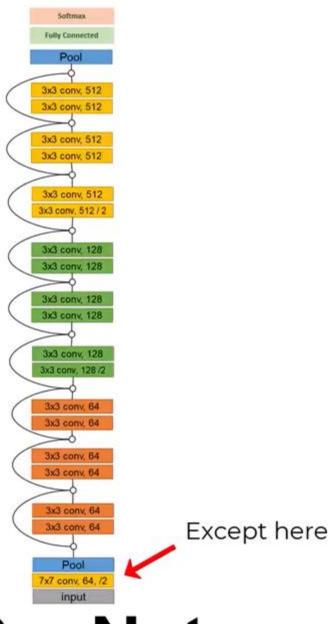
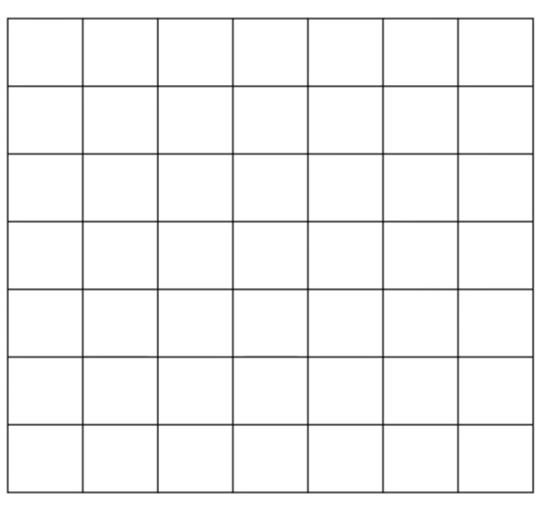
Why 3x3 conv layers!?

Softmax **Fully Connected Layers** Max Pool 3x3 conv, 512 3x3 conv, 512 3x3 conv, 512 Max Pool 3x3 conv, 512 3x3 conv, 512 3x3 conv, 512 Max Pool 3x3 conv, 256 3x3 conv, 256 3x3 conv, 256 Max Pool 3x3 conv, 128 3x3 conv, 128 Max Pool 3x3 conv, 64 3x3 conv, 64 Input Image

VGG



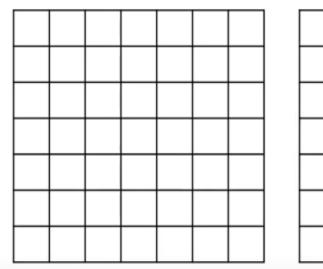
ResNet

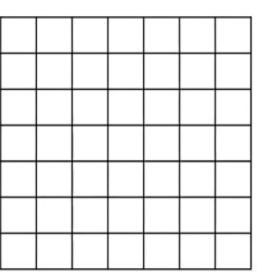


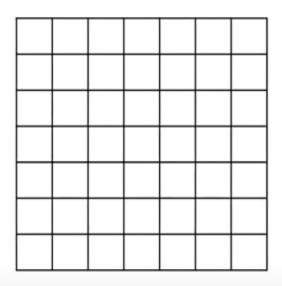
We only use 3x3 conv layers with padding = 1, stride = 1

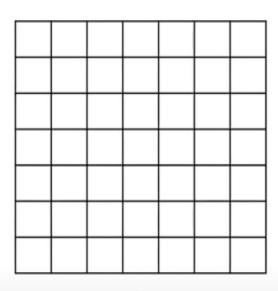
input

We only use 3x3 conv layers with padding = 1, stride = 1



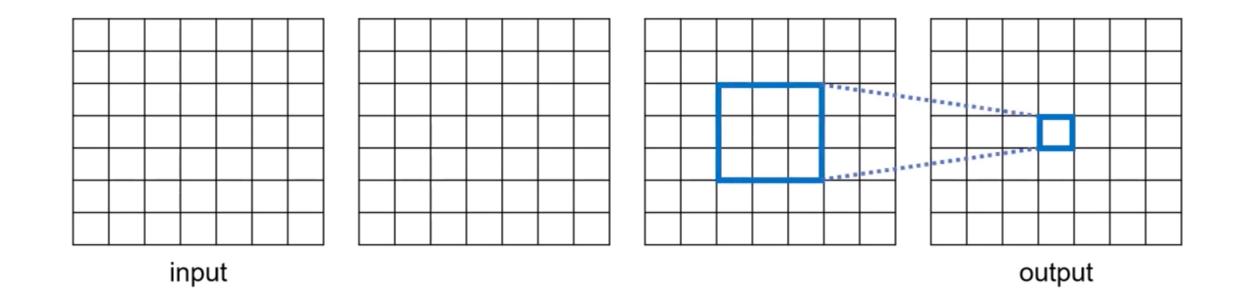


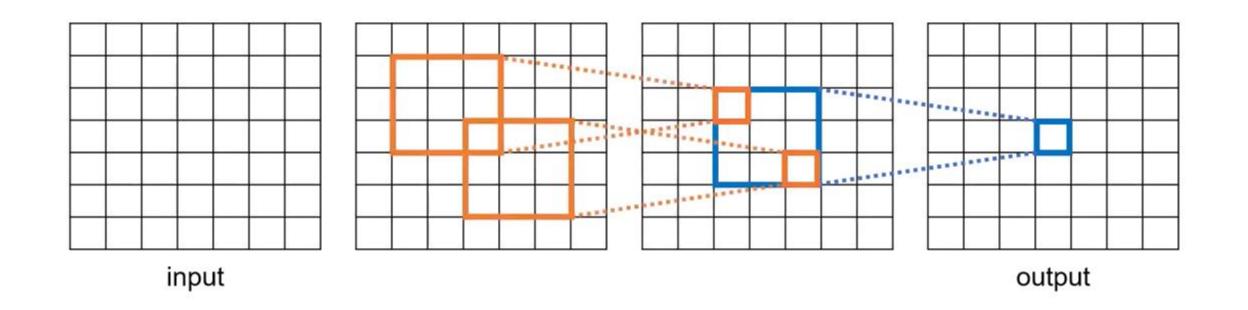


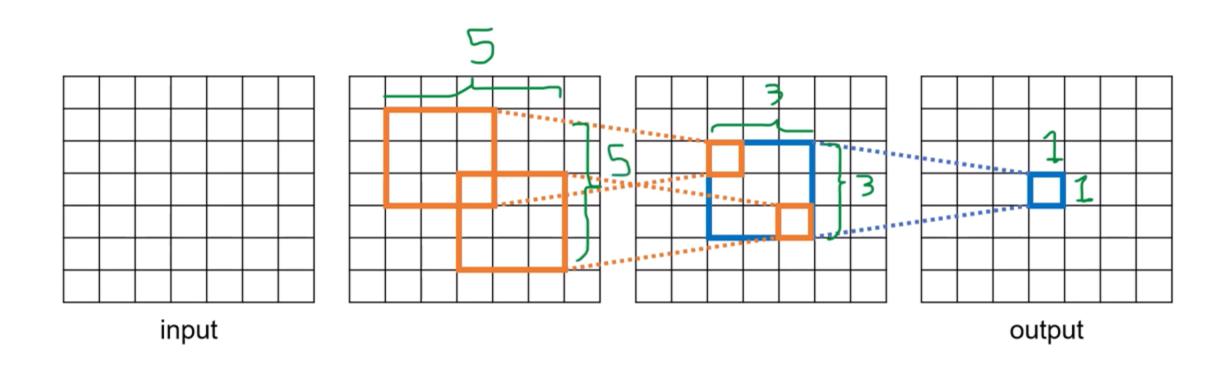


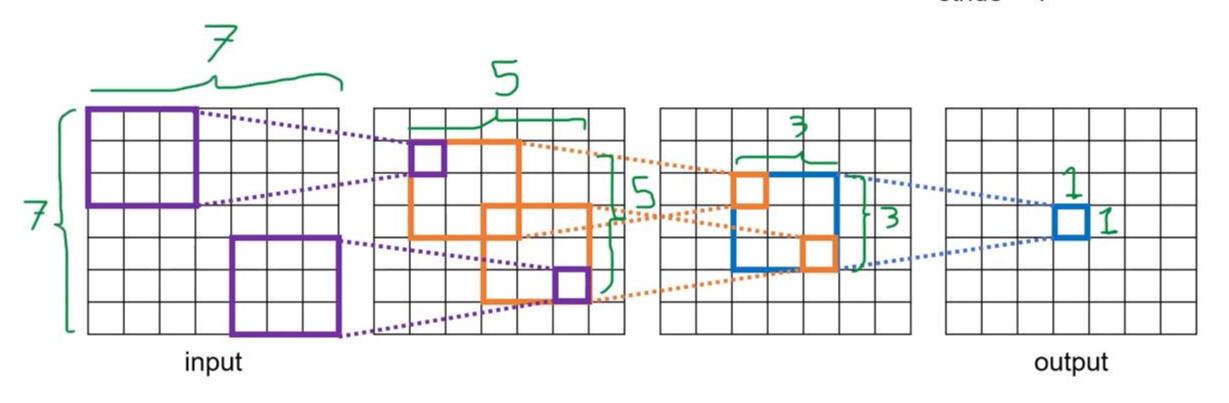
input

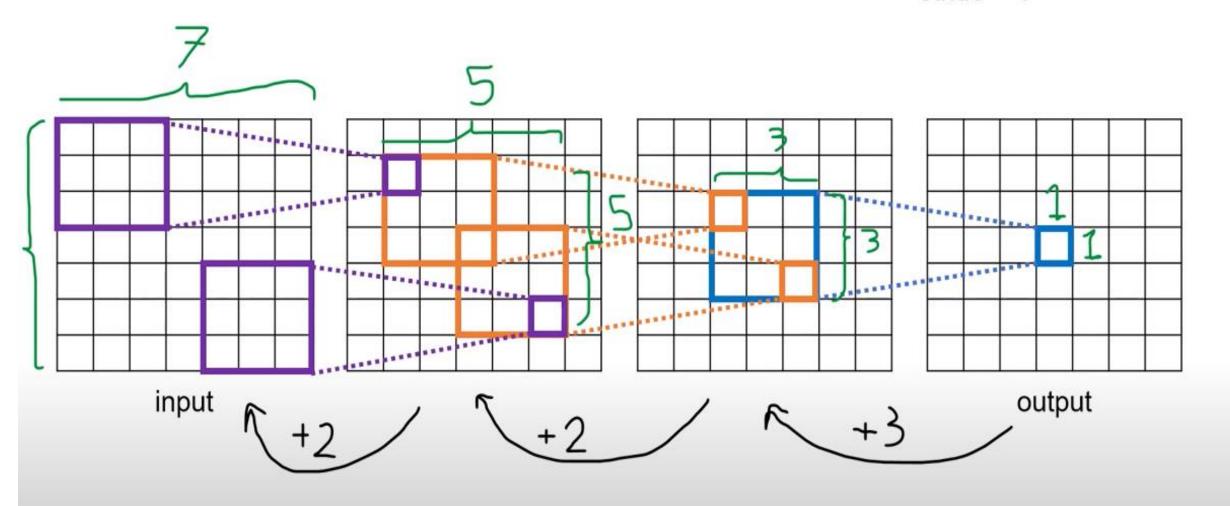
output



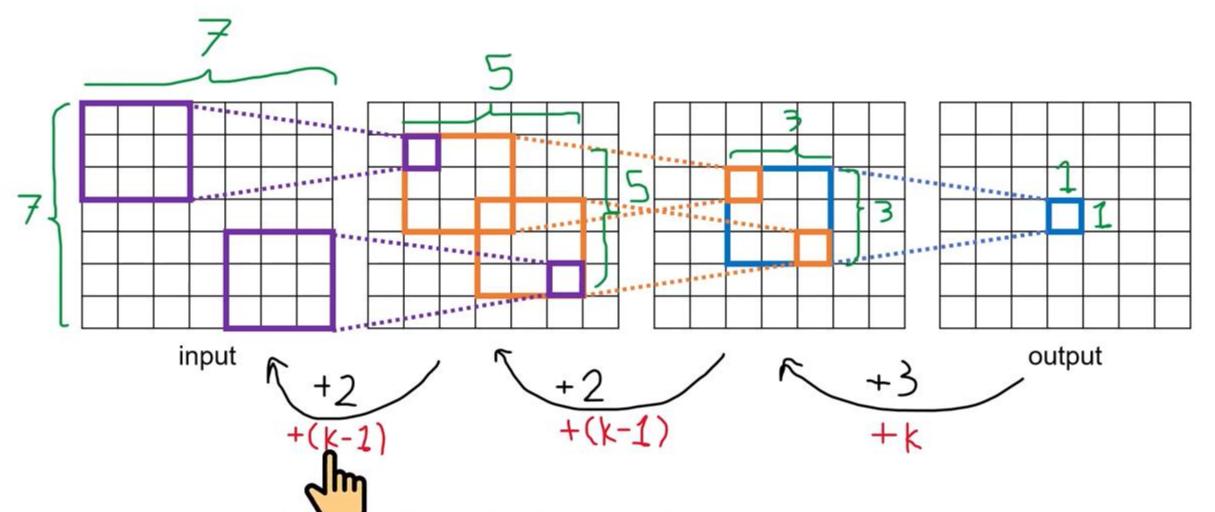






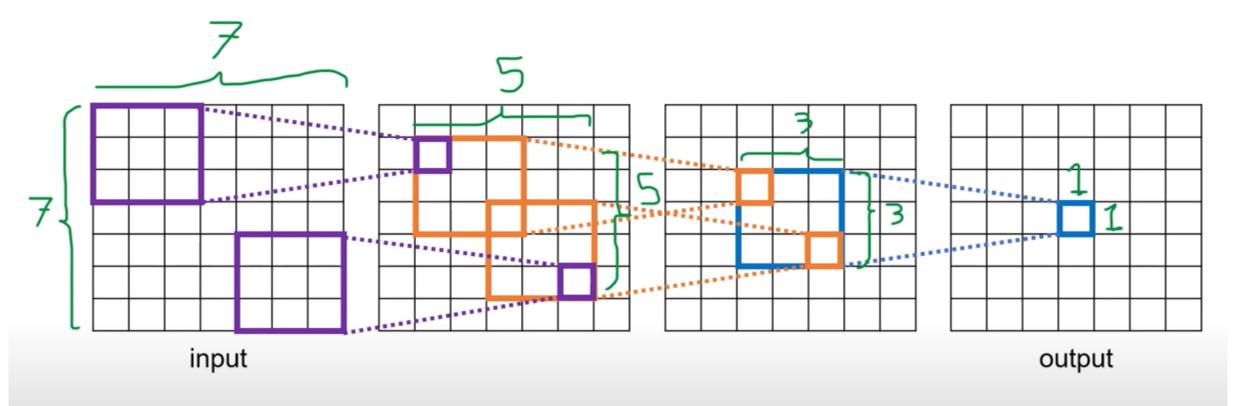


We only use 3x3 conv layers with padding = 1, stride = 1



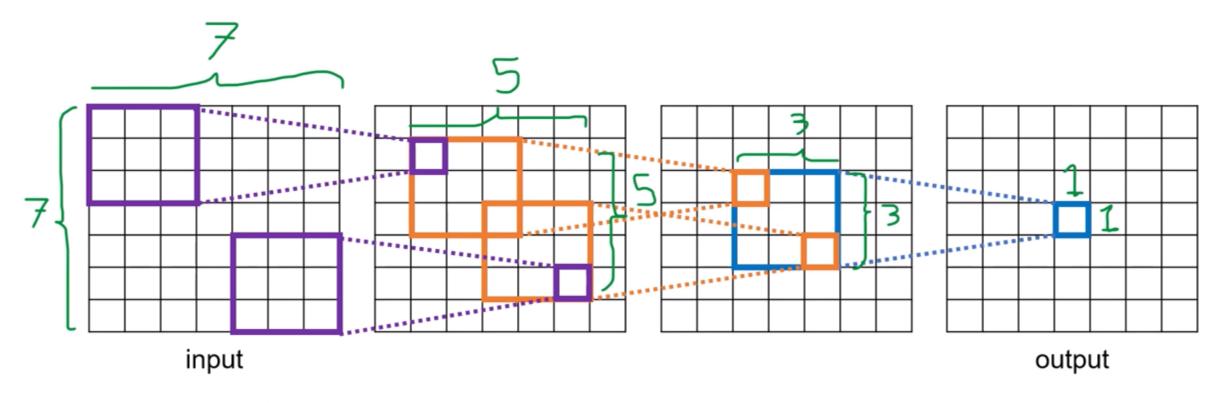
Assume having L layers, Receptive field size: k + (L-1)(k-1)

We only use 3x3 conv layers with padding = 1, stride = 1



Assume having L layers, Receptive field size: 1 + L(k-1)

We only use 3x3 conv layers with padding = 1, stride = 1



If input size is 224x224 And kernel size is 3:

$$1 + L(3 - 1) = 224 \rightarrow L \cong 112$$

Assume having L layers, Receptive field size: 1 + L(k-1)

(max-pooling)

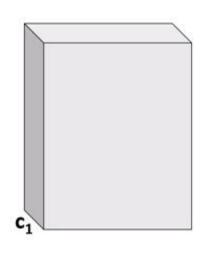
•				
	99	81	51	54
	93	25	73	29
	33	8	66	68
	39	68	59	31

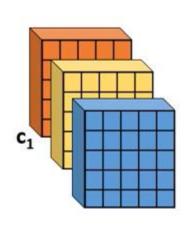
7	
99	73
68	68

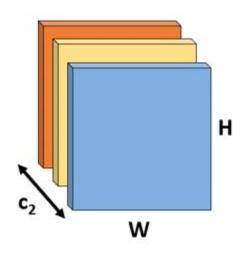


Why 3x3 conv layers!?

Why 3x3 conv layers!?

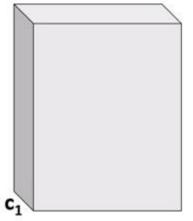


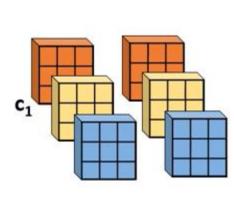


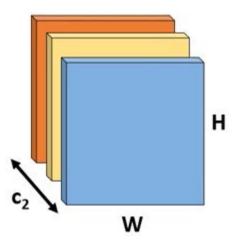


Parameters: 25C₁C₂

FLOPs: 25C₁C₂HW







Parameters: 18C₁C₂

FLOPs: 18C₁C₂HW