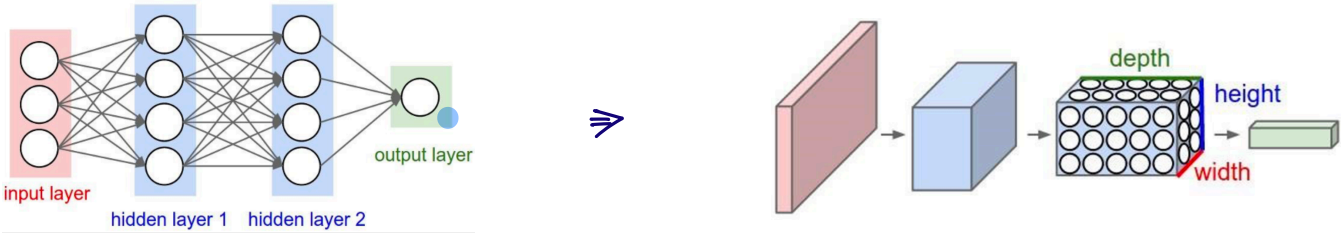
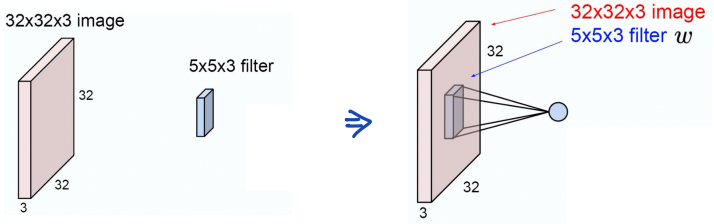
Ch 8. Convolutional Neural Networks

A1. Convolutional Layers

~ arrange its neurons in 3 dimensions (3D output tensor)



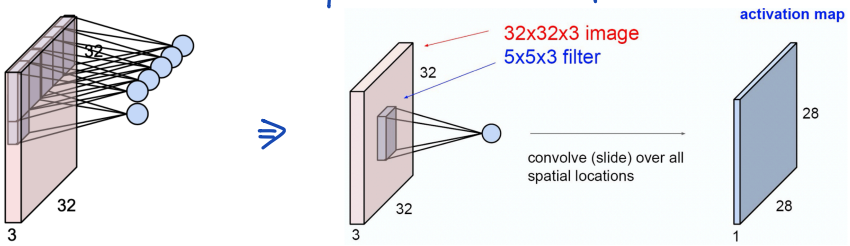
Given a 32x32x5 image and a 5x5x3 filter ⇒ Convole the filter with the image



1 number : $w^T x + b$

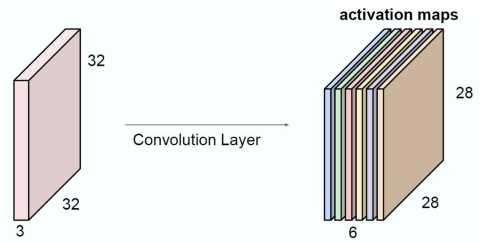
(the result of taking one dot product)

Construct the activation map based on one filter

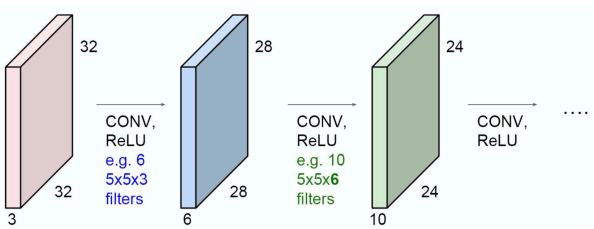


If we have more filters, we can get more seperate activation maps

And we can stack them up and get a new image

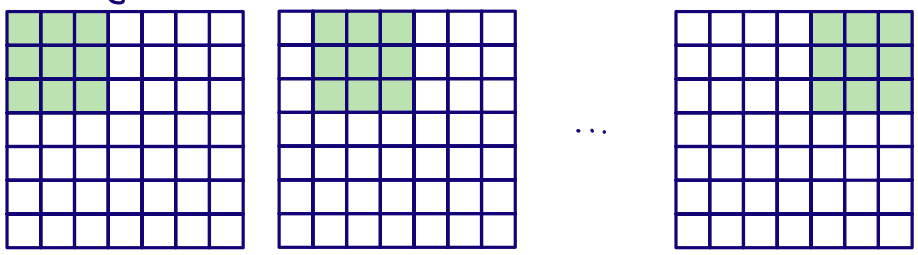


ConvNet: a sequence of Convolutional Layers, interspersed with activation function



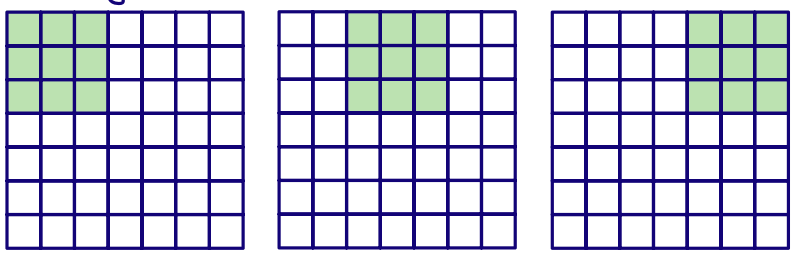
- Spatial Dimension Analysis

Given 7x7 image and 3x3 filter



⇒ 5x5 outputs

Given 7x7 image and 3x3 filter with stride 2



⇒ 3x3 outputs

Given 7x7 image with 1 zero padding to the border

0	0	0	0	0	0	0	0	0
0								0
0								0
0								0
0								0
0								0
0								0
0								0
0	0	0	0	0	0	0	0	0

N : Input dimension F : Filter size P : Width of zero-pad

$$\text{Output Size} = (N + 2 \times P - F)/\text{stride} + 1$$

Given a 3D Volume with size $W_1 \times H_1 \times D_1$ and 4 hyper-parameters :

filters K, filter size F, convolution stride S, # zero padding P

⇒ Produce a volume with size $W_2 \times H_2 \times D_2$ where

$W_2 = (W_1 + 2P - F)/S + 1$, $H_2 = (H_1 + 2P - F)/S + 1$, $D_2 = K$

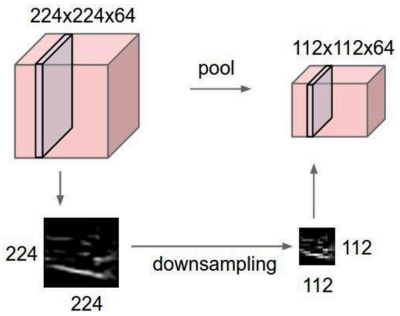
⇒ # parameters in the layer = $(F \times F \times D_1) \times K + \boxed{K}$ (bias)

☆ It has $(F \times F \times D_1) \times K$ weights and K bias

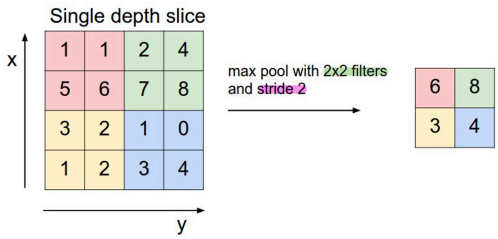
A2. Pooling Layers

~ make the representation smaller and more manageable

↳ operate over each activation map independently



☆ Max. pooling layer



EG. Input volume of size $224 \times 224 \times 64$ is pooled with filter size 2, stride 2 into output volume of size $112 \times 112 \times 64$

Given a 3D Volume with size $W_1 \times H_1 \times D$ and 2 hyper-parameters:

filter size F, convolution stride S

⇒ Produce a volume with size $W_2 \times H_2 \times D$ where

$W_2 = (W_1 - F)/S + 1$, $H_2 = (H_1 - F)/S + 1$

⇒ # parameters in the layer = 0