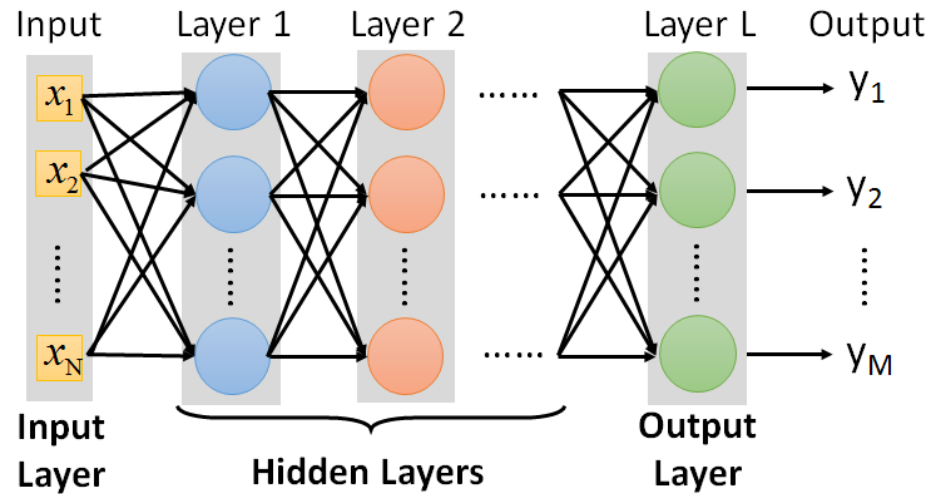


Machine Learning

Convolutional Neural Network

Can the network be simplified by considering the properties of images?

Recap



- Q: How many layers? How many neurons for each layer?

Trial and Error

+

Intuition

- Q: Can the structure be automatically determined?
 - E.g. Evolutionary Artificial Neural Networks
- Q: Can we design the network structure?

Convolutional Neural Network (CNN)

Why CNN for Image

- Some patterns are much smaller than the whole image

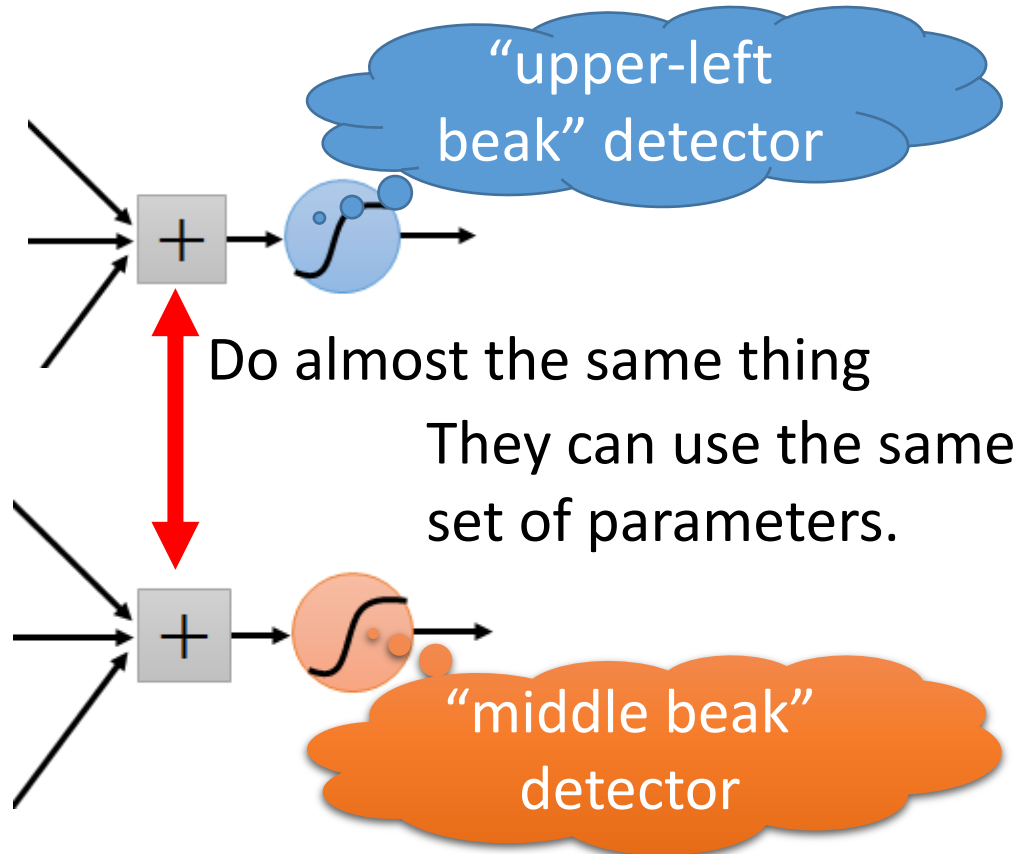
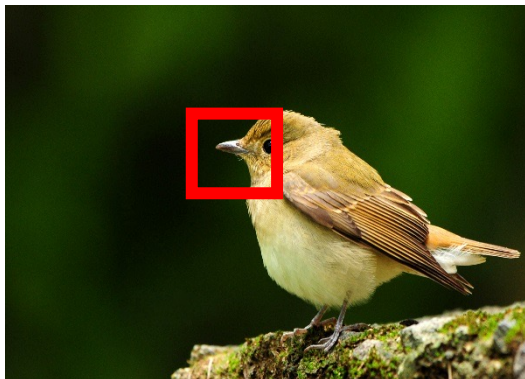
A neuron does not have to see the whole image to discover the pattern.

Connecting to small region with less parameters



Why CNN for Image

- The same patterns appear in different regions.



Why CNN for Image

- Subsampling the pixels will not change the object

bird



subsampling

bird

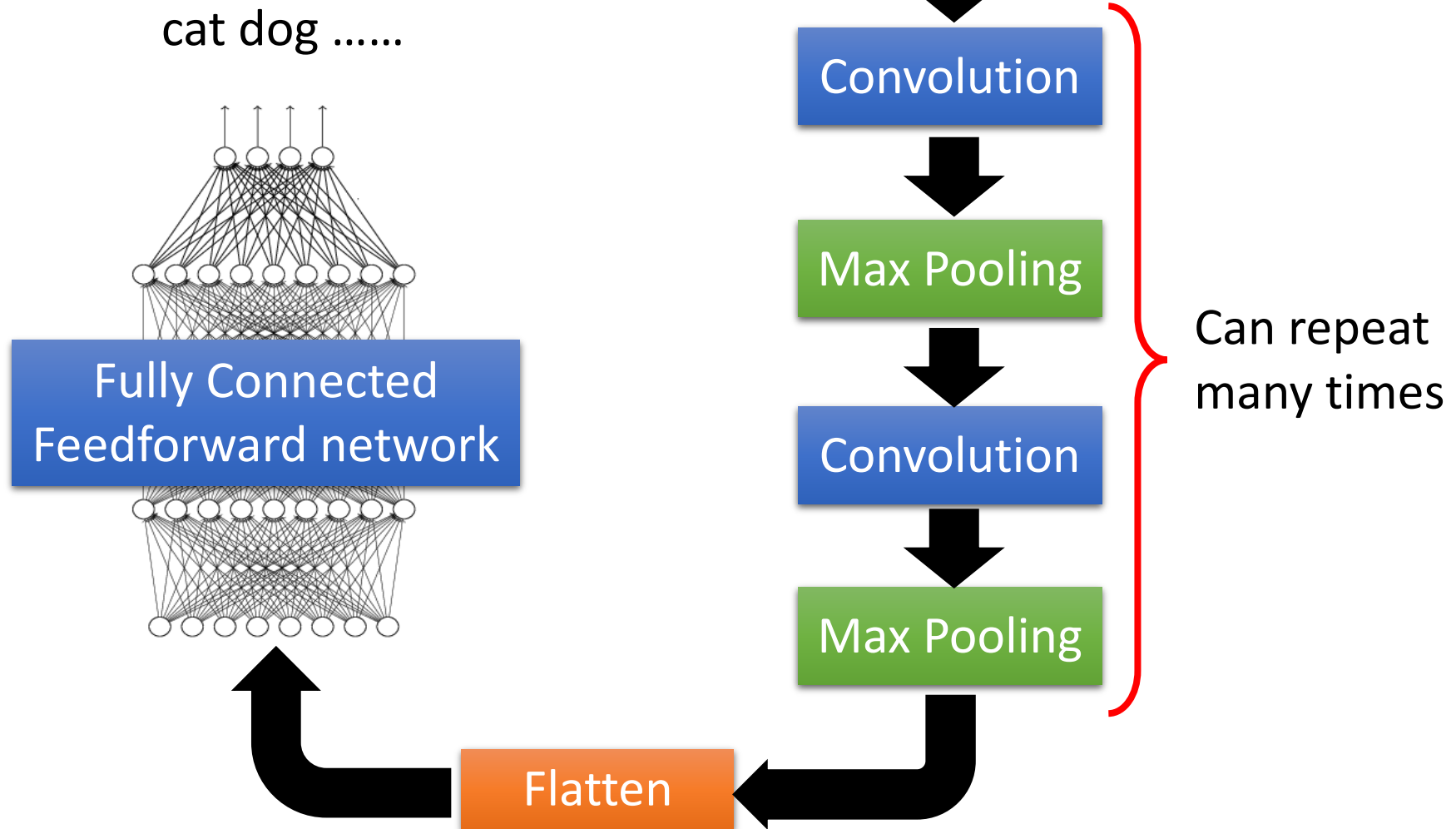


We can subsample the pixels to make image smaller



Less parameters for the network to process the image

The whole CNN



The whole CNN

Property 1

- Some patterns are much smaller than the whole image

Property 2

- The same patterns appear in different regions.

Property 3

- Subsampling the pixels will not change the object



Convolution

Max Pooling

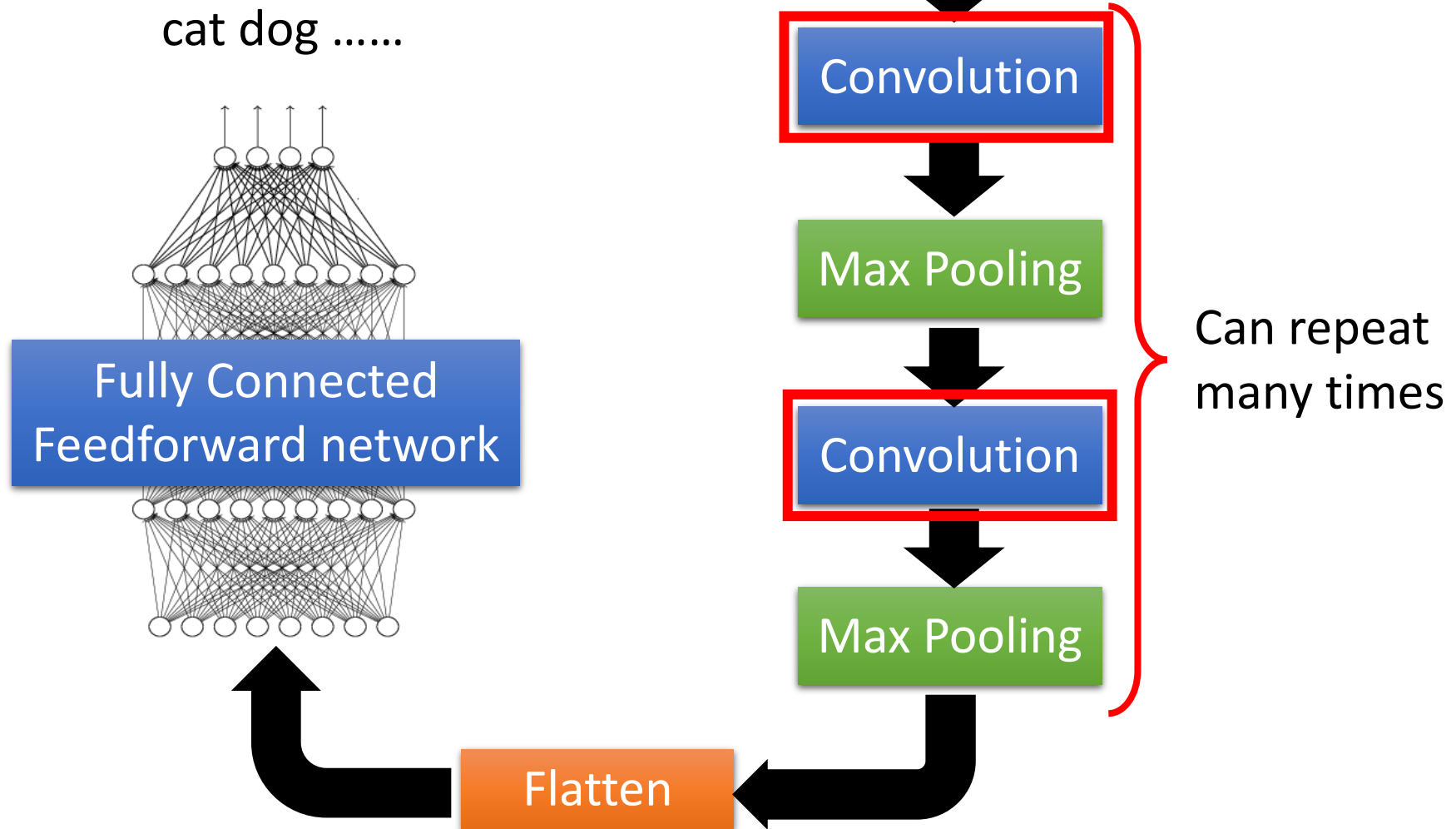
Convolution

Max Pooling

Flatten

Can repeat many times

The whole CNN



CNN – Convolution

Those are the network parameters to be learned.

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

Matrix

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

Matrix

⋮

Property 1

Each filter detects a small pattern (3 x 3).

CNN – Convolution

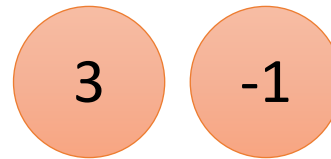
stride=1

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1



CNN – Convolution

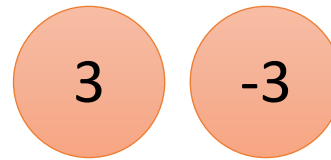
1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

If stride=2

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

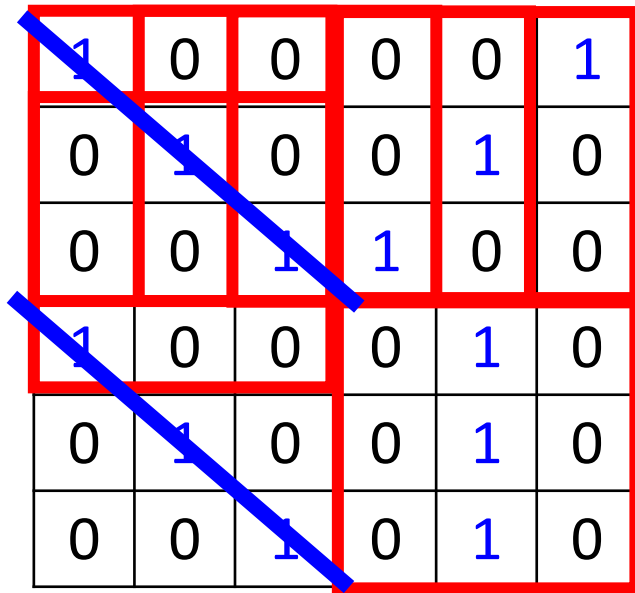
6 x 6 image



We set stride=1 below

CNN – Convolution

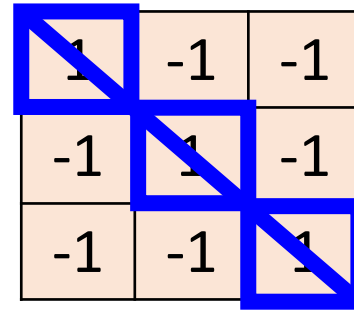
stride=1



A 6x6 grid of numbers representing an image. The values are: Row 1: 1, 0, 0, 0, 0, 1; Row 2: 0, 1, 0, 0, 1, 0; Row 3: 0, 0, 1, 1, 0, 0; Row 4: 1, 0, 0, 0, 1, 0; Row 5: 0, 1, 0, 0, 1, 0; Row 6: 0, 0, 1, 0, 1, 0. A 3x3 red bounding box highlights the top-left corner (rows 1-3, columns 1-3). A blue diagonal line runs from the top-left to the bottom-right of the entire grid.

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

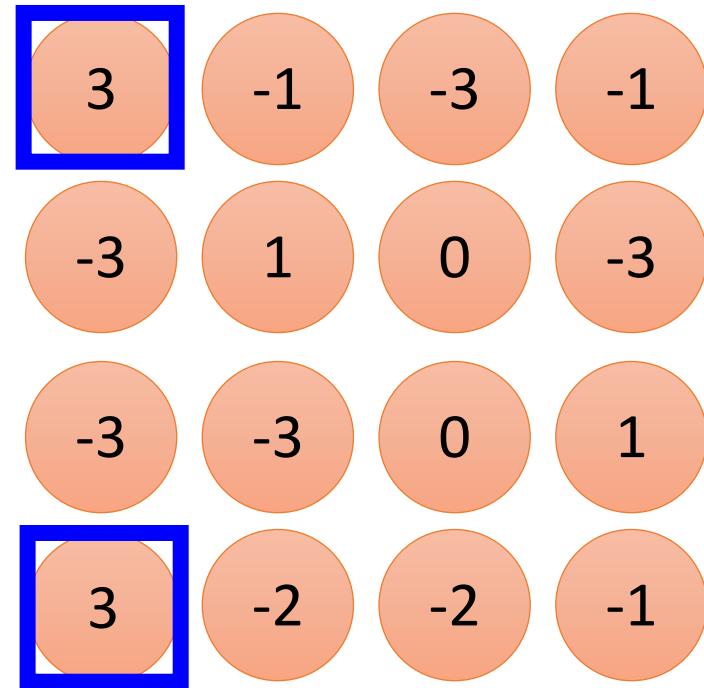
6 x 6 image



A 3x3 grid of numbers representing a filter. The values are: Row 1: 1, -1, -1; Row 2: -1, 1, -1; Row 3: -1, -1, 1. A blue diagonal line runs from the top-left to the bottom-right. A blue bounding box highlights the top-left 2x2 sub-region.

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1



A 4x4 grid of circles containing numbers. The values are: Row 1: 3, -1, -3, -1; Row 2: -3, 1, 0, -3; Row 3: -3, -3, 0, 1; Row 4: 3, -2, -2, -1. Two blue bounding boxes highlight the top-left and bottom-left circles, both containing the value 3.

3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

Property 2

CNN – Convolution

-1	1	-1
-1	1	-1
-1	1	-1

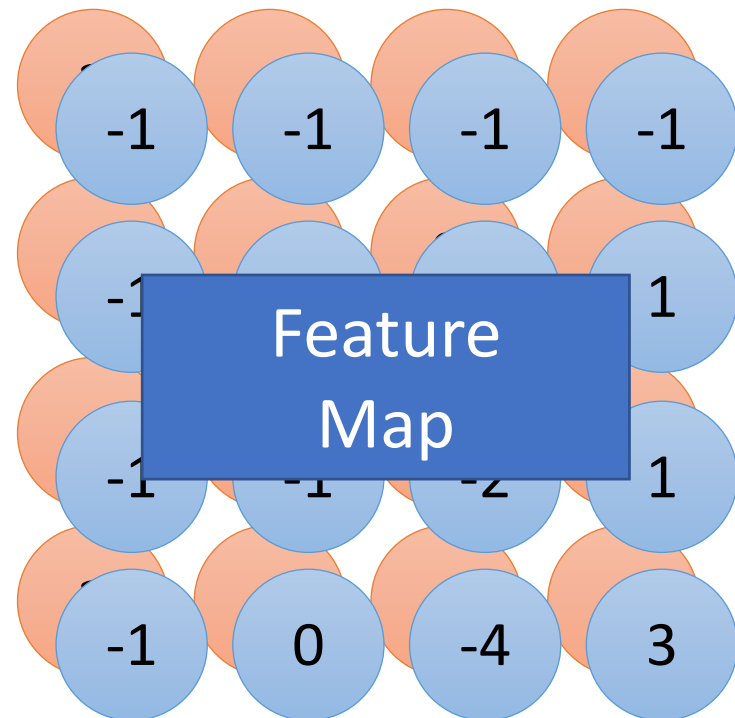
Filter 2

stride=1

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

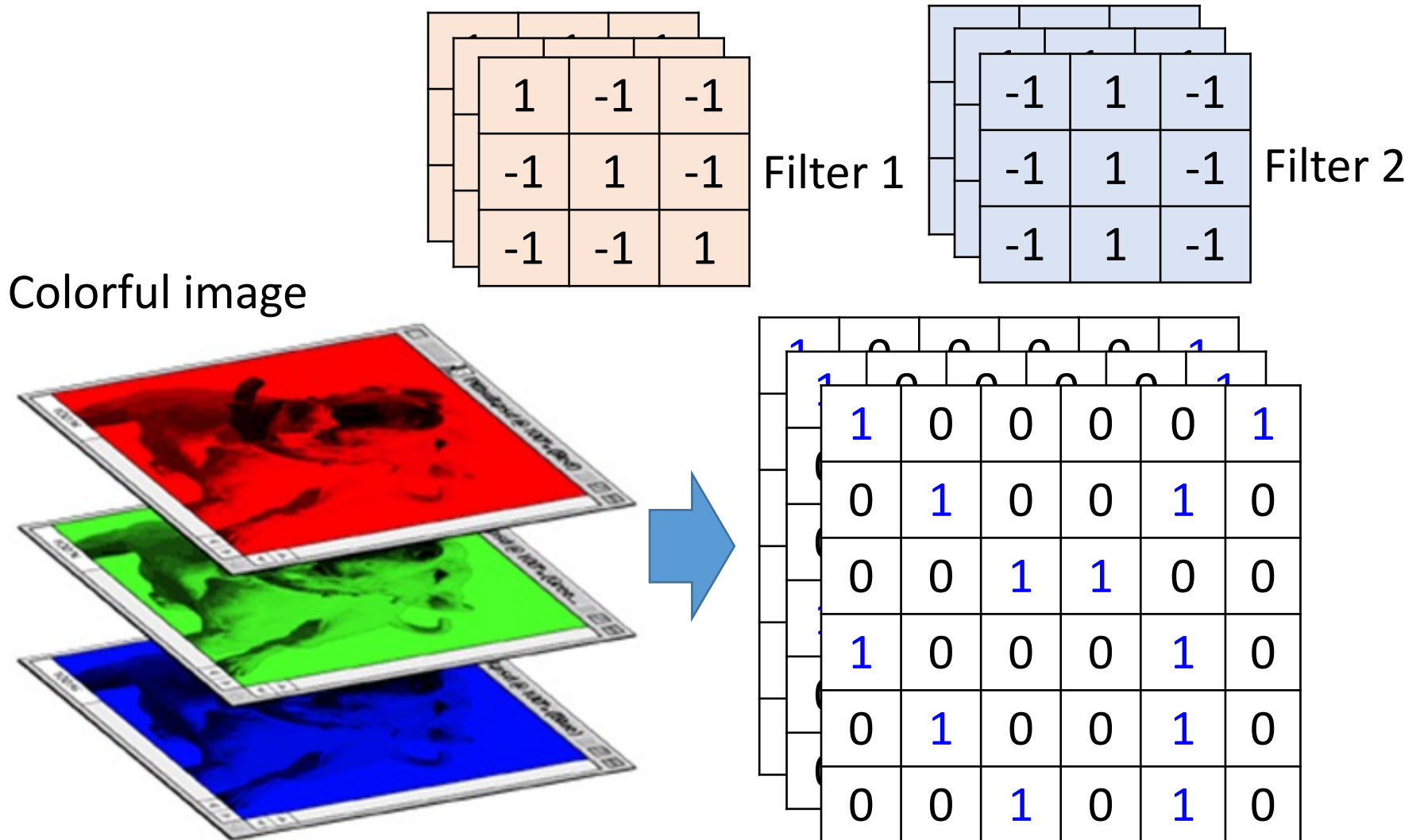
6 x 6 image

Do the same process for every filter

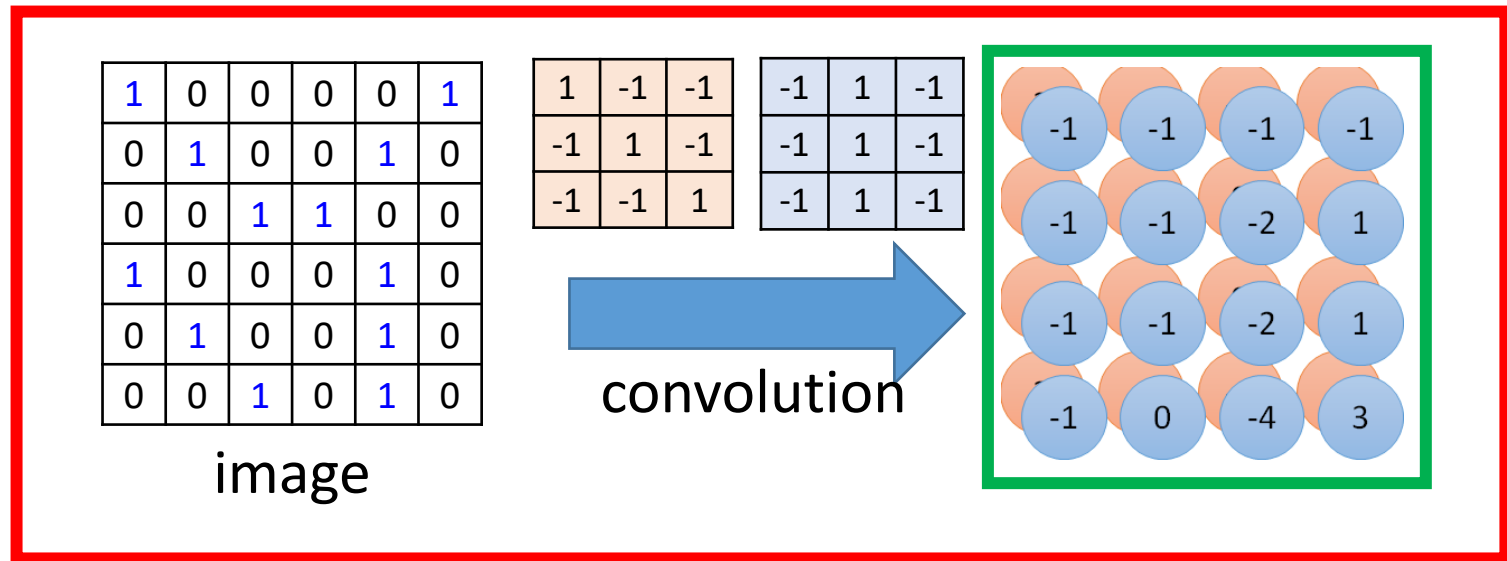


4 x 4 image

CNN – Colorful image

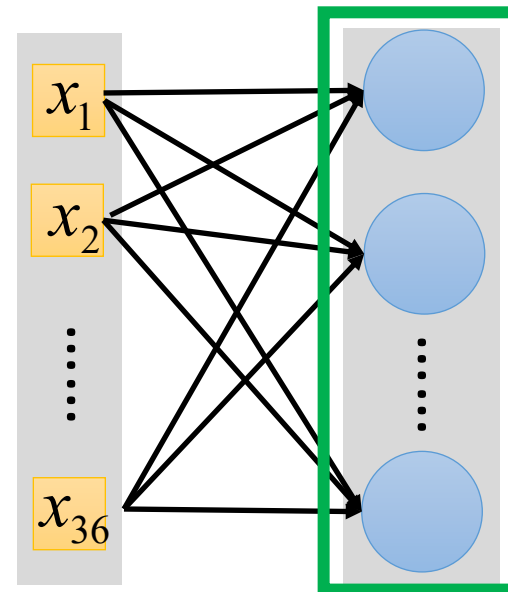


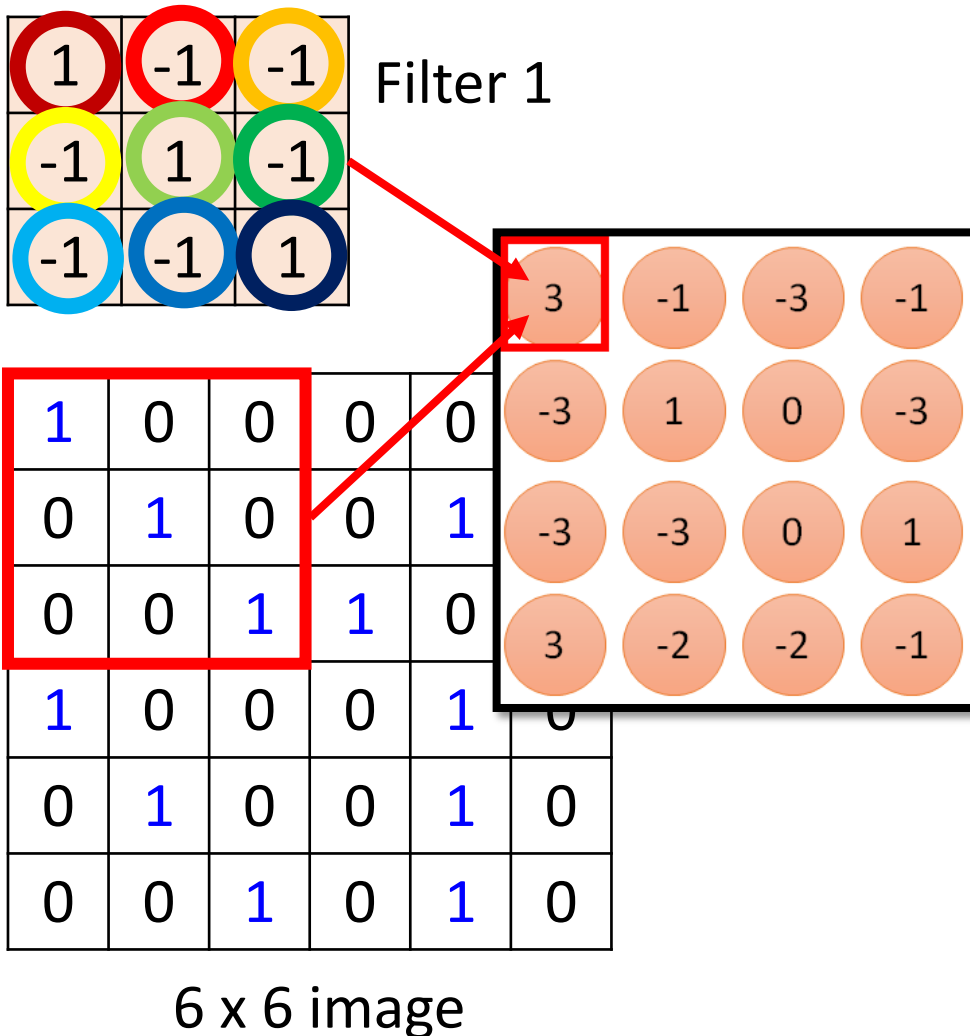
Convolution v.s. Fully Connected



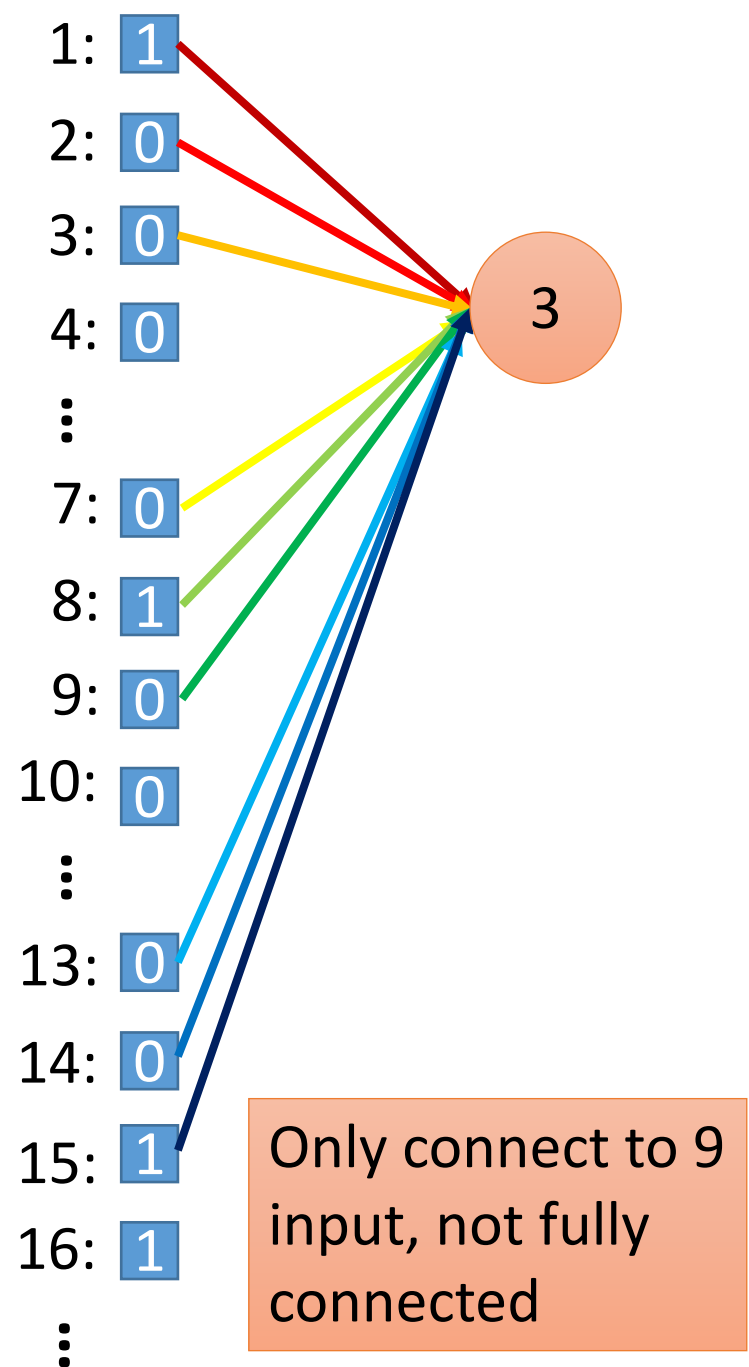
Fully-
connected

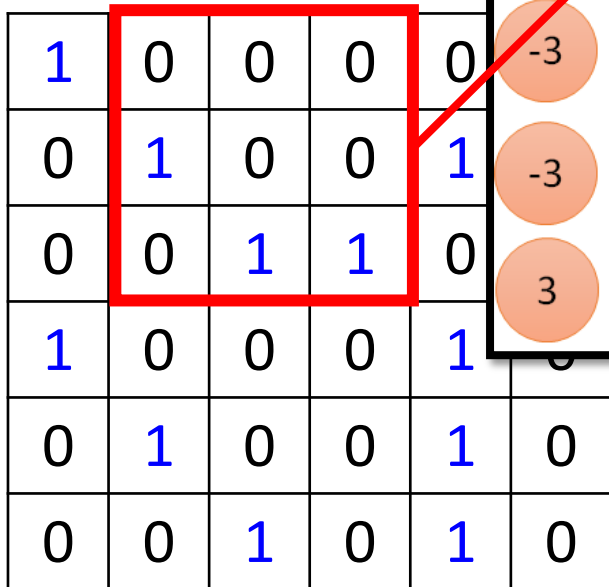
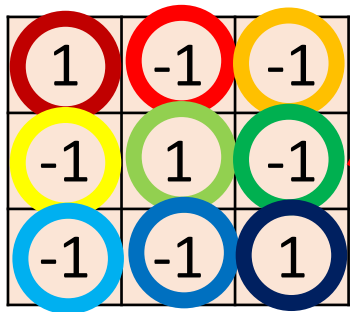
1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0





Less parameters!

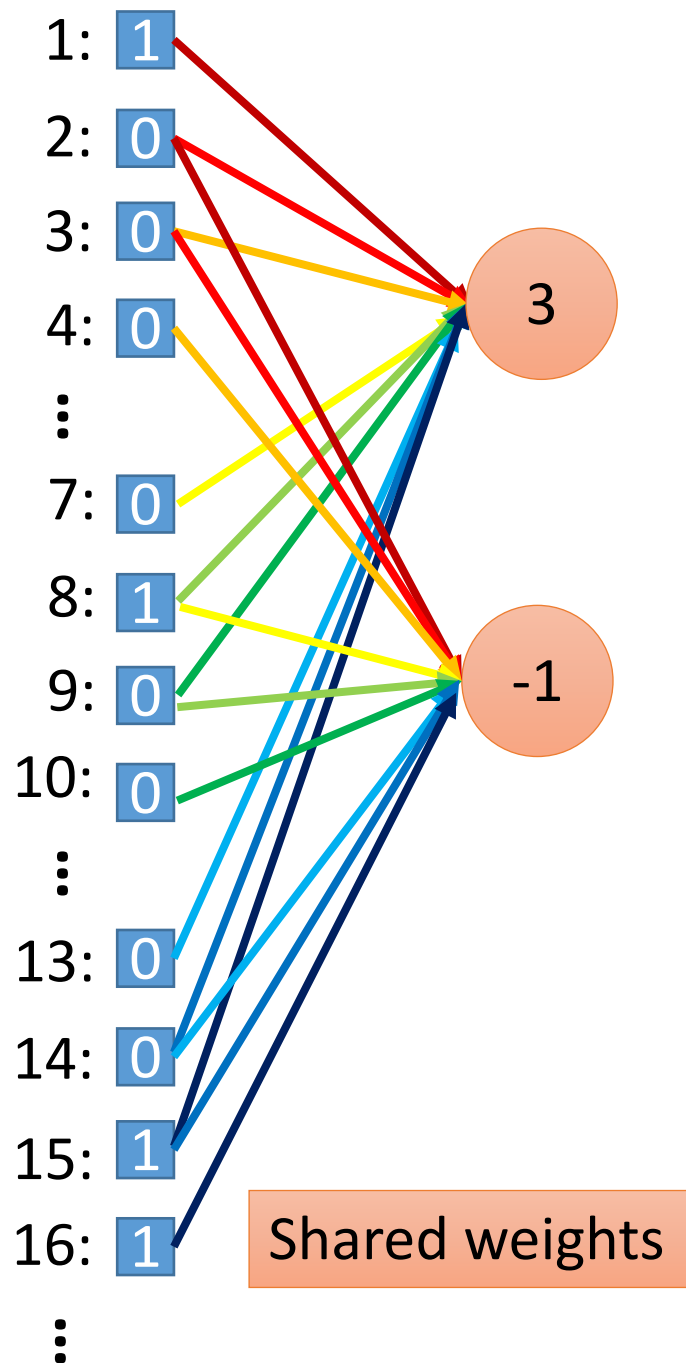
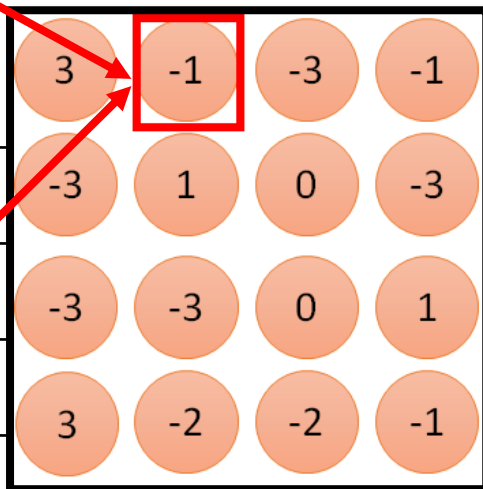




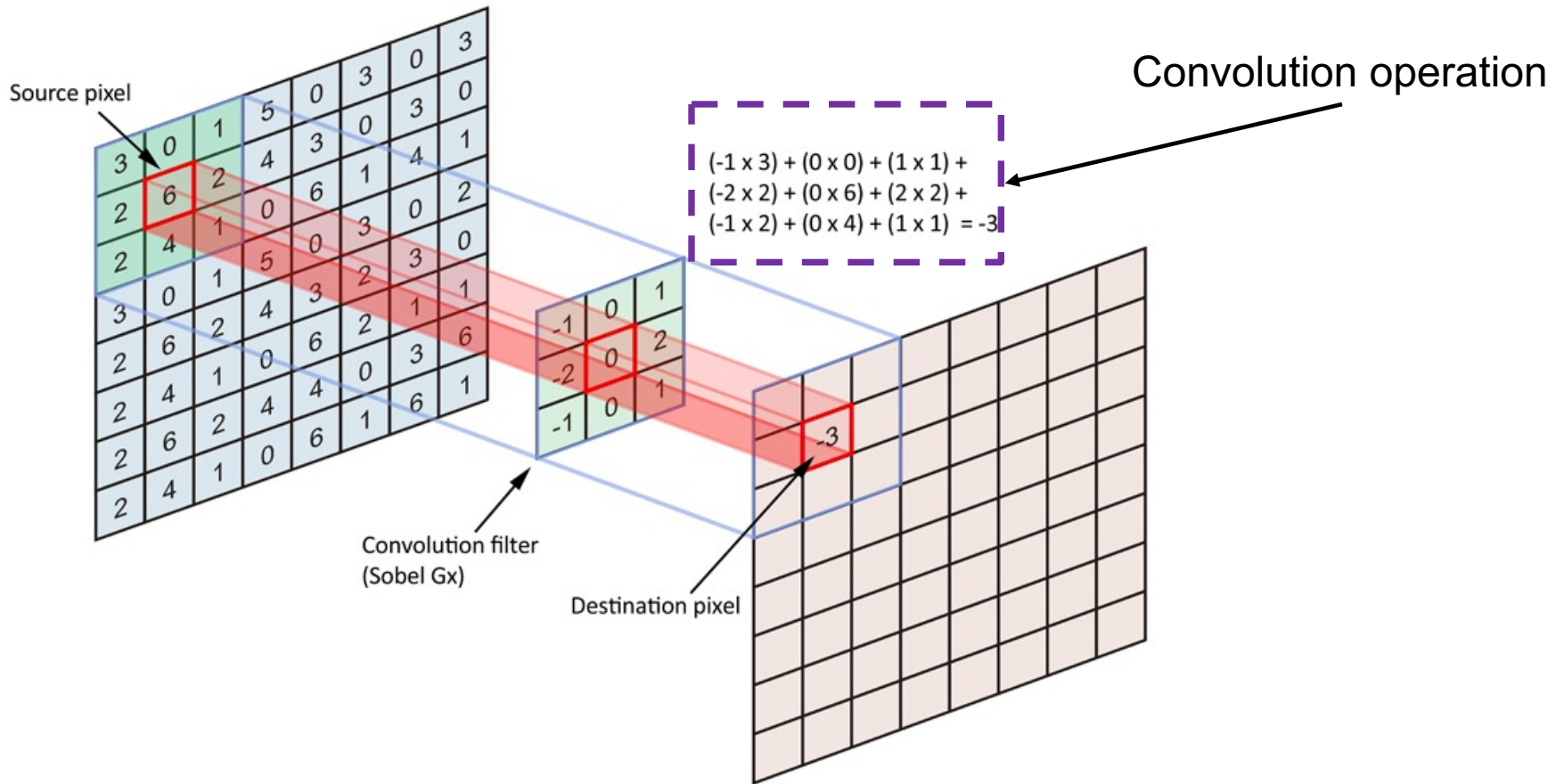
6 x 6 image

Less parameters!

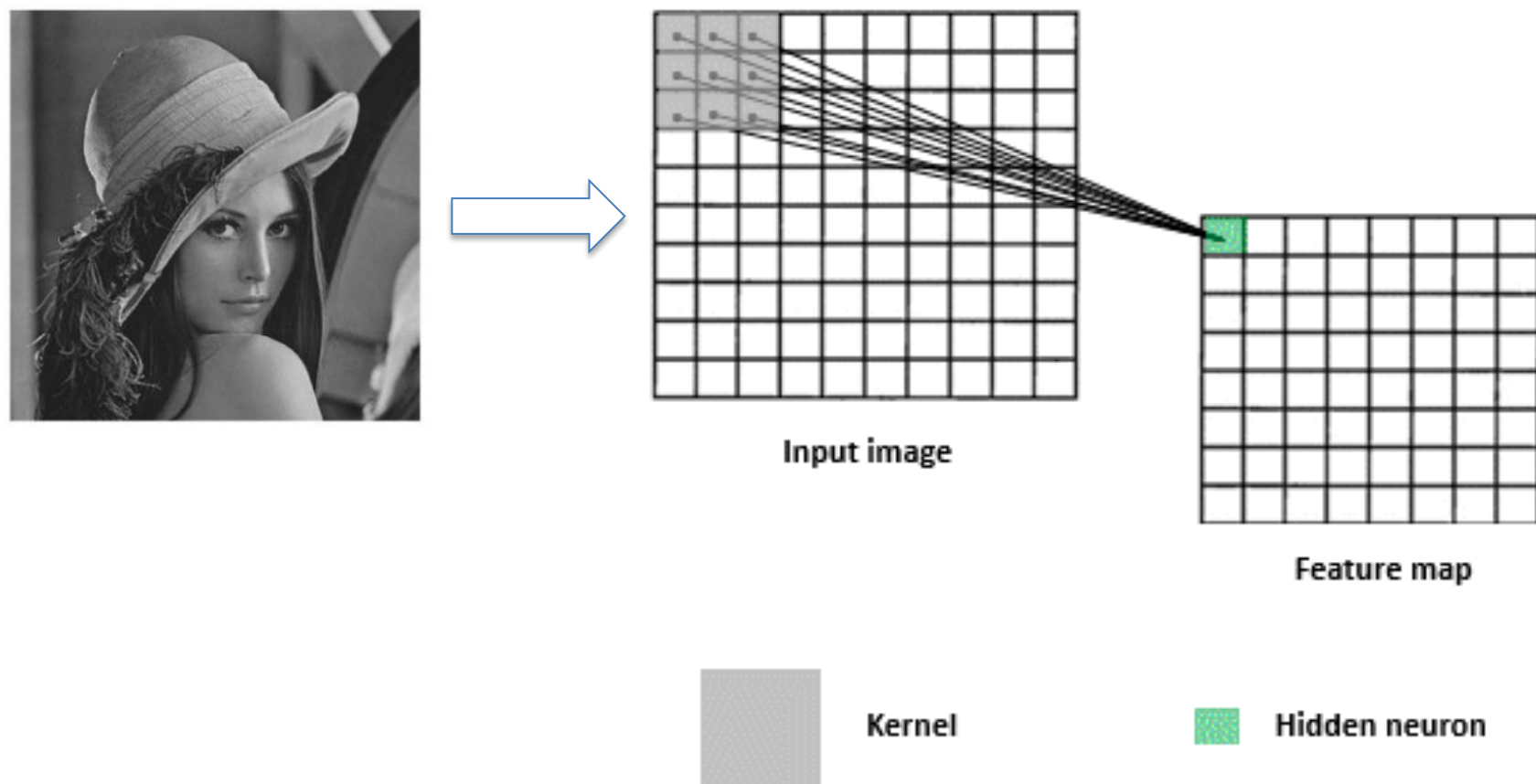
Even less parameters!



Convolution in Image

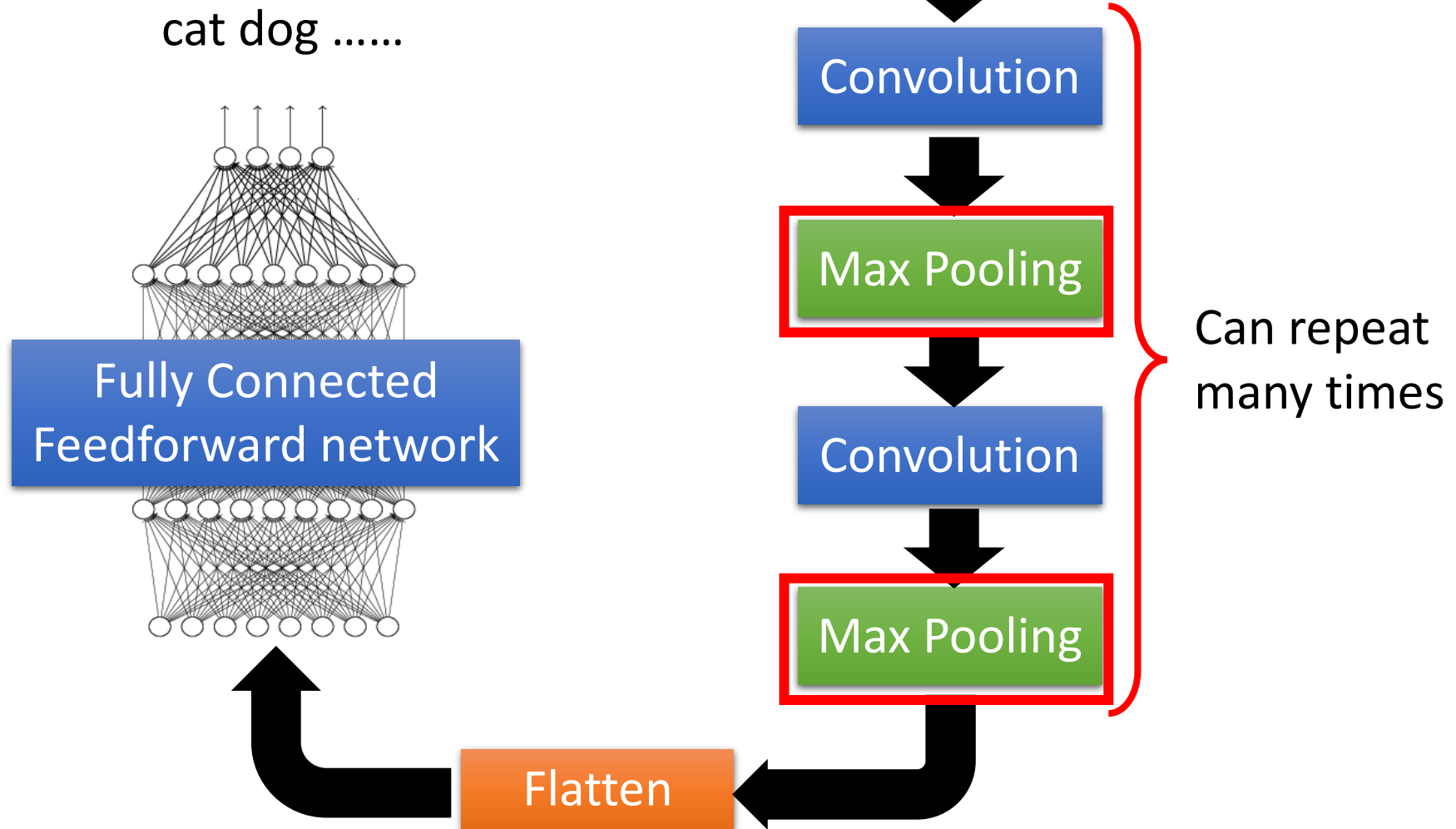


Standard Convolution Operation



Source: <https://www.ahmedbesbes.com/blog/introduction-to-cnns>

The whole CNN



CNN – Max Pooling

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

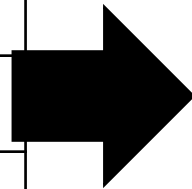
3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

-1	-1	-1	-1
-1	-1	-2	1
-1	-1	-2	1
-1	0	-4	3

CNN – Max Pooling

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

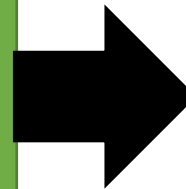
6 x 6 image



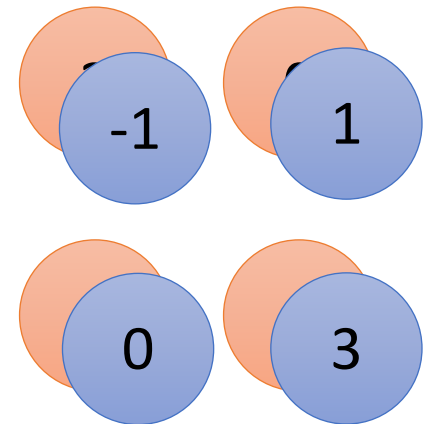
Conv



Max
Pooling



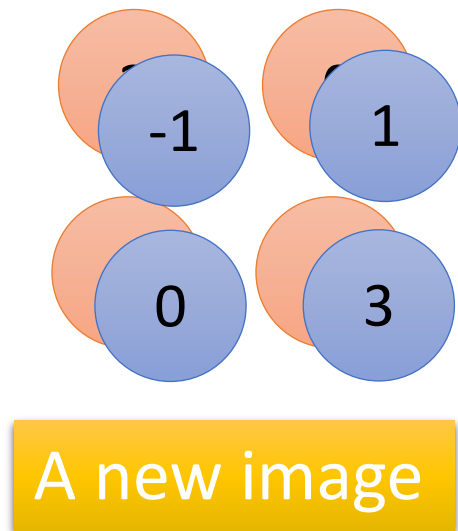
New image
but smaller



2 x 2 image

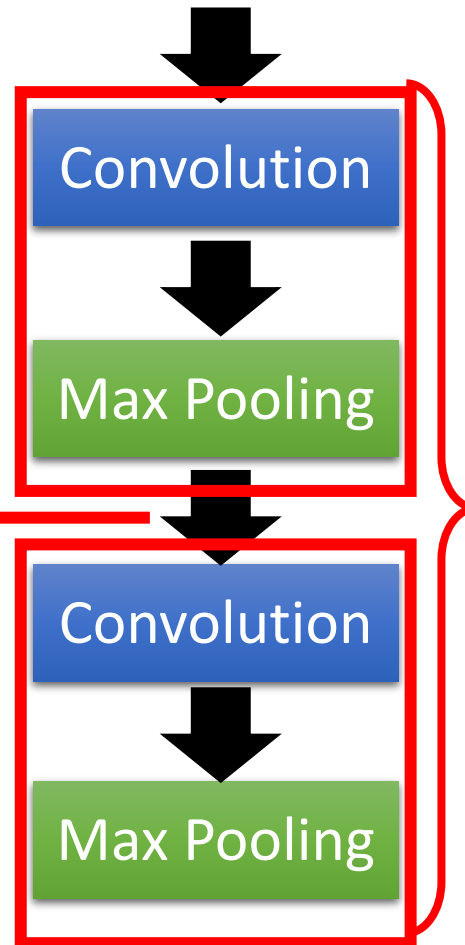
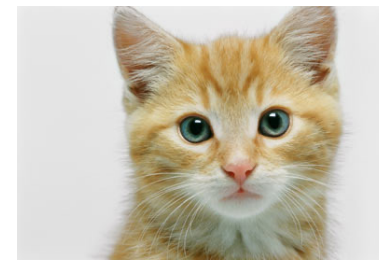
Each filter
is a channel

The whole CNN



Smaller than the original image

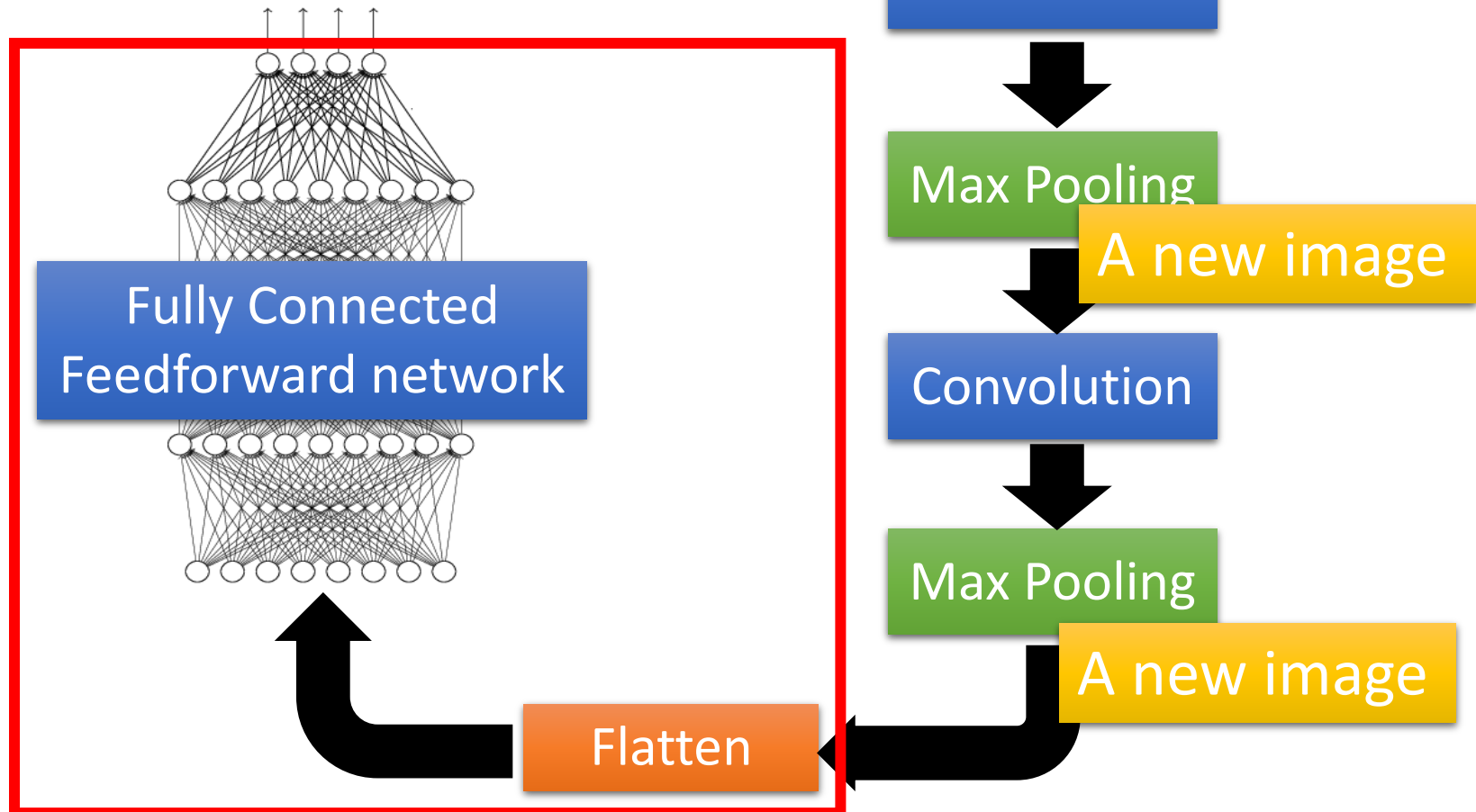
The number of the channel is the number of filters



Can repeat many times

The whole CNN

cat dog



Flatten

