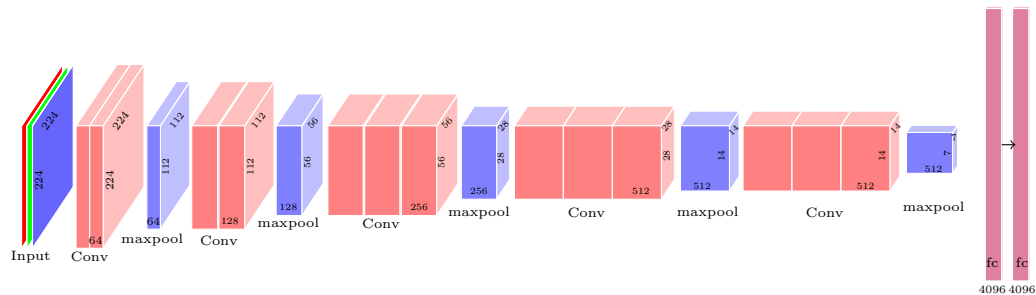


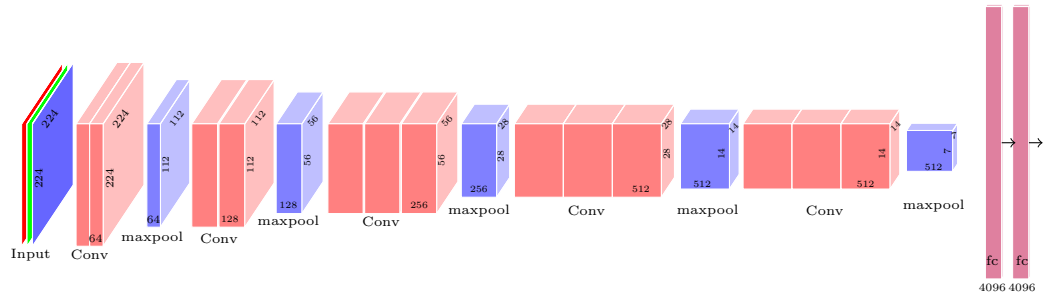


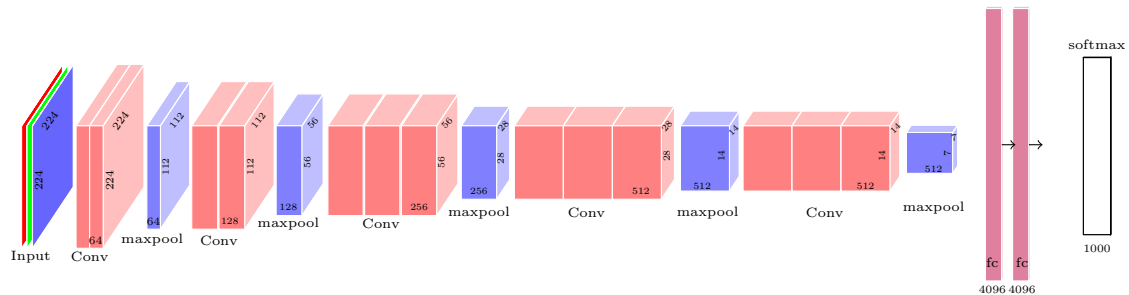


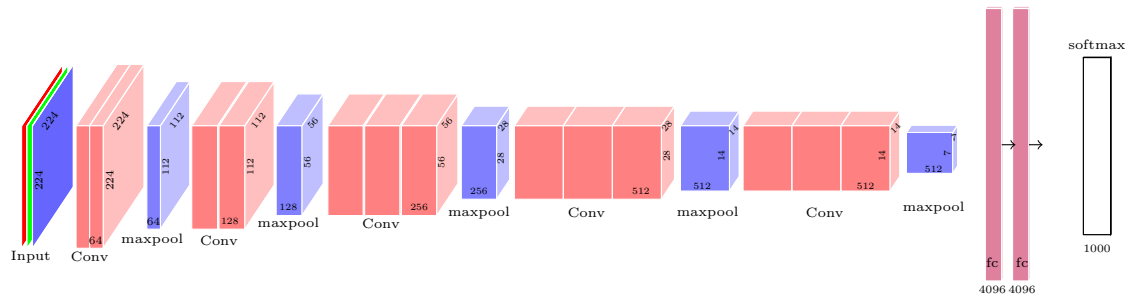




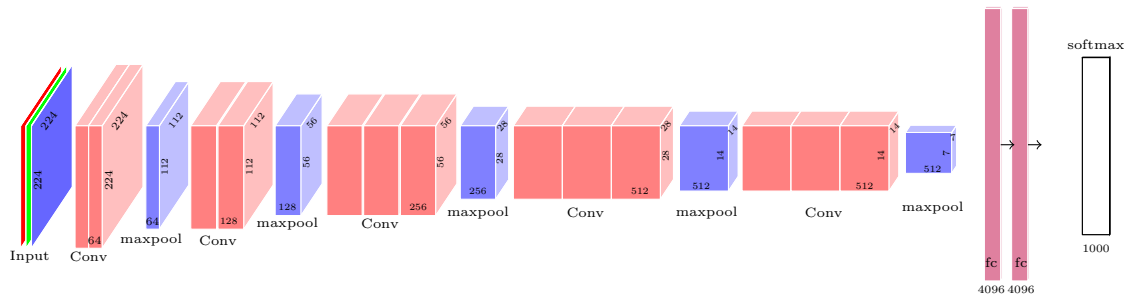




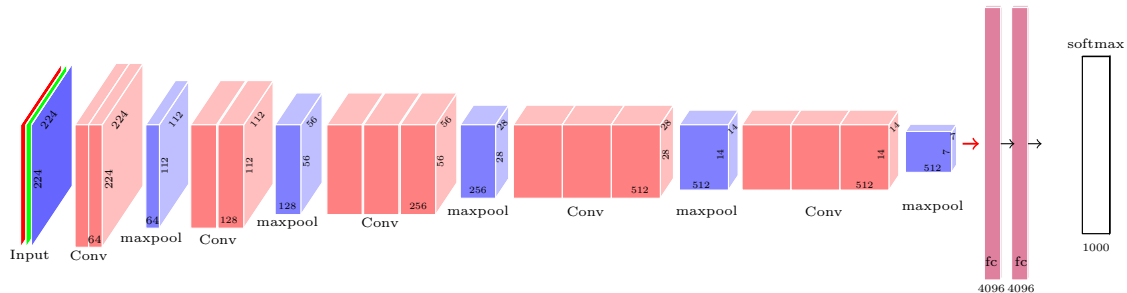




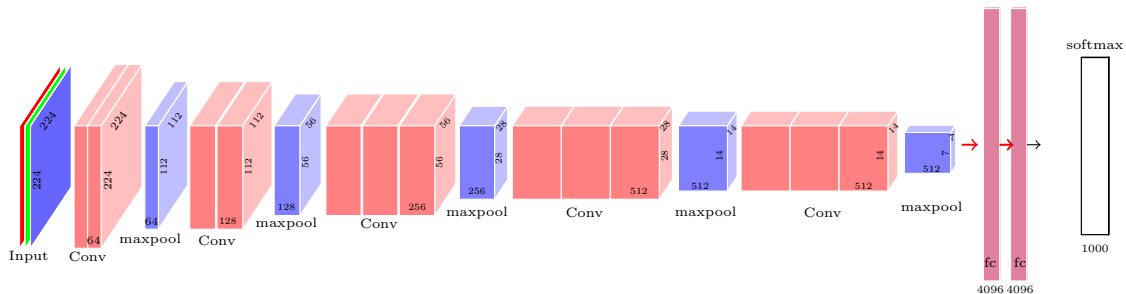
- Kernel size is 3×3 throughout



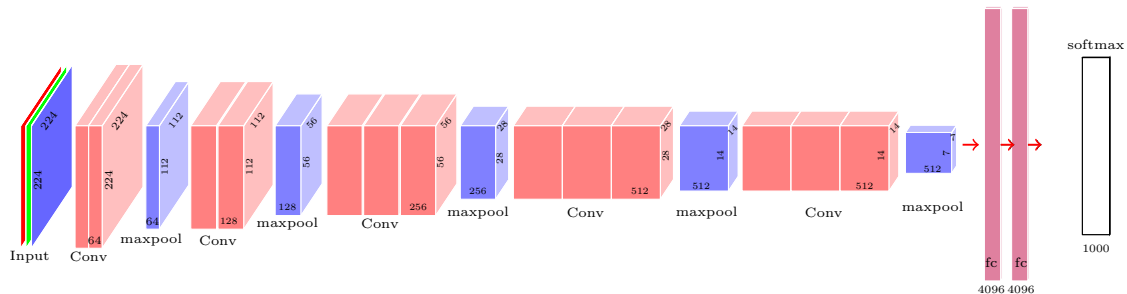
- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$



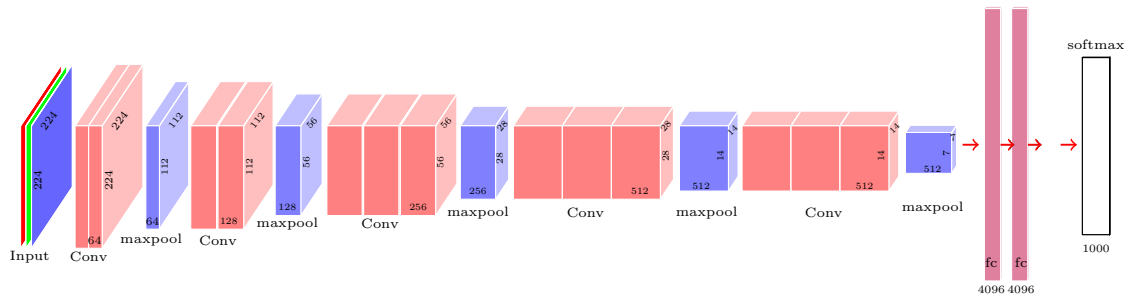
- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$
- Total Parameters in FC layers =



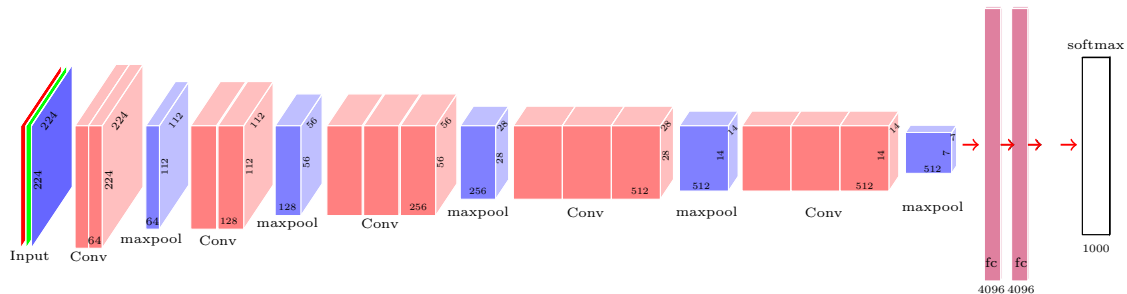
- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$
- Total Parameters in FC layers = $(512 \times 7 \times 7 \times 4096)$



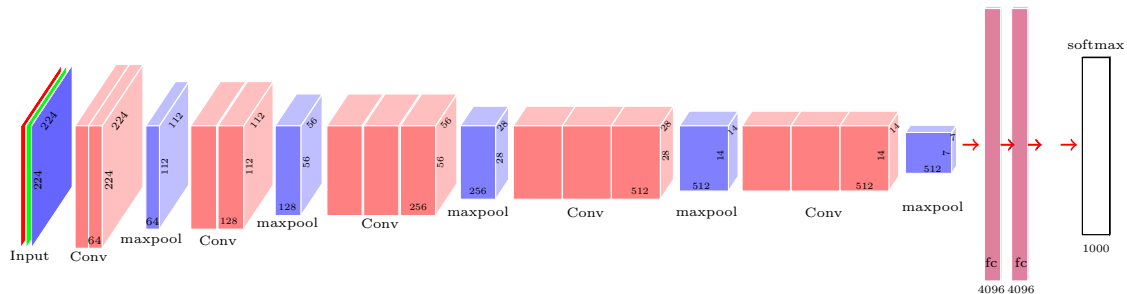
- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$
- Total Parameters in FC layers = $(512 \times 7 \times 7 \times 4096) + (4096 \times 4096)$



- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$
- Total Parameters in FC layers = $(512 \times 7 \times 7 \times 4096) + (4096 \times 4096) + (4096 \times 1024)$

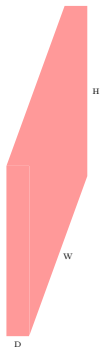


- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$
- Total Parameters in FC layers = $(512 \times 7 \times 7 \times 4096) + (4096 \times 4096) + (4096 \times 1024) = \sim 122M$

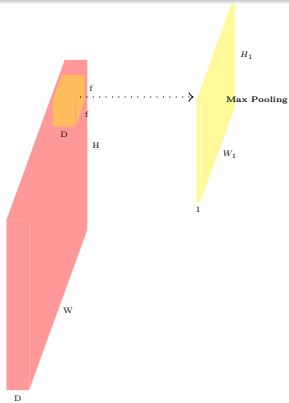


- Kernel size is 3×3 throughout
- Total parameters in non FC layers = $\sim 16M$
- Total Parameters in FC layers = $(512 \times 7 \times 7 \times 4096) + (4096 \times 4096) + (4096 \times 1024) = \sim 122M$
- Most parameters are in the first FC layer ($\sim 102M$)

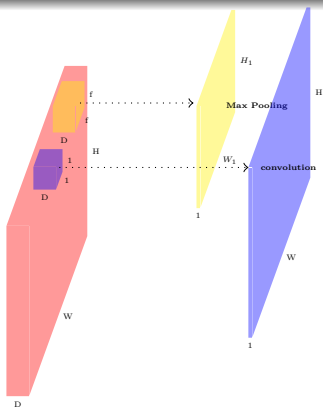
Module 11.5 : Image Classification continued (GoogLeNet and ResNet)



- Consider the output at a certain layer of a convolutional neural network



- Consider the output at a certain layer of a convolutional neural network
- After this layer we could apply a max-pooling layer



- Consider the output at a certain layer of a convolutional neural network
- After this layer we could apply a max-pooling layer
- Or a 1×1 convolution