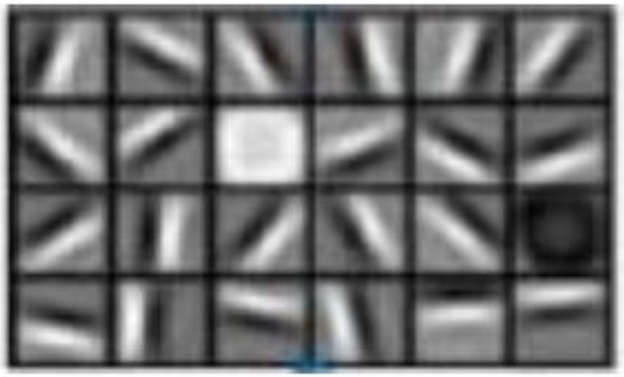
Convolutional Neural Network(CNN) model

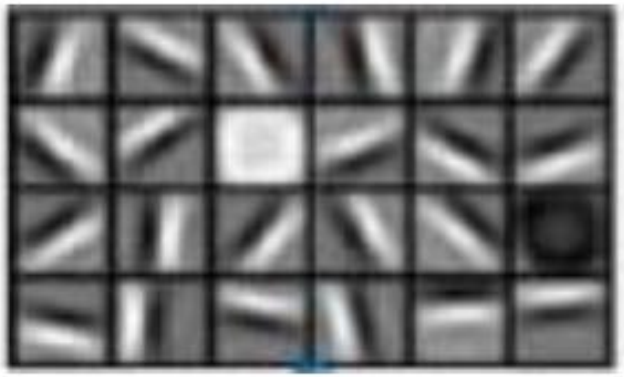
Issues with FFNN model

- Images are recognised by human brain inFFNN does not exploit redundancy(edges, blobs, etc., at different places) in image structure while learning. It tries to relearn those redundant patterns if appeared at different places.



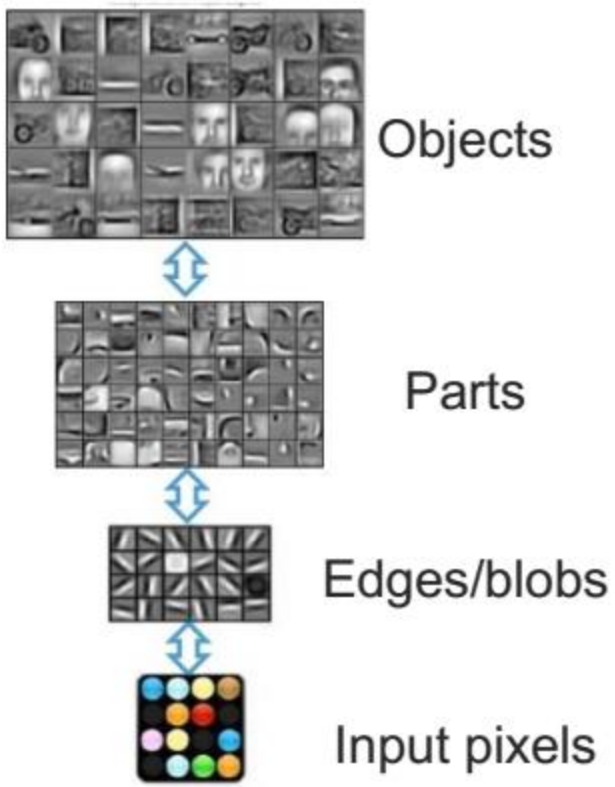
Issues with FFNN model

- FFNN does not exploit redundancy(edges, blobs, etc., at different places) in image structure while learning. It tries to relearn those redundant patterns if appeared at different places.



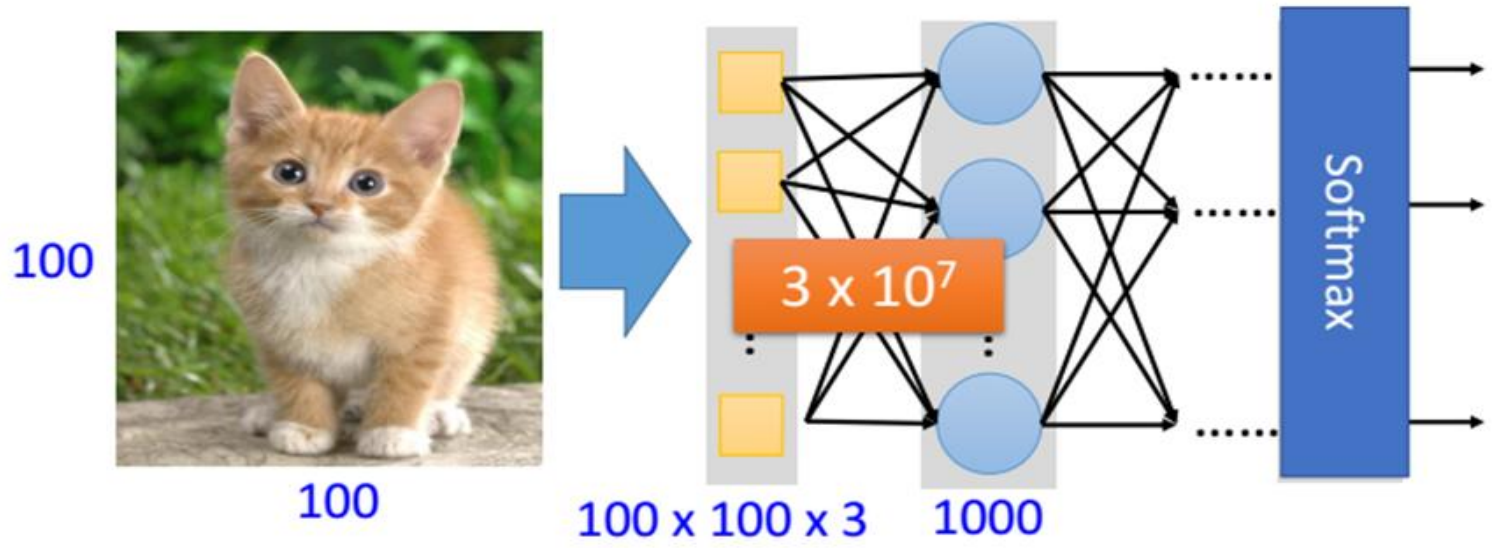
Issues with FFNN model

- Human visual system works in a filtered fashion. The hierarchy of features starts from basic low-level features(edges/blobs/..) to more high-level features. FFNN does not necessarily encourage visual abstraction.



Issues with FFNN model

- Too many parameters to learn in FFNN. For a small 100 × 100 pixel RGB image, we would have 30000 × n parameters for first layer alone.



Can the fully connected network be simplified by considering the above issues?

Interesting properties about image

Property 1

- Some patterns are much smaller(i.e.,local) than the whole image

Property 2

- The same patterns appear in different regions.

Property 3

- Subsampling the pixels will not change the object

Why CNN for Image?

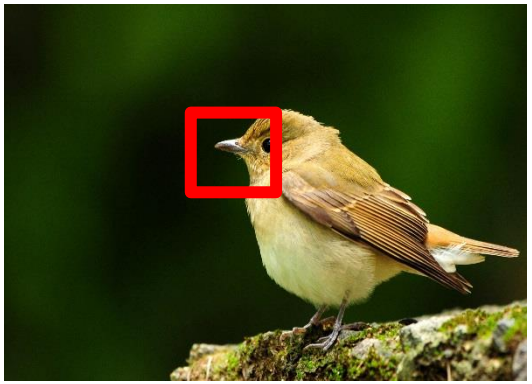
- Some patterns are much smaller than the whole image

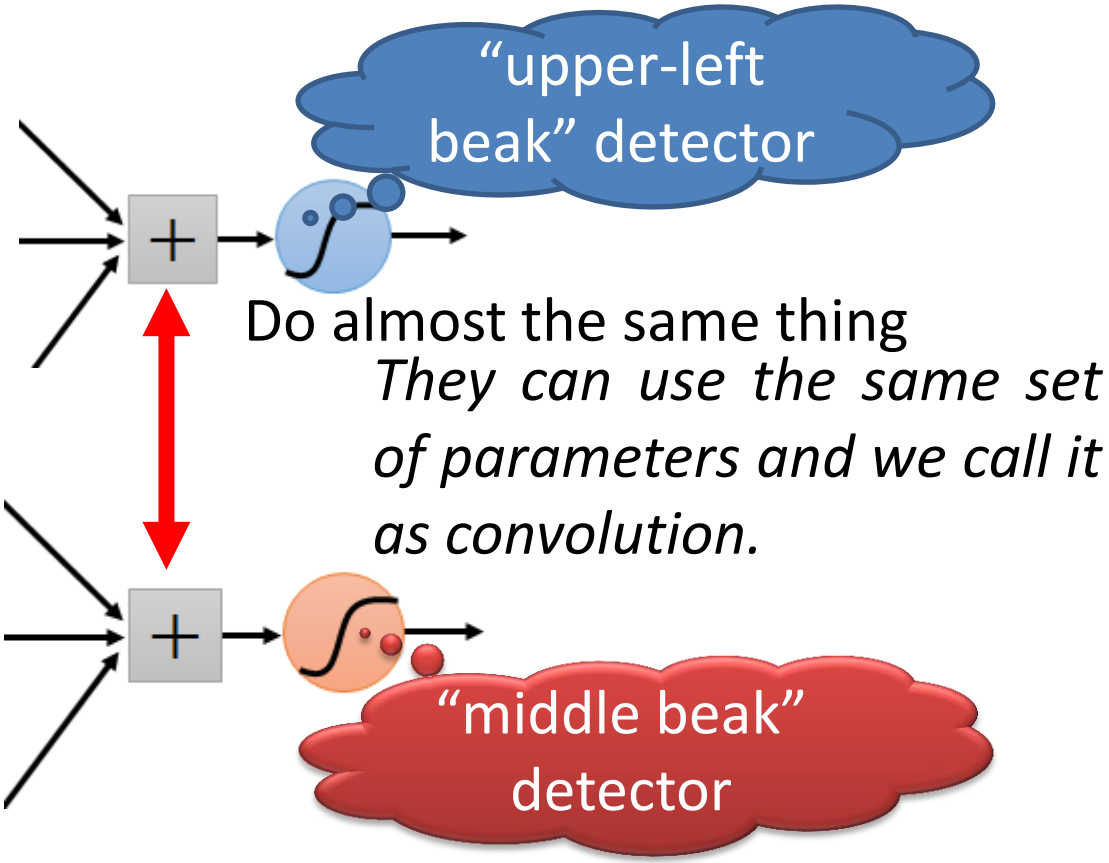
A neuron does not have to see the whole image to discover the pattern.

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

bird

subsampling

We can subsample the pixels to make image smaller

Less parameters for the network to process the image

The whole CNN

Convolution Layer

Max Pooling Layer

Convolution Layer

Max Pooling Layer

Can repeat many times

Flatten

Fully Connected Feedforward network

cat dog ……

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

Can repeat many times

Flatten

Convolution

The whole CNN

Convolution Layer

Max Pooling Layer

Convolution Layer

Max Pooling Layer

Can repeat many times

Flatten

Fully Connected Feedforward network

cat dog ……

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).

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

3 -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

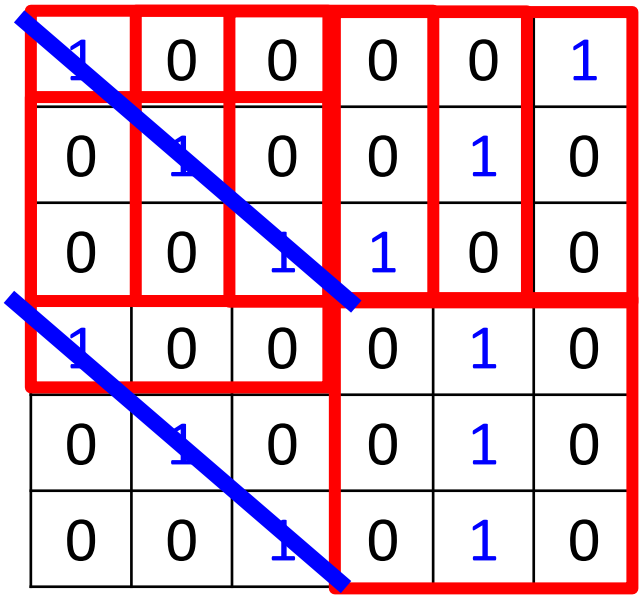
1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

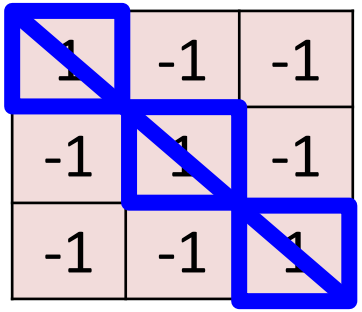
3 -3

We set stride=1 below

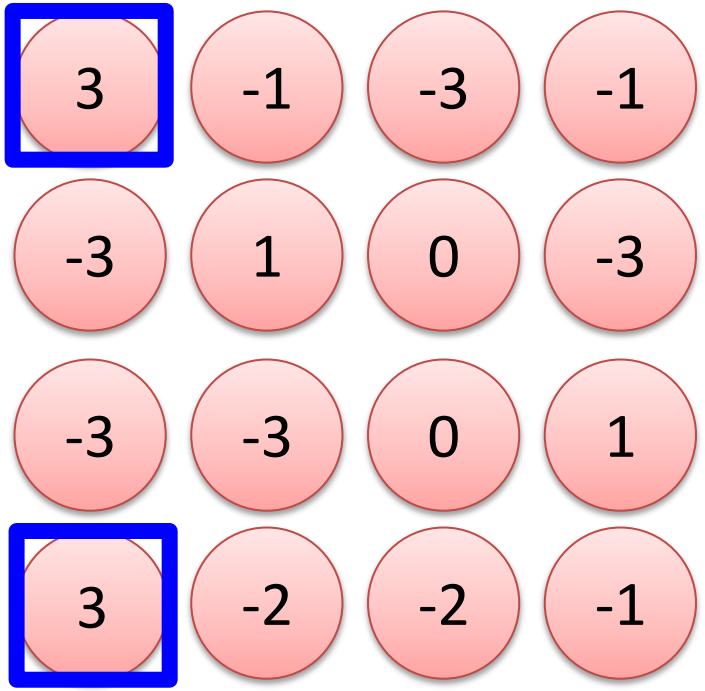
stride=1



6 x 6 image



Filter 1



Property 2

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

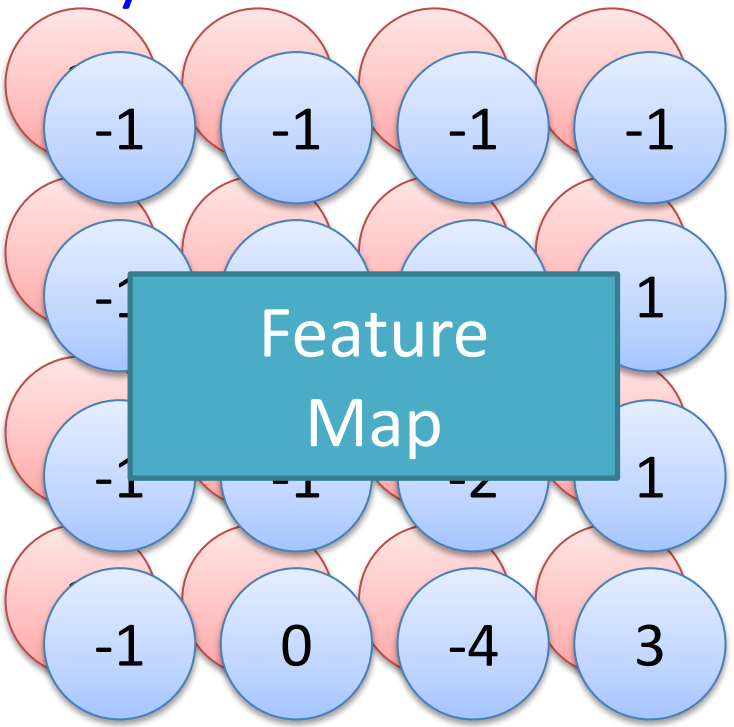
Filter 2

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 – Zero Padding

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

Filter 1

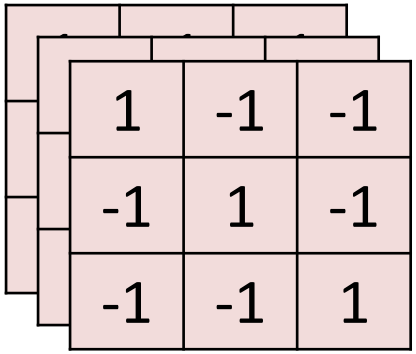
0	0	0				
0	1	0	0	0	0	1
0	0	1	0	0	1	0
	0	0	1	1	0	0
	1	0	0	0	1	0
	0	1	0	0	1	0
	0	0	1	0	1	0
				0	0	0

6 x 6 image

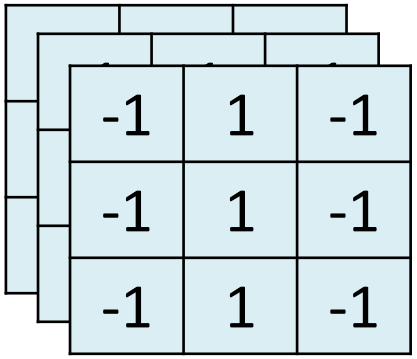
You will get another 6 x 6 images in this way

Zero padding

CNN – Color image

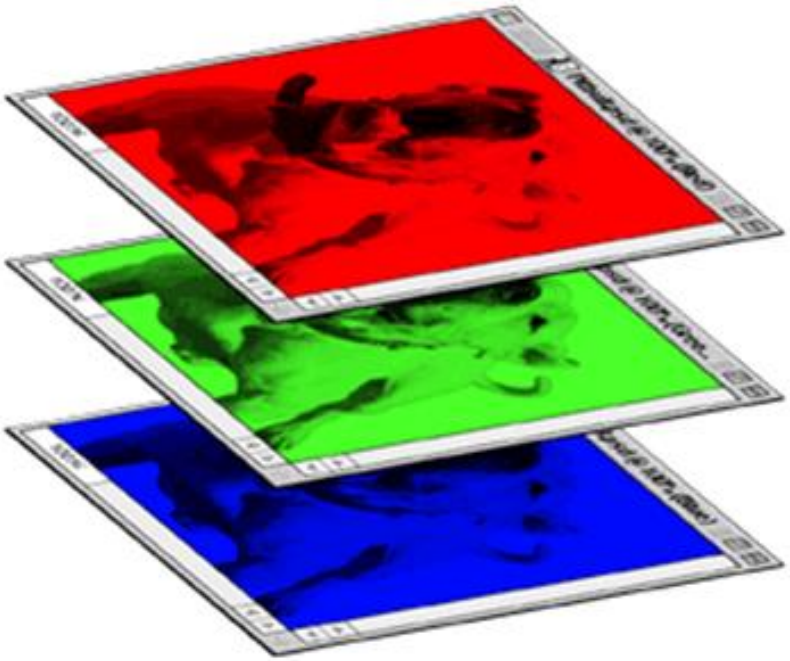


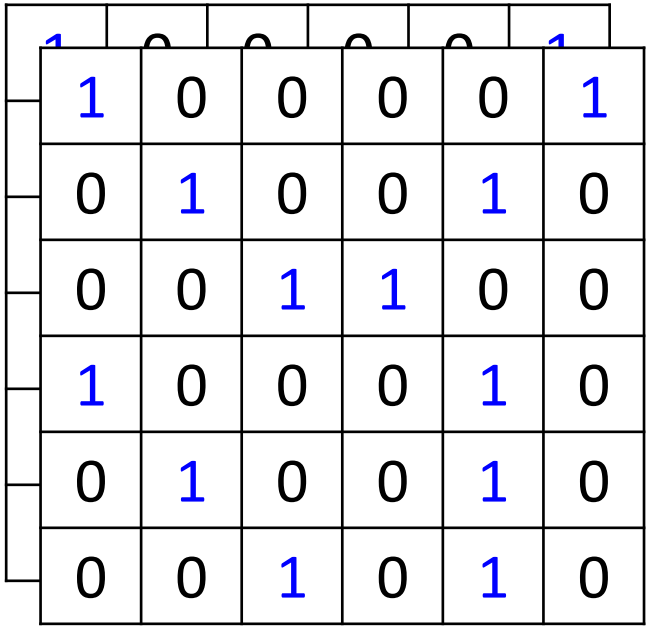
Filter 1



Filter 2

Colorful image





Pooling

The whole CNN

Convolution Layer

Max Pooling Layer

Convolution Layer

Max Pooling Layer

Can repeat many times

Flatten

Fully Connected Feedforward network

cat dog ……

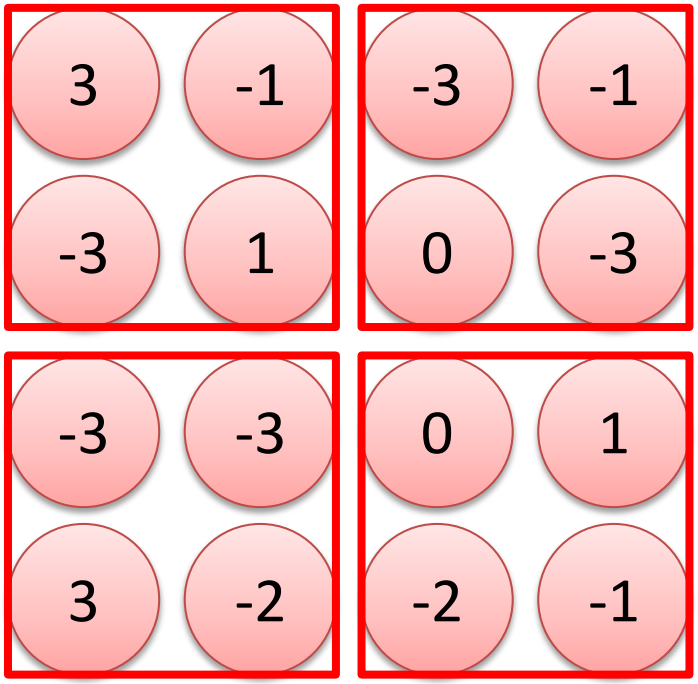
CNN – Max Pooling

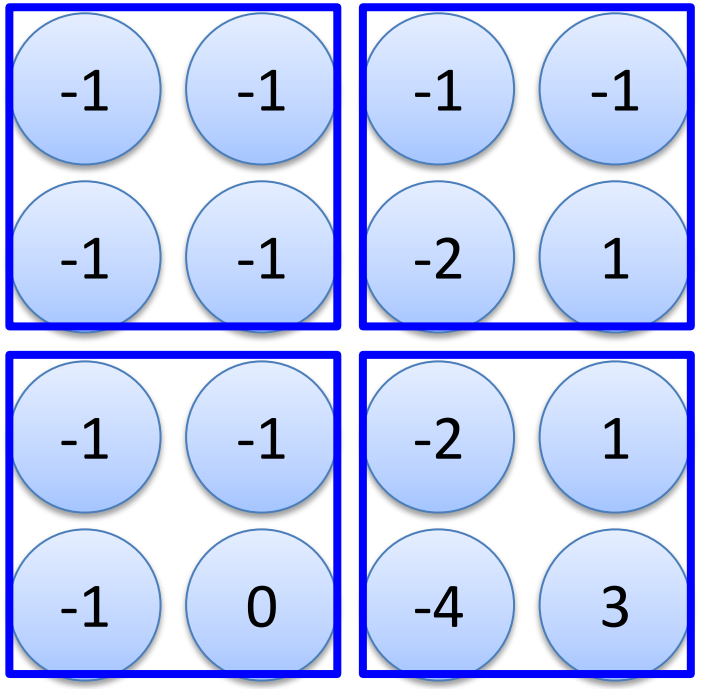
1	-1	-1
-1	1	-1
-1	-1	1

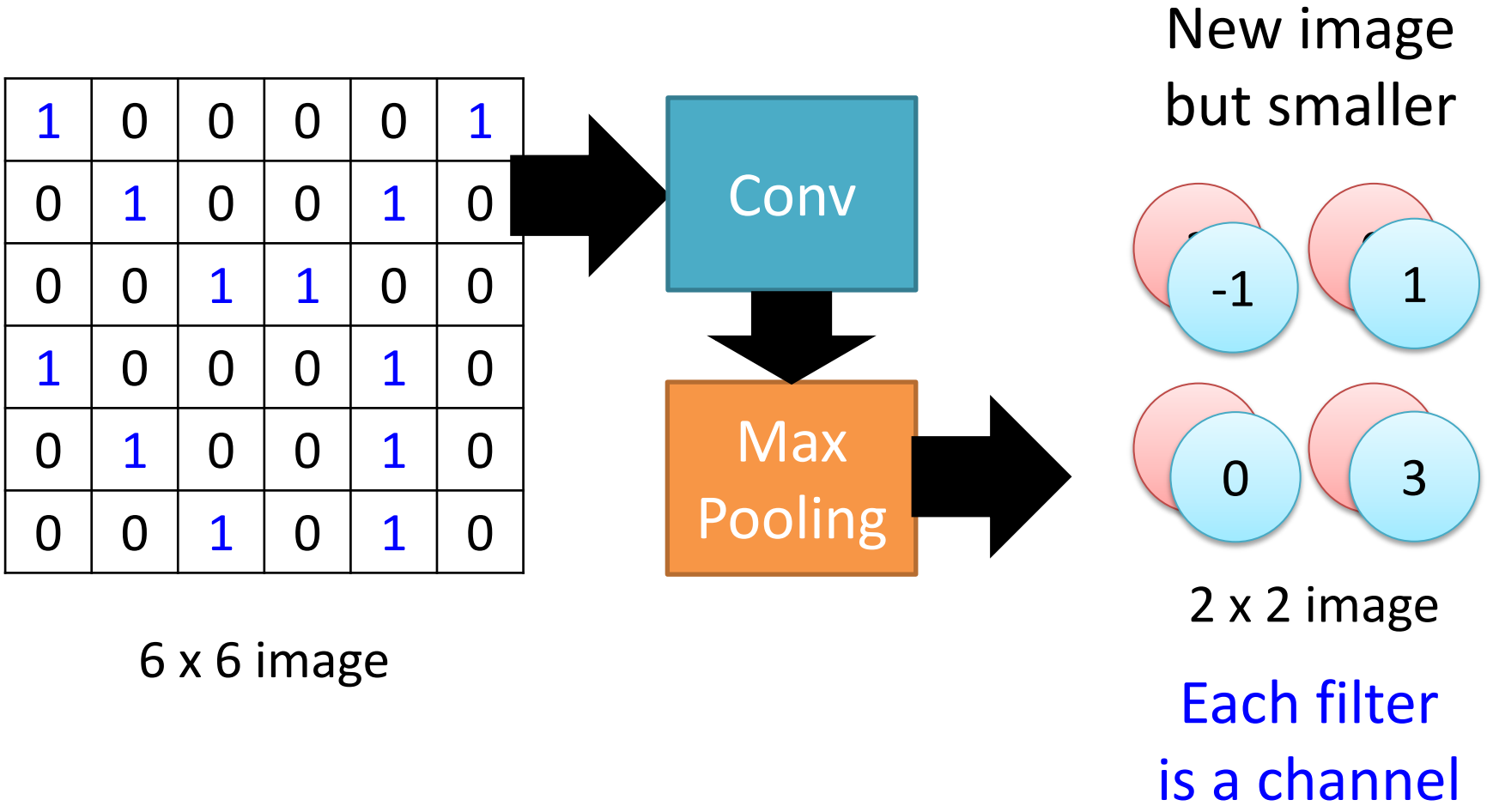
Filter 1

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

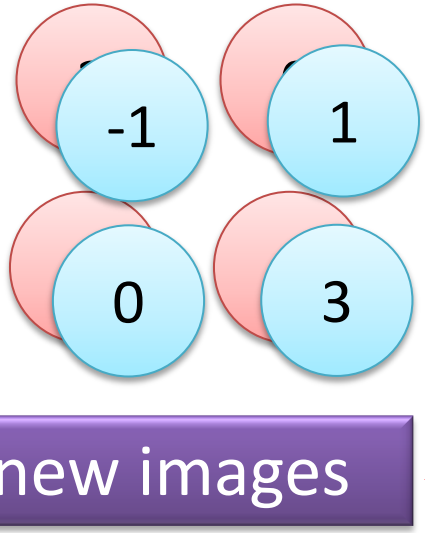
Filter 2

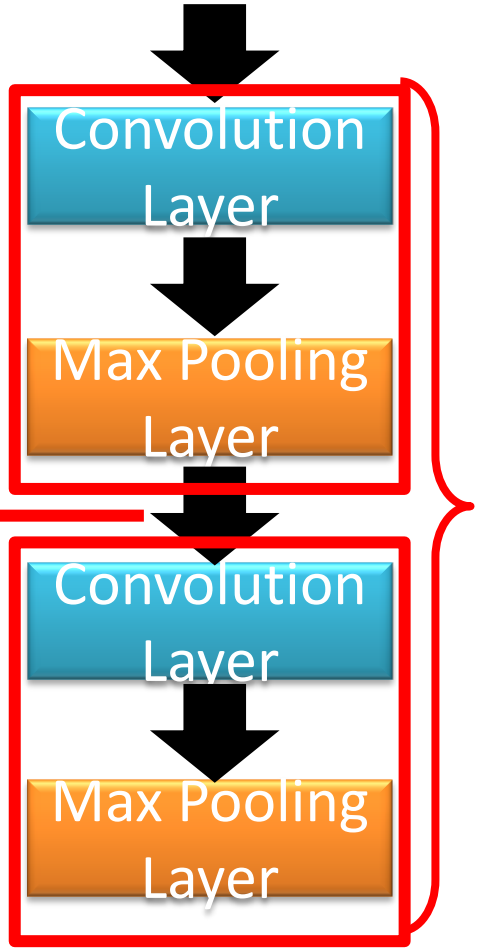





1 0 0 0 0 1
0 1 0 0 1 0
0 0 1 1 0 0
1 0 0 0 1 0
0 1 0 0 1 0
0 0 1 0 1 0
6 x 6 image
Conv
Max Pooling
New image but smaller
-1
1
0
3
2 x 2 image
Each filter is a channel

The whole CNN





Smaller than the original image

The number of the channels is the number of filters

Flatten

The whole CNN

cat dog ……

Convolution Layer

Max Pooling Layer

new images

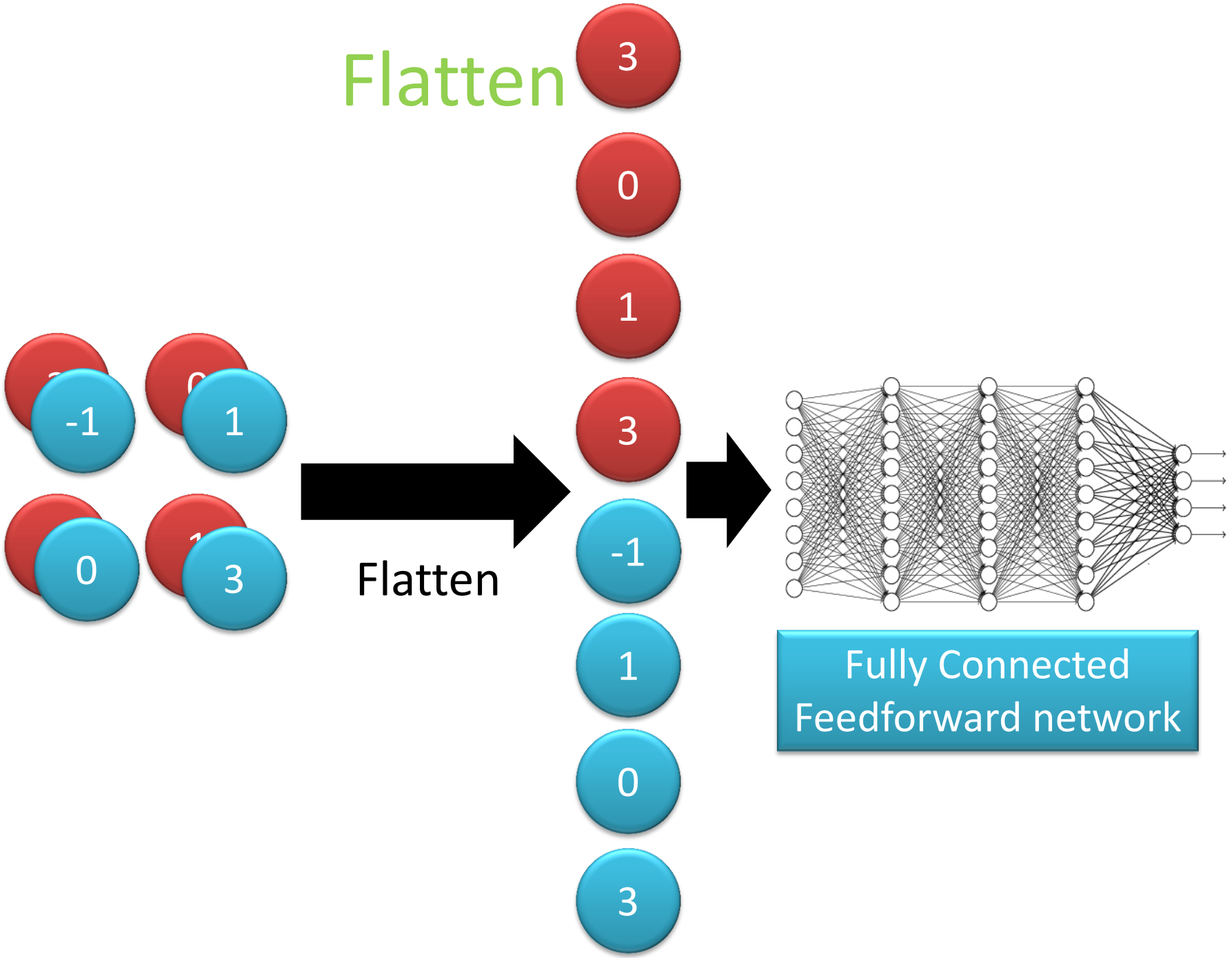
Convolution Layer

Max Pooling Layer

new images

Flatten

Fully Connected Feedforward network


Flatten
3
0
1
3
-1
1
0
3
-1
1
0
3
Flatten
Fully Connected
Feedforward network