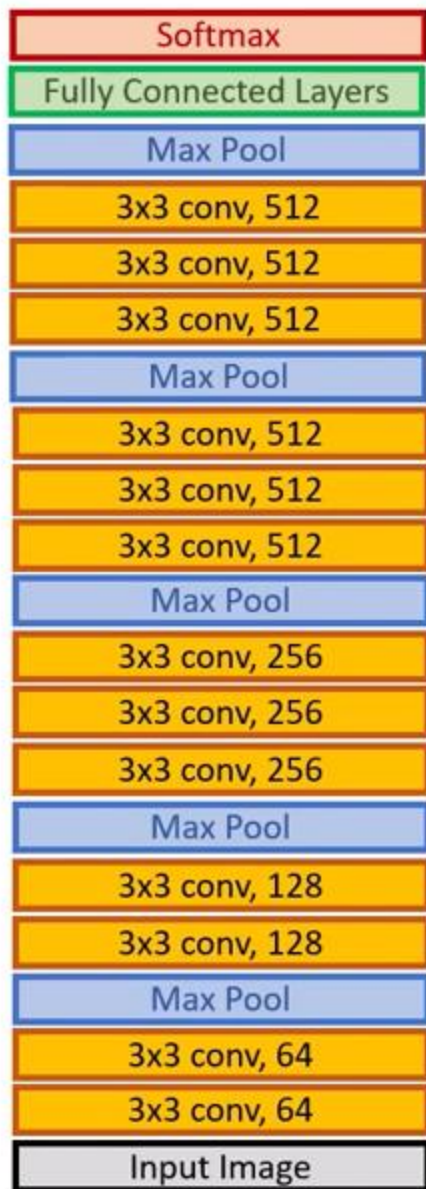
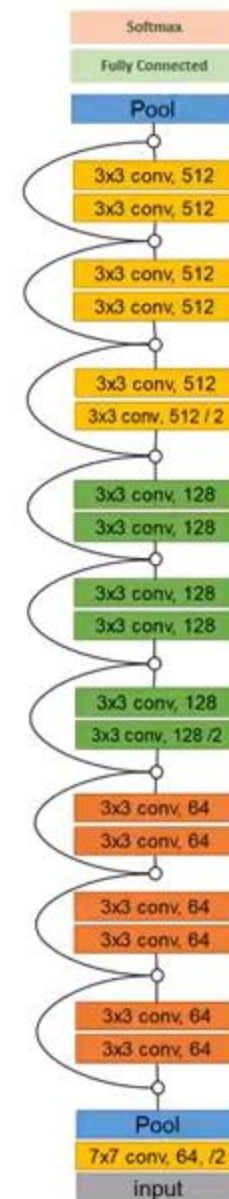


**Why 3x3 conv layers!?**



**VGG**

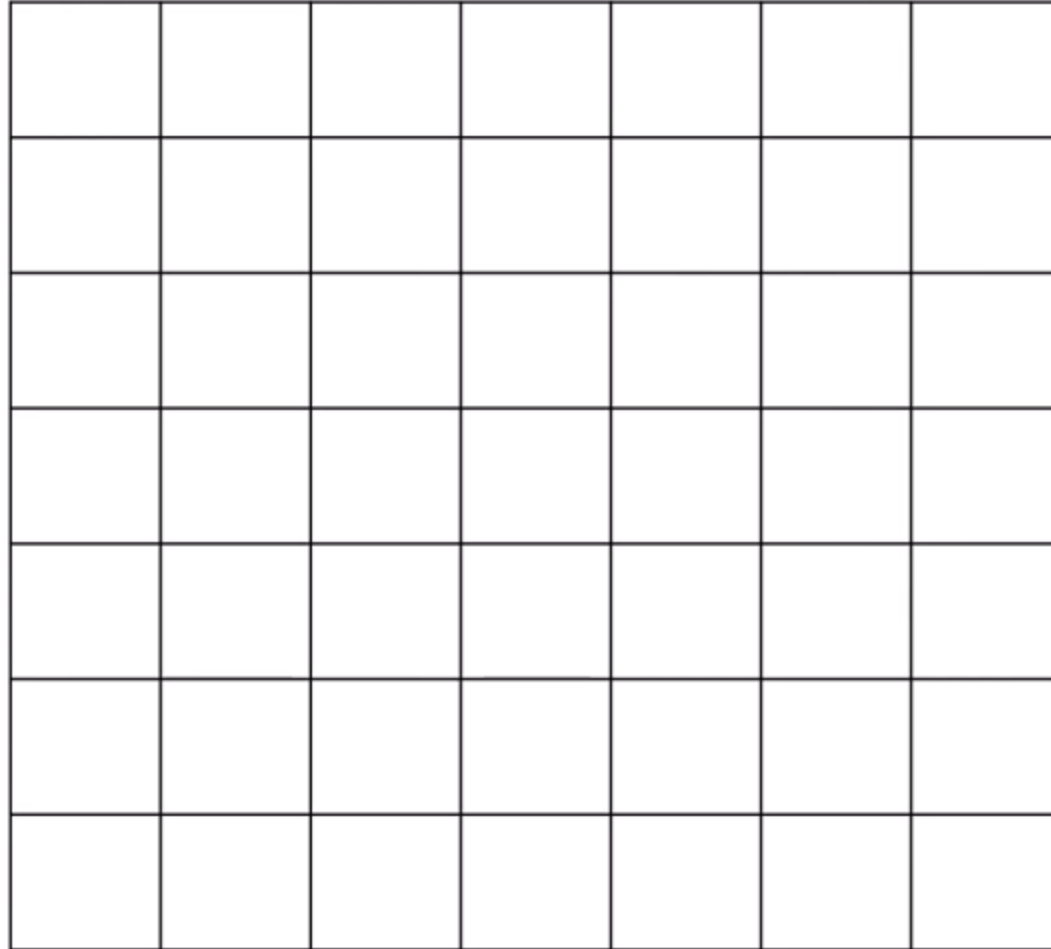


**ResNet**

Receptive **fields**

# Receptive fields

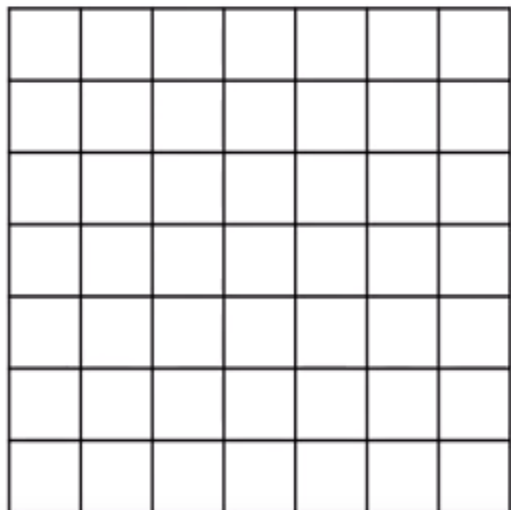
We only use 3x3 conv layers with padding = 1, stride = 1



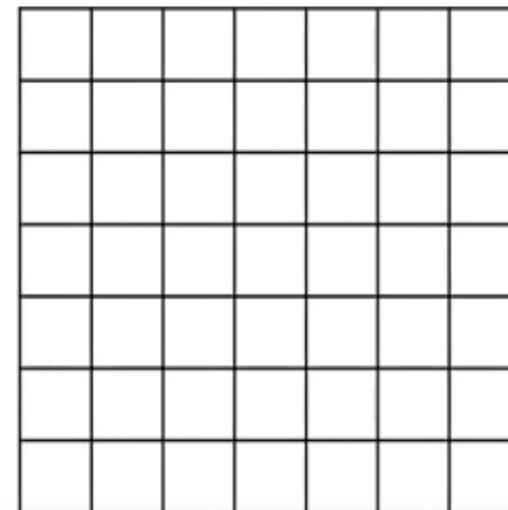
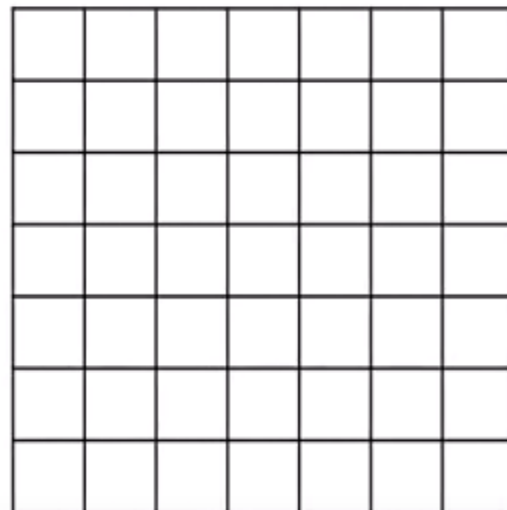
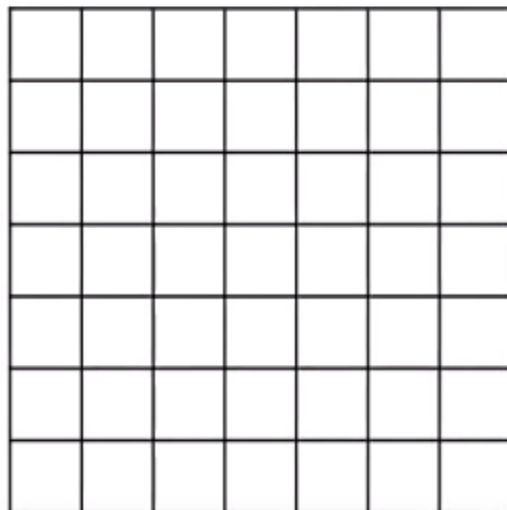
input

# Receptive fields

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



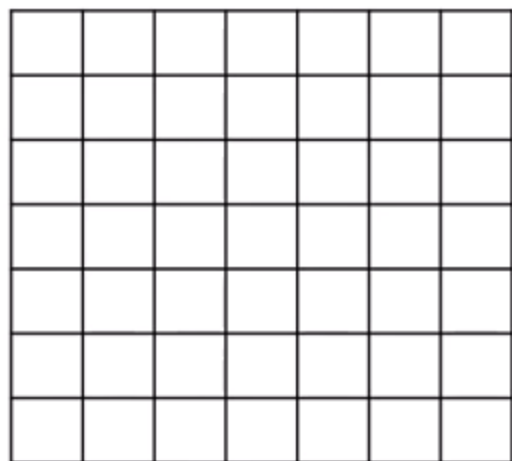
input



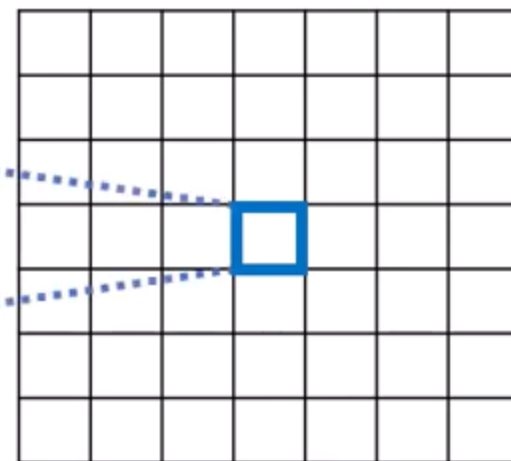
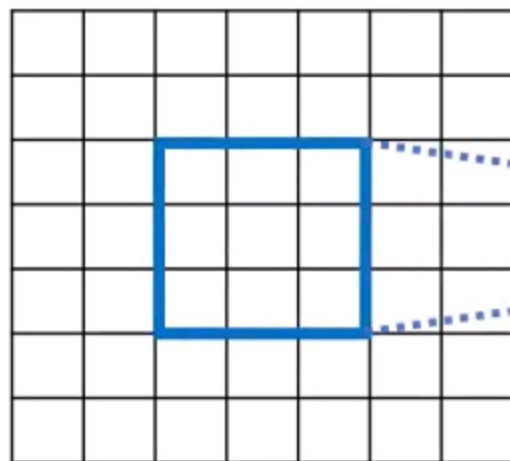
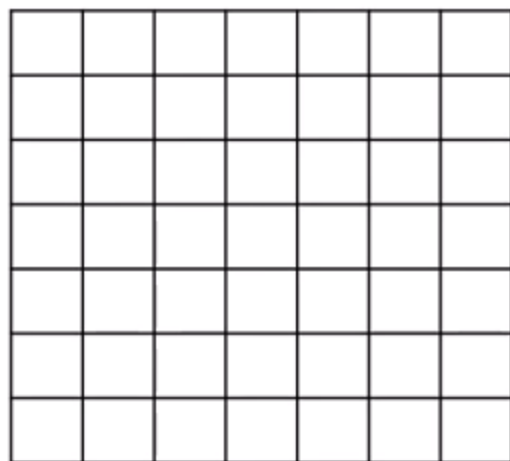
output

# Receptive fields

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



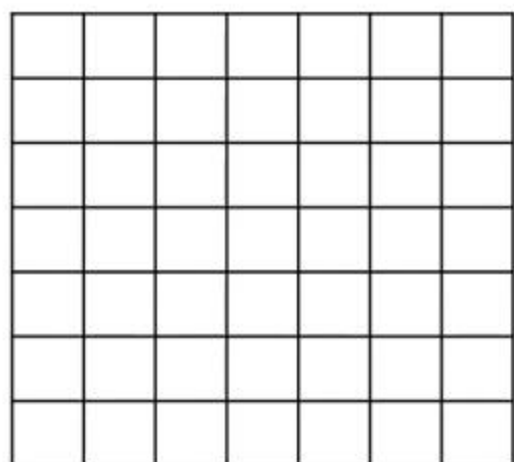
input



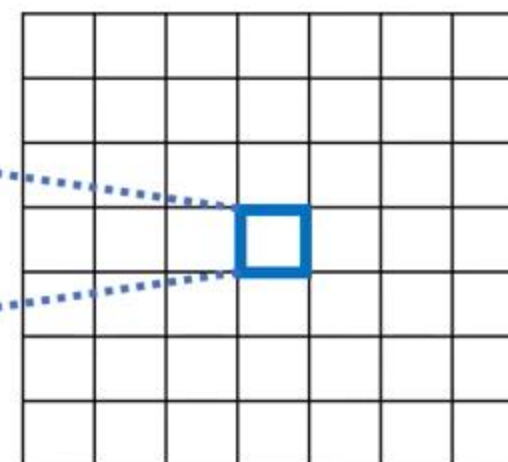
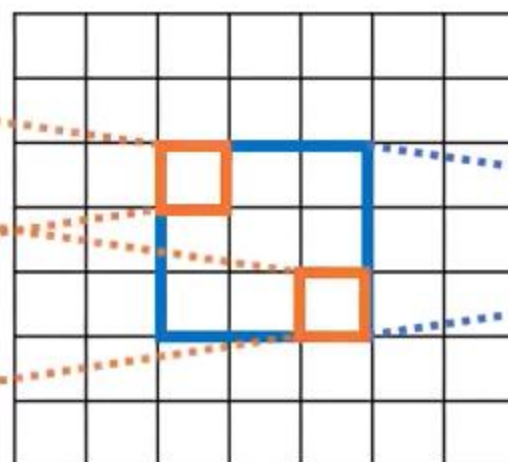
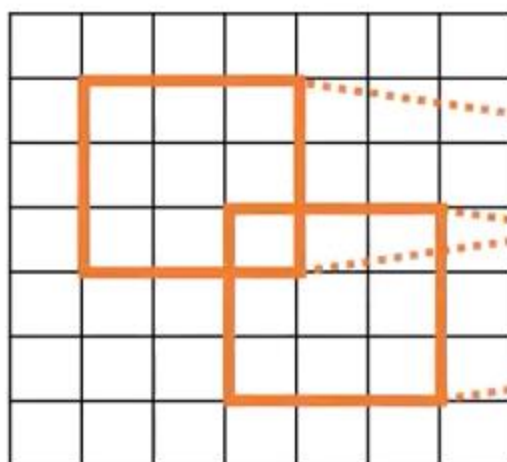
output

# Receptive fields

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



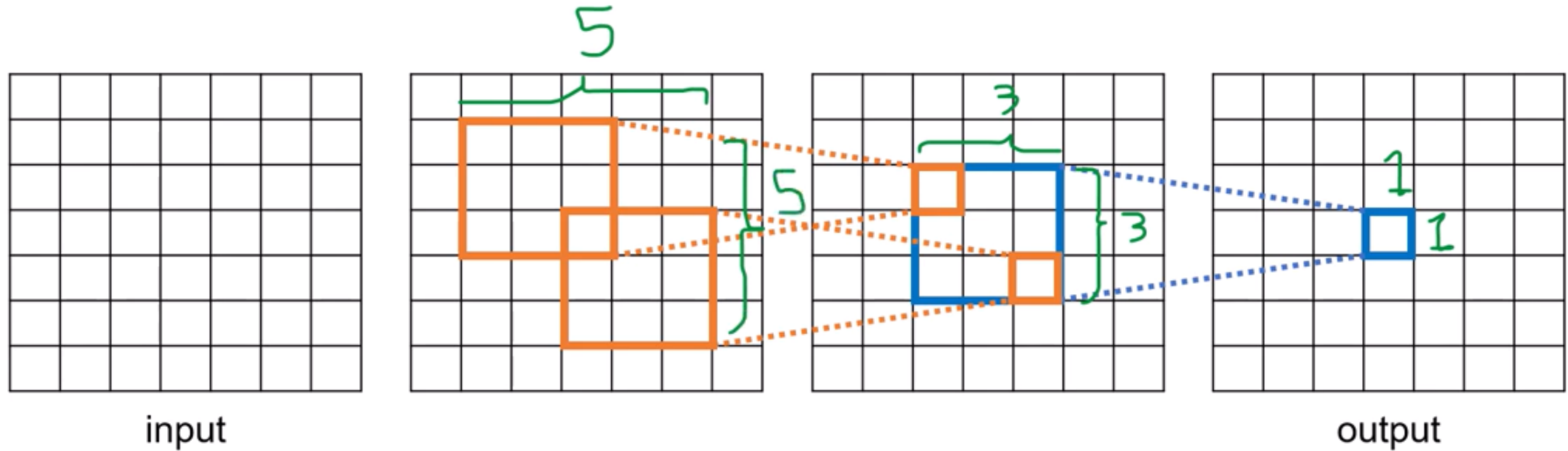
input



output

# Receptive fields

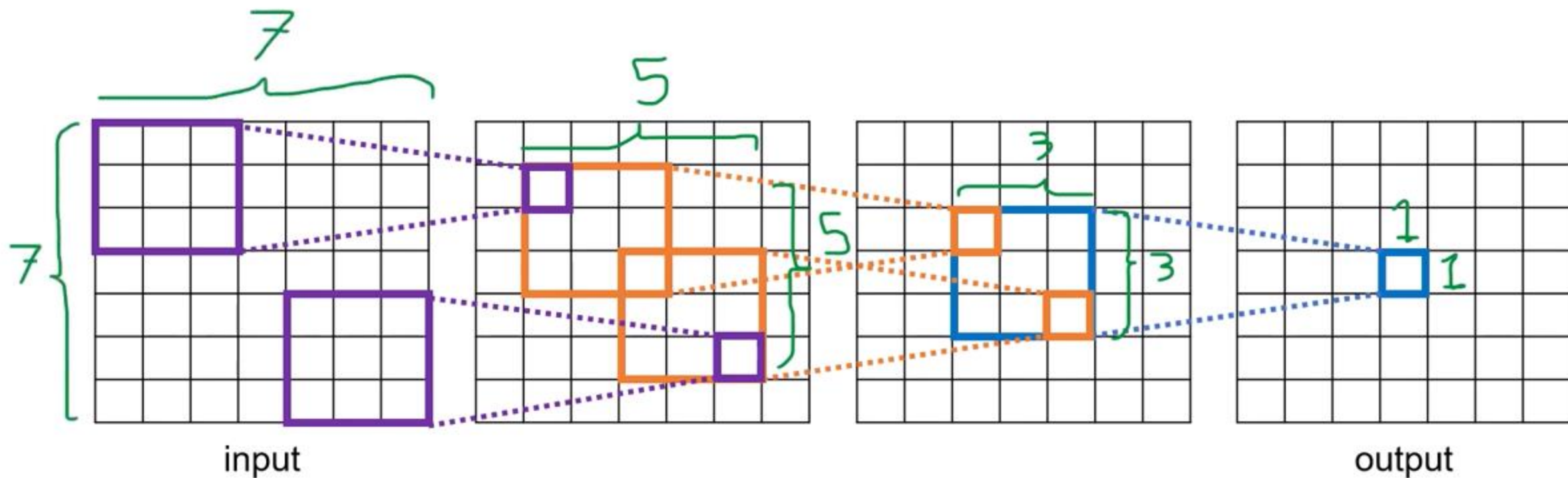
We only use 3x3 conv layers with padding = 1, stride = 1





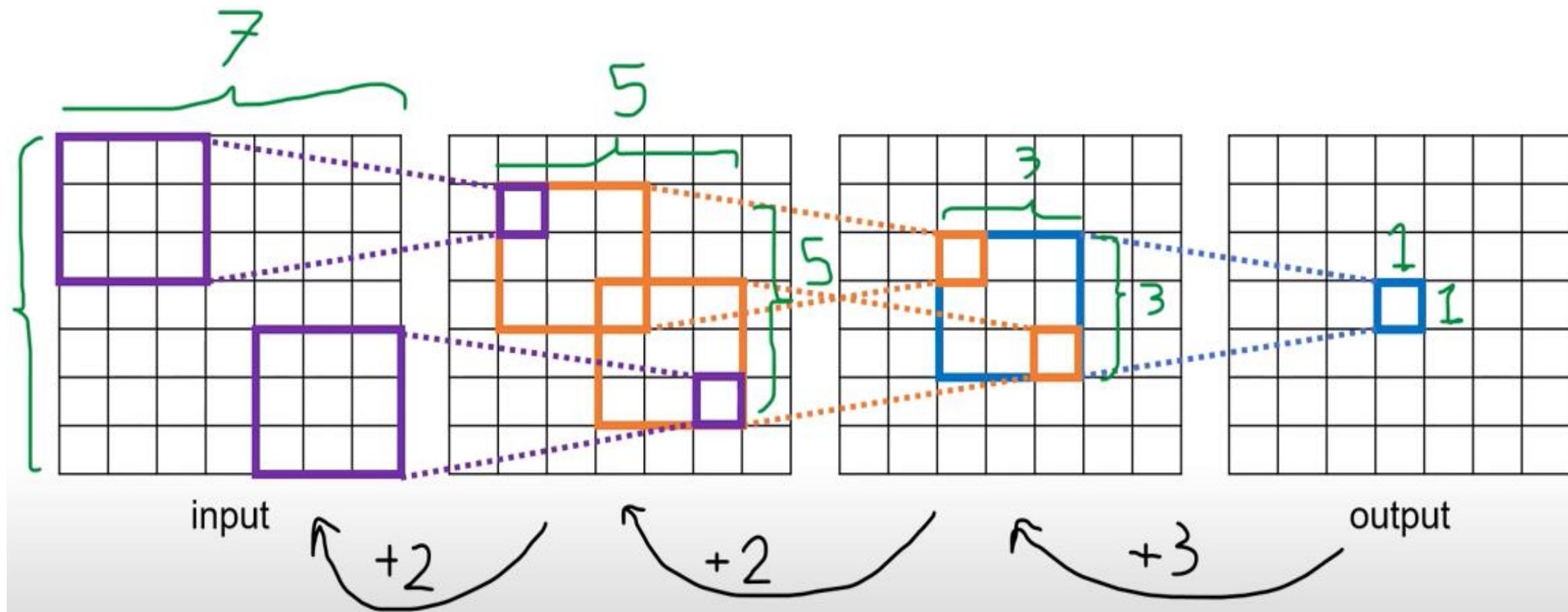
# Receptive fields

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



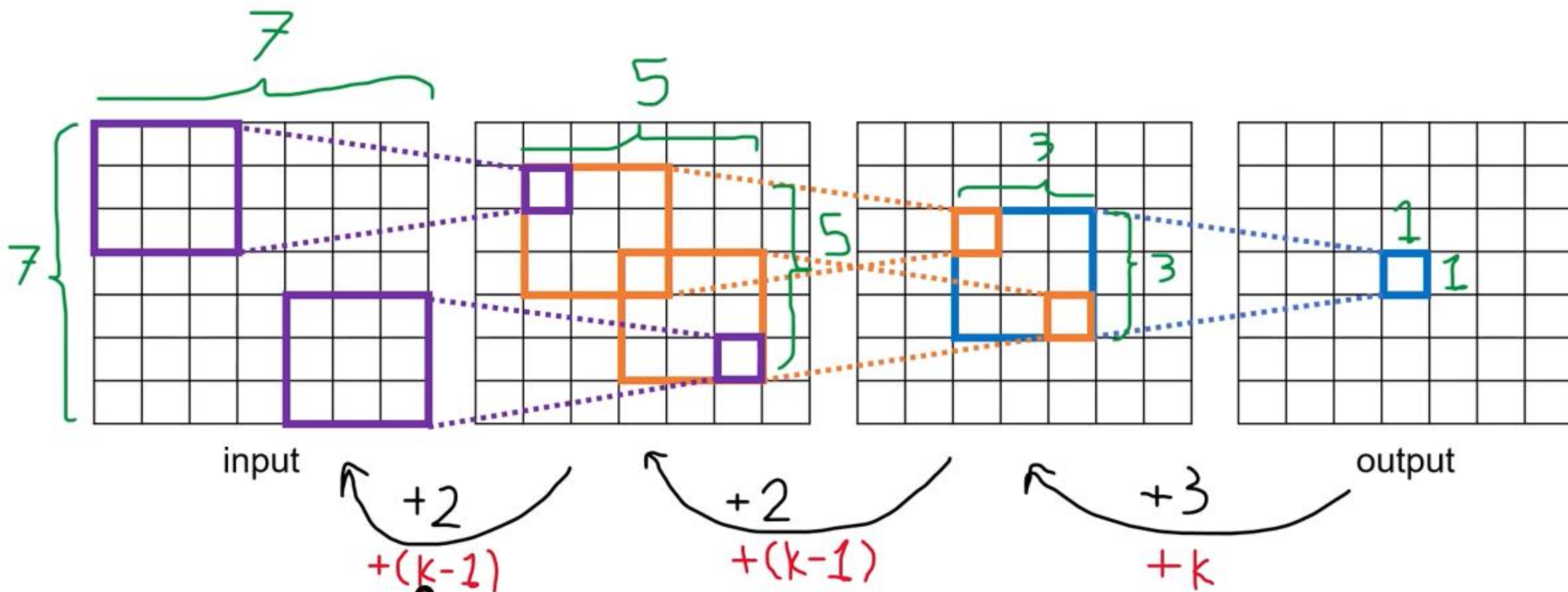
# Receptive fields

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



# Receptive fields

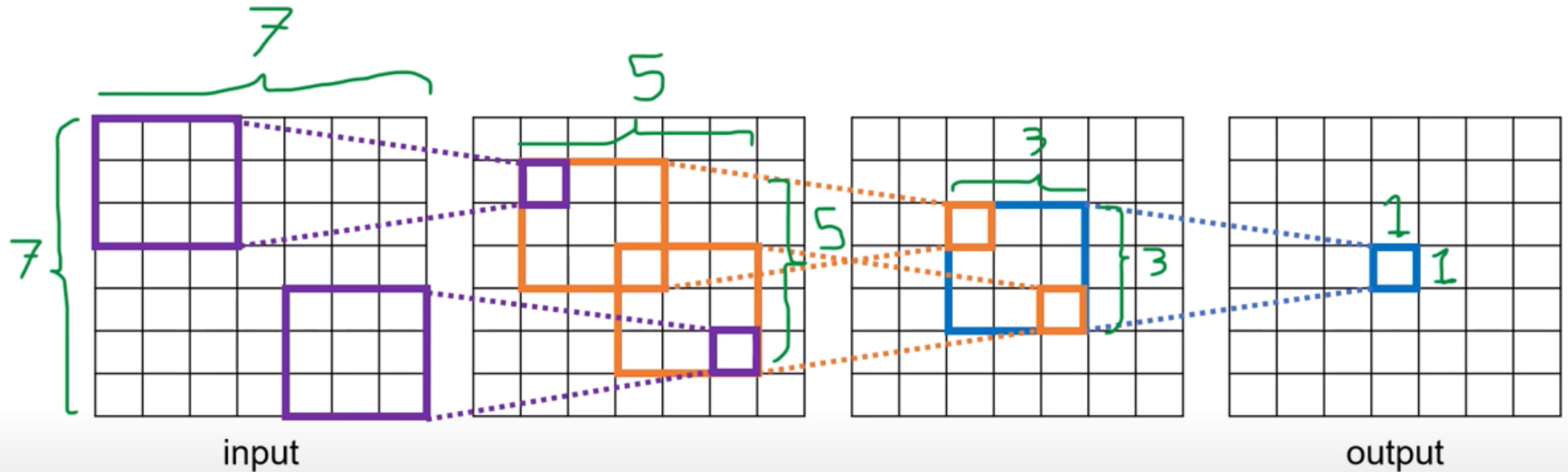
We only use 3x3 conv layers with padding = 1, stride = 1



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

# Receptive fields

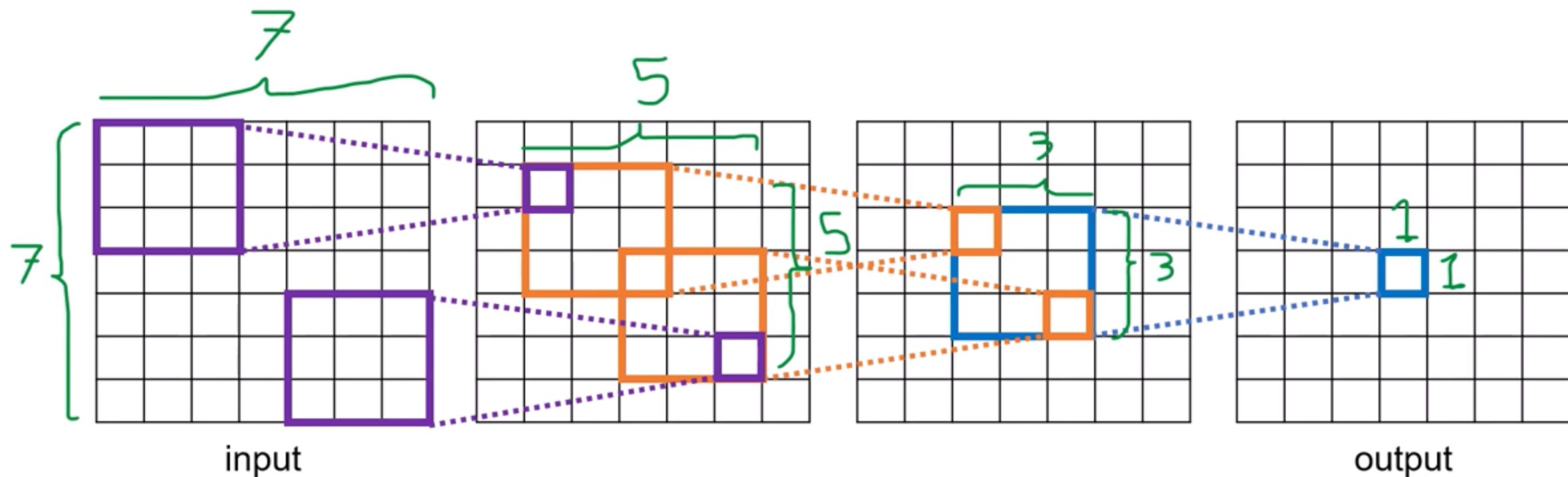
We only use 3x3 conv layers with padding = 1, stride = 1



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

# Receptive fields

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



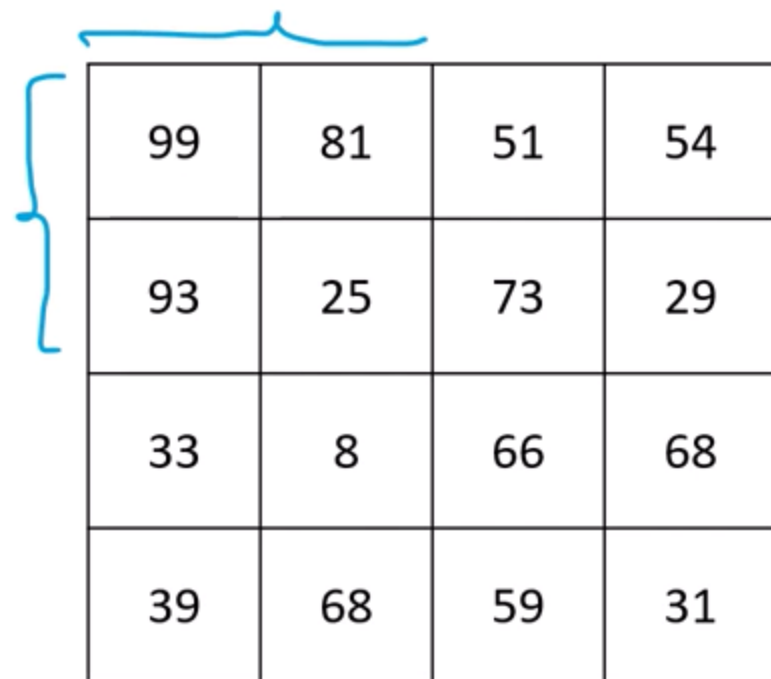
If input size is 224x224 And kernel size is 3:

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

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

# Receptive fields

(max-pooling)



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



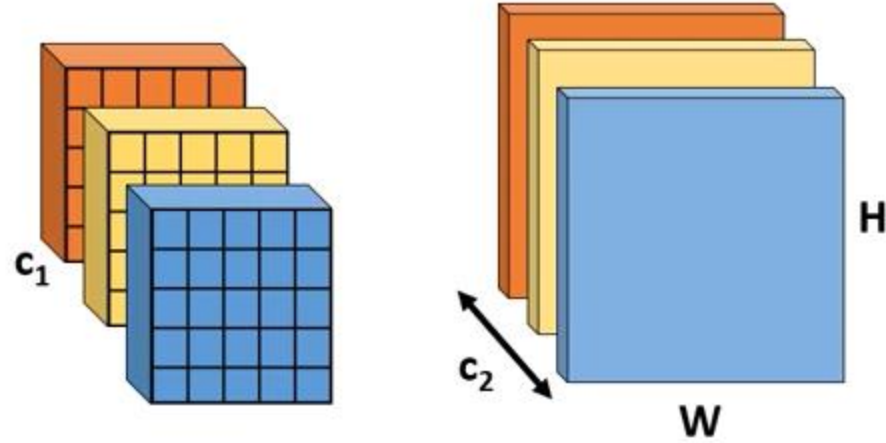
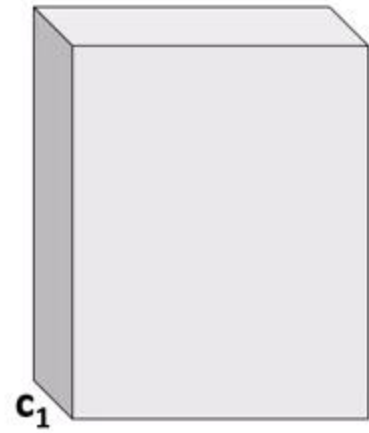
99	73
68	68

$\times 2$



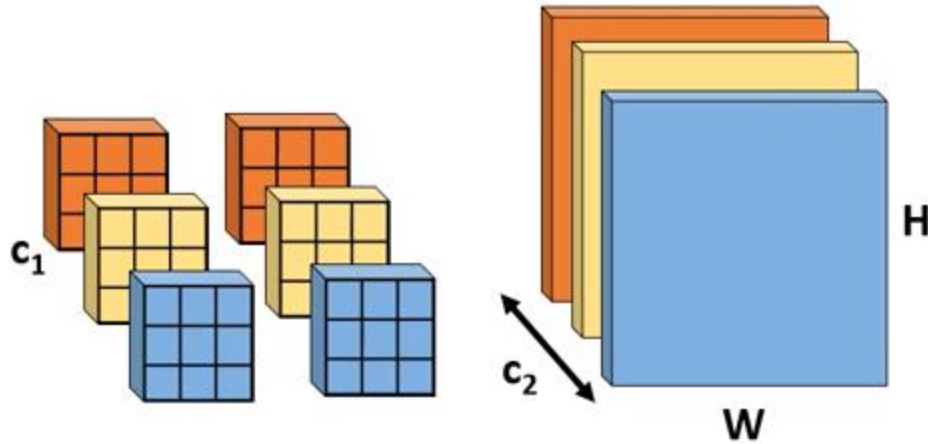
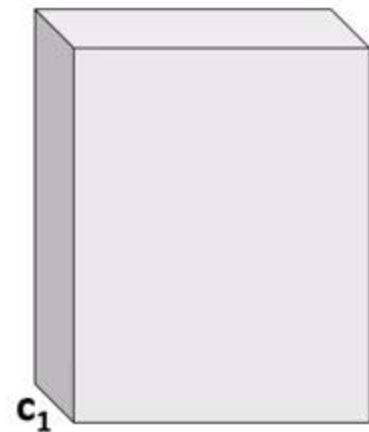
# **Why 3x3 conv layers!?**

# Why 3x3 conv layers!?



**Parameters:**  $25C_1C_2$

**FLOPs:**  $25C_1C_2HW$



**Parameters:**  $18C_1C_2$

**FLOPs:**  $18C_1C_2HW$