CS 4602

Introduction to Machine Learning

Convolutional Neural Network

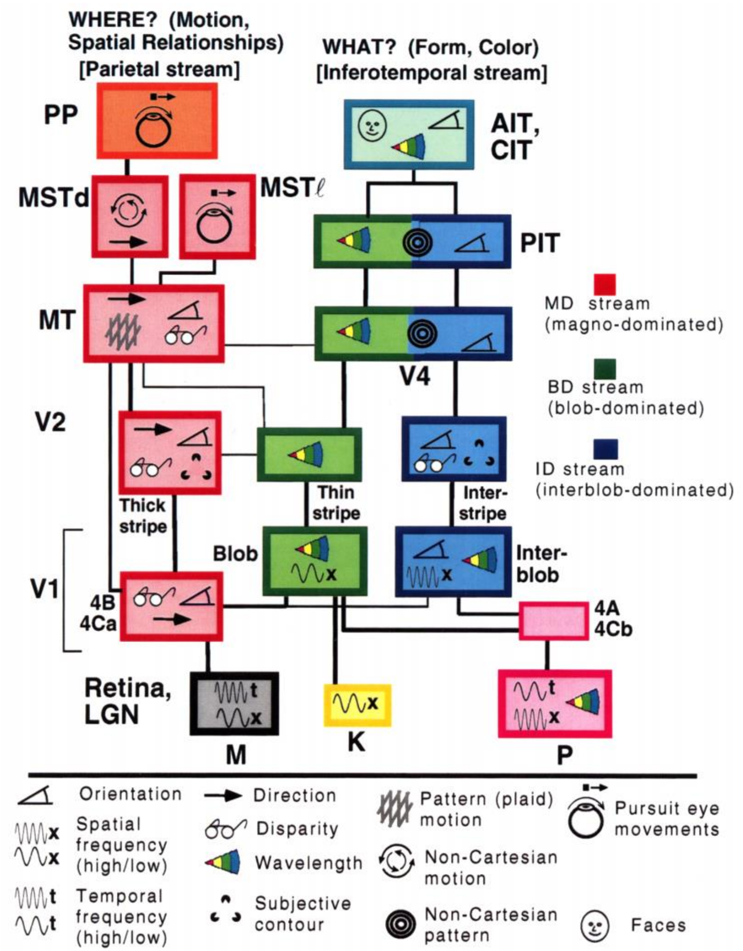
Instructor: Po-Chih Kuo

Roadmap

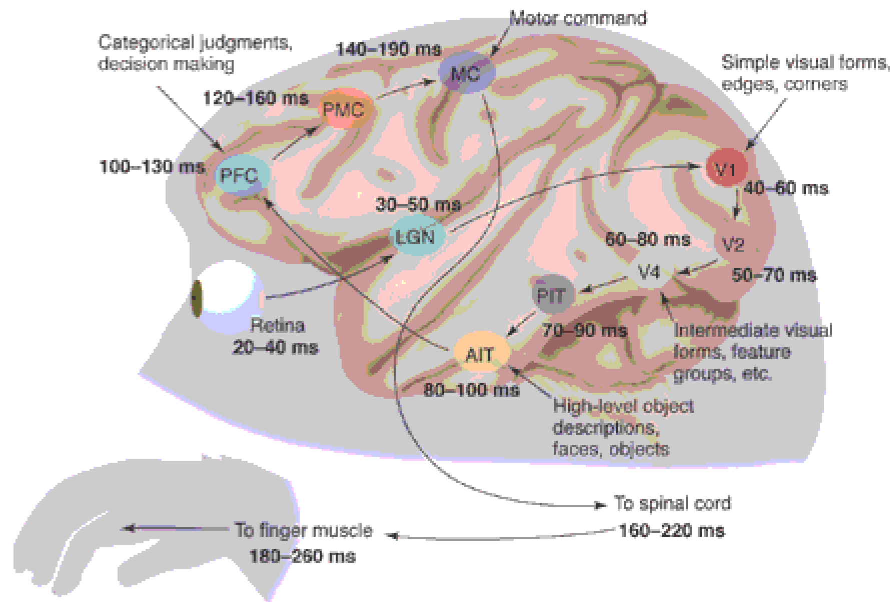
- Introduction and Basic Concepts
- Regression
- Bayesian Classifiers
- Decision Trees
- Linear Classifier
- Neural Networks
- Deep learning
- Convolutional Neural Networks
- Recurrent Neural Networks, Transformer, Reinforcement Learning…
- Clustering
- Dimensionality reduction
- Model Selection and Evaluation

How does the brain interpret images?

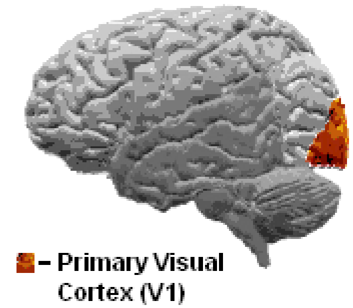
- The ventral (recognition) pathway in the visual cortex has multiple stages
- Retina - LGN - V1 - V2 - V4 - PIT - AIT ....



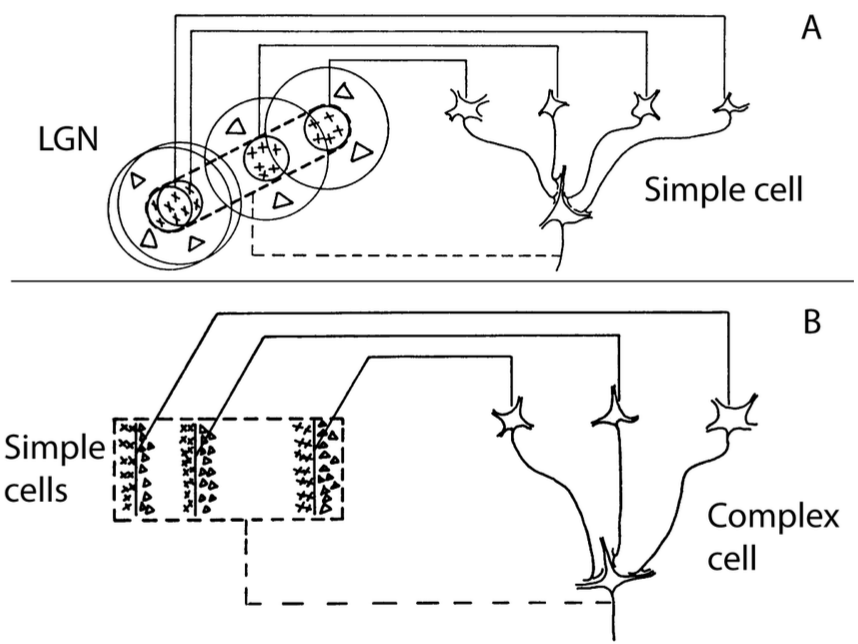
Gallant & Van Essen, 1994

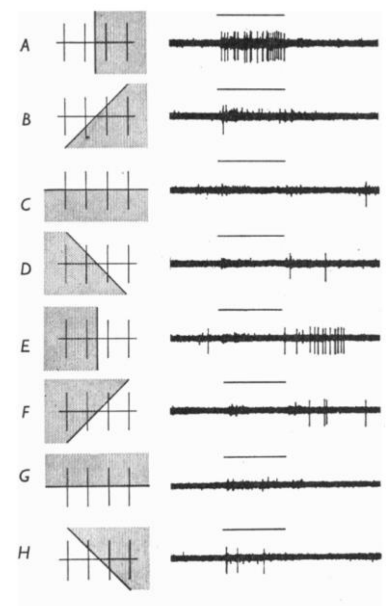


Visual Cortex



- Hubel & Wiesel 1962
- Simple cells detect local features
- Complex cells “pool” the outputs of simple cells



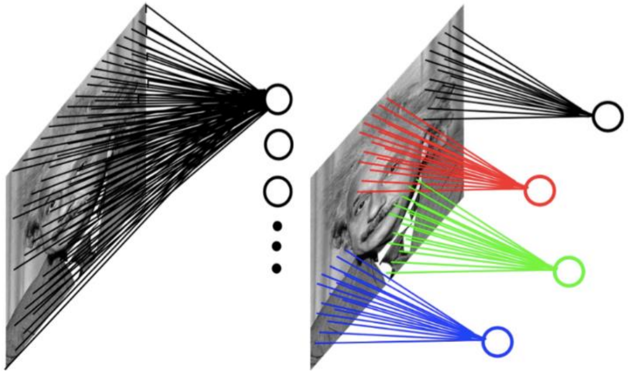


Convolutional Neural Network (CNN)

- Fully-connected neural network needs a large amount of parameters.
- CNNs are a special type of neural network whose hidden units are only connected to **local receptive field**
  - The number of parameters needed by CNNs is much smaller.

Example: 200x200 image

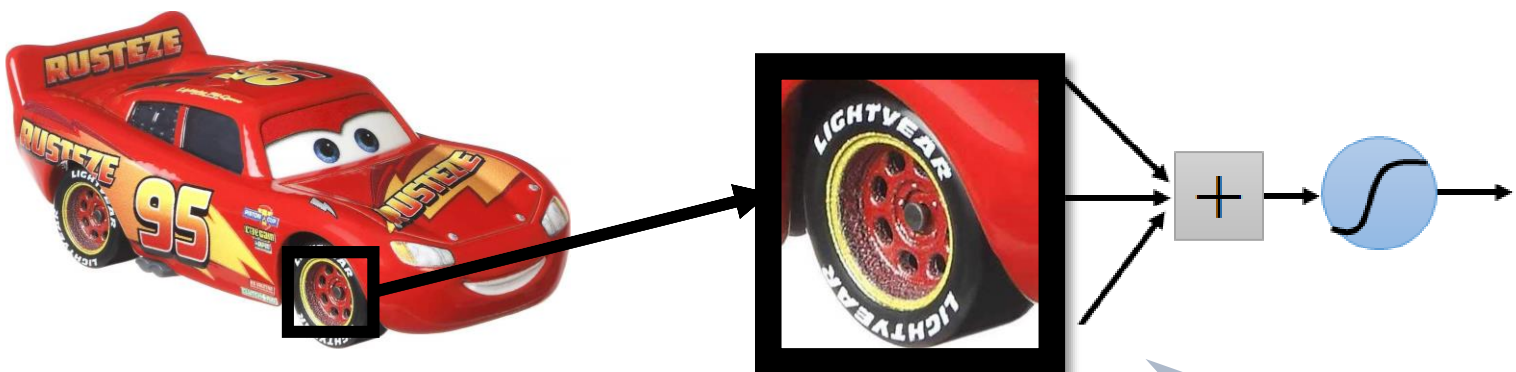
fully connected:
100 hidden units
=> 4,000,000 parameters



CNN:
5x5 kernel, 100 filters
=> 2,500 parameters

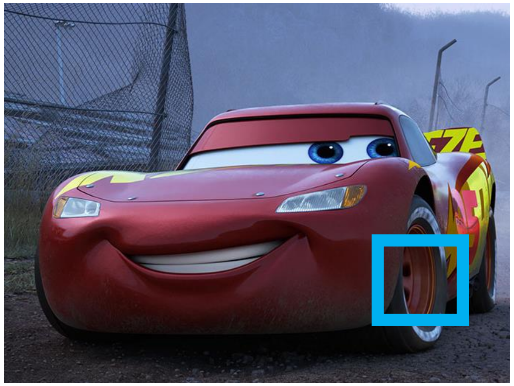
Learning a pattern

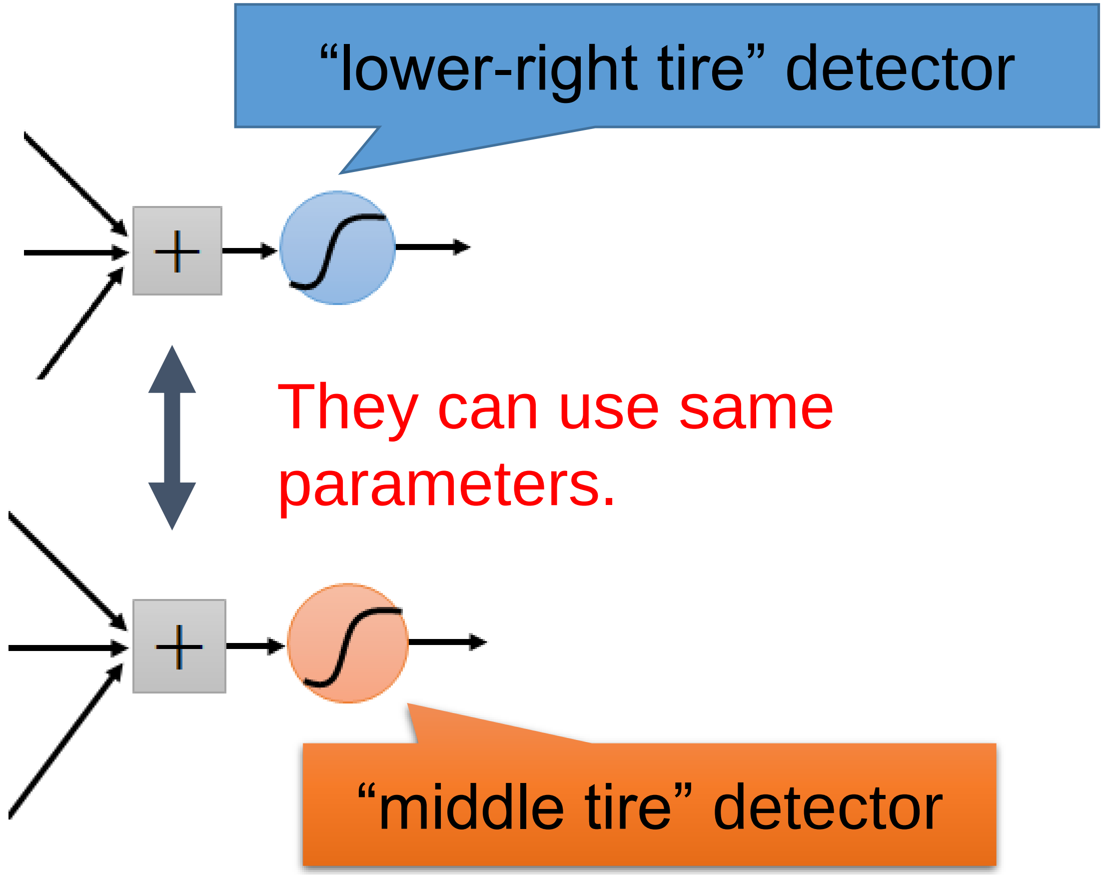
- Some patterns are much smaller than the whole image
- Can represent a small region with fewer parameters



“tire” detector

- Same pattern appears in different places

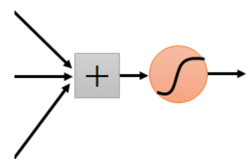




They can use same parameters.

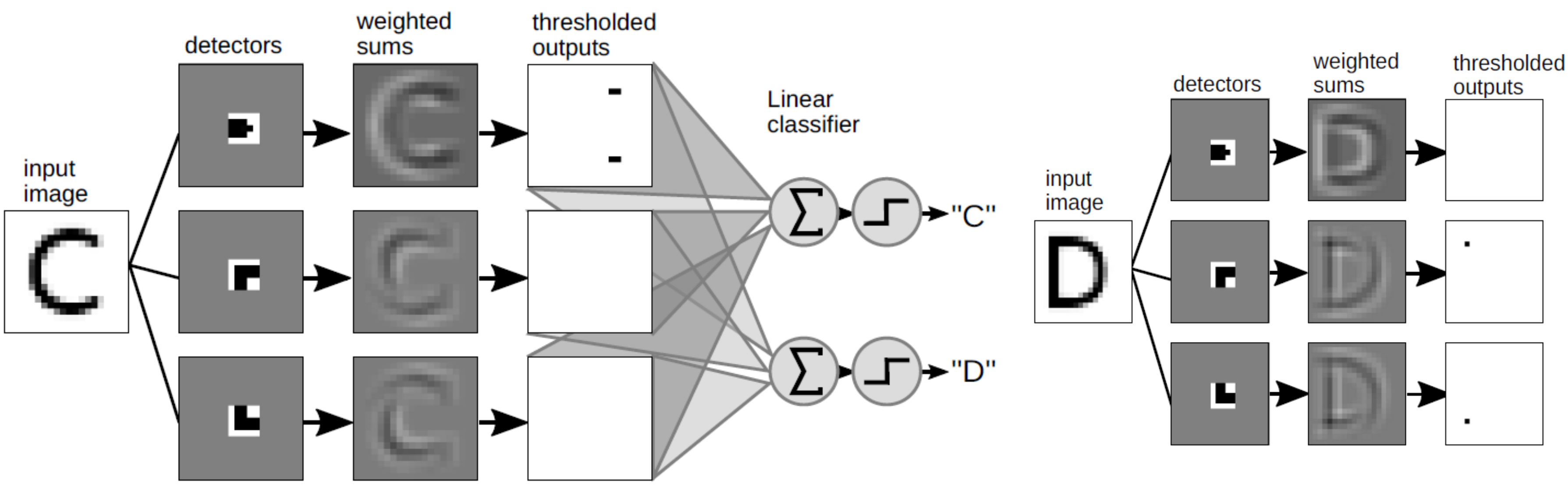
What about each detector can move around?

- A moving detector



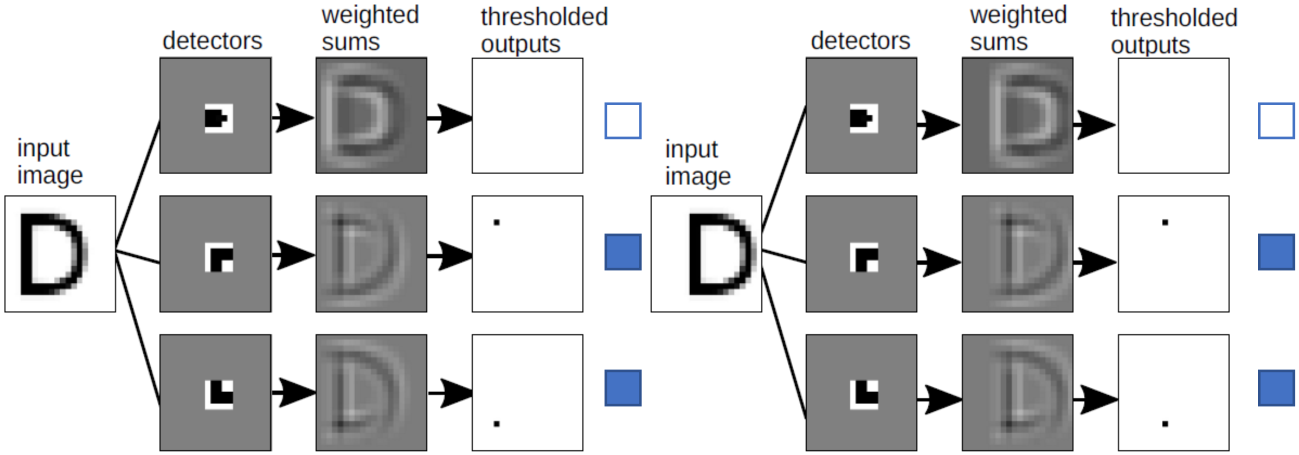
Detecting Motifs in Images

- Swipe “templates” over the image to detect motifs



Detecting Motifs in Images

- Shift invariance

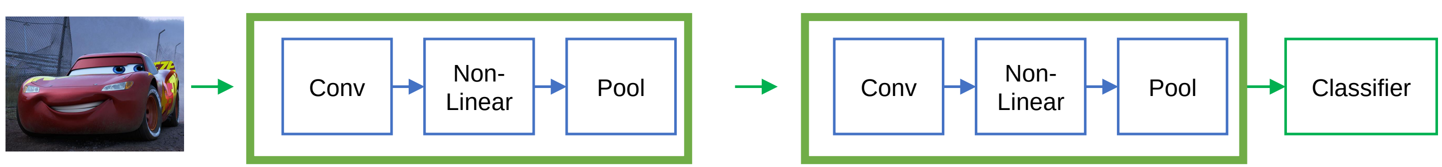


How can we have similar outputs from the model?

To combine outputs in a common pool!

Overall Architecture

- Multiple stages: Normalization→ Convolution→Non-Linearity→Pooling
  - Normalization: average removal, variance normalization...
  - Convolution: dimension expansion, projection on basis…
  - Non-Linearity: Rectification (ReLU), tanh…
  - Pooling: Max, average…



Convolution

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

Filter 1

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

Dot product → 3 -1

6 x 6 image

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

3 -3

Convolution

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

Filter 1

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

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

Convolution

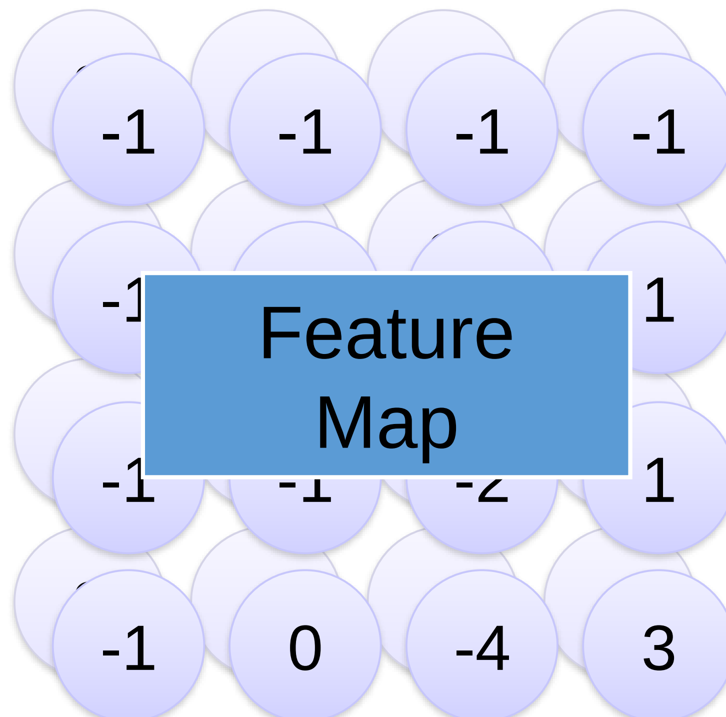
-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

stride=1

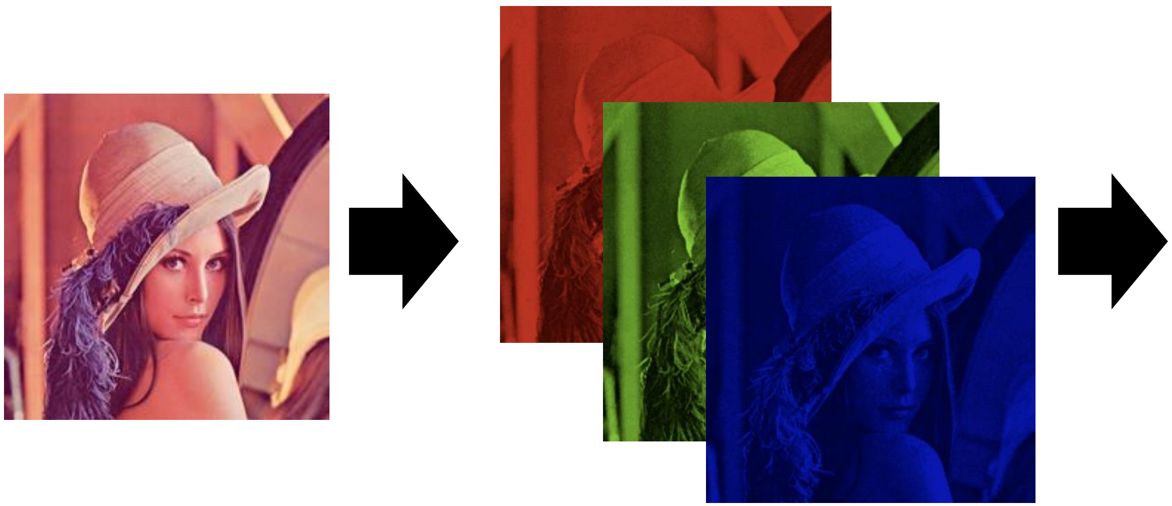
1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image



Two 4 x 4 images
Forming 2 x 4 x 4 matrix

Color image (3 channels)



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

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

Filter 1

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

Filter 2

Padding

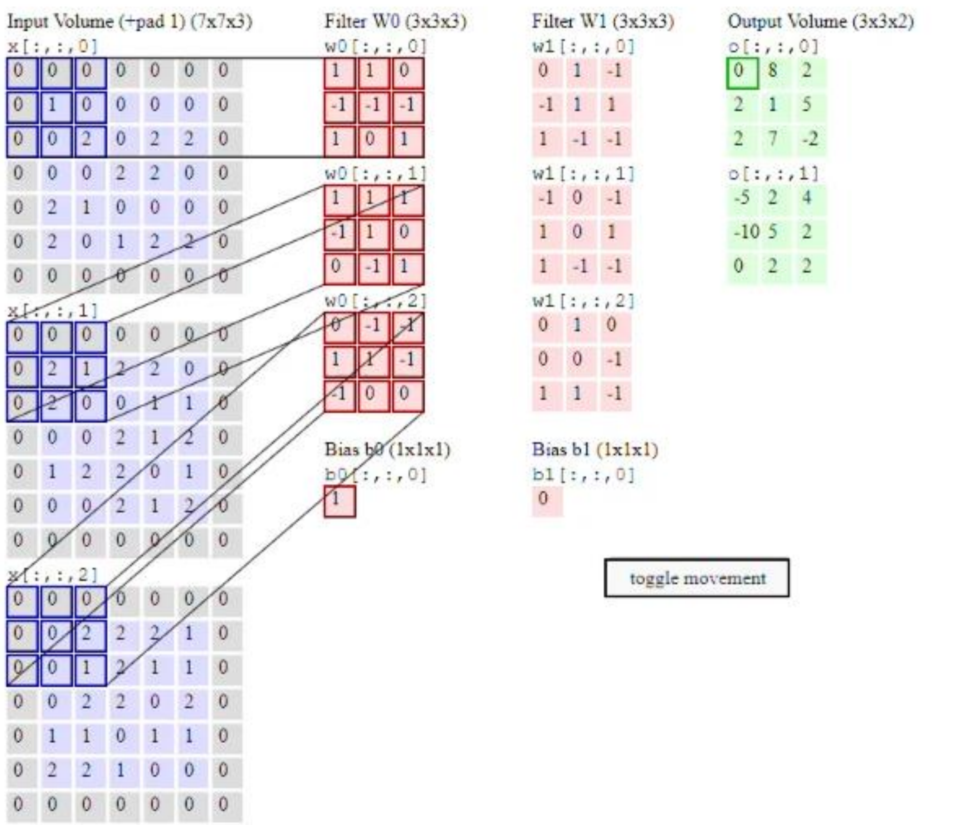
- Conv 3x3 with stride=1, padding=1

0	0	0	0	0	0
0	1	5	3	9	0
0	4	4	3	5	0
0	6	4	2	6	0
0	6	5	2	1	0
0	0	0	0	0	0

4 x 4 image

14	24	33	24
27	41	32	25
33	34	32	26
26	32	27	16

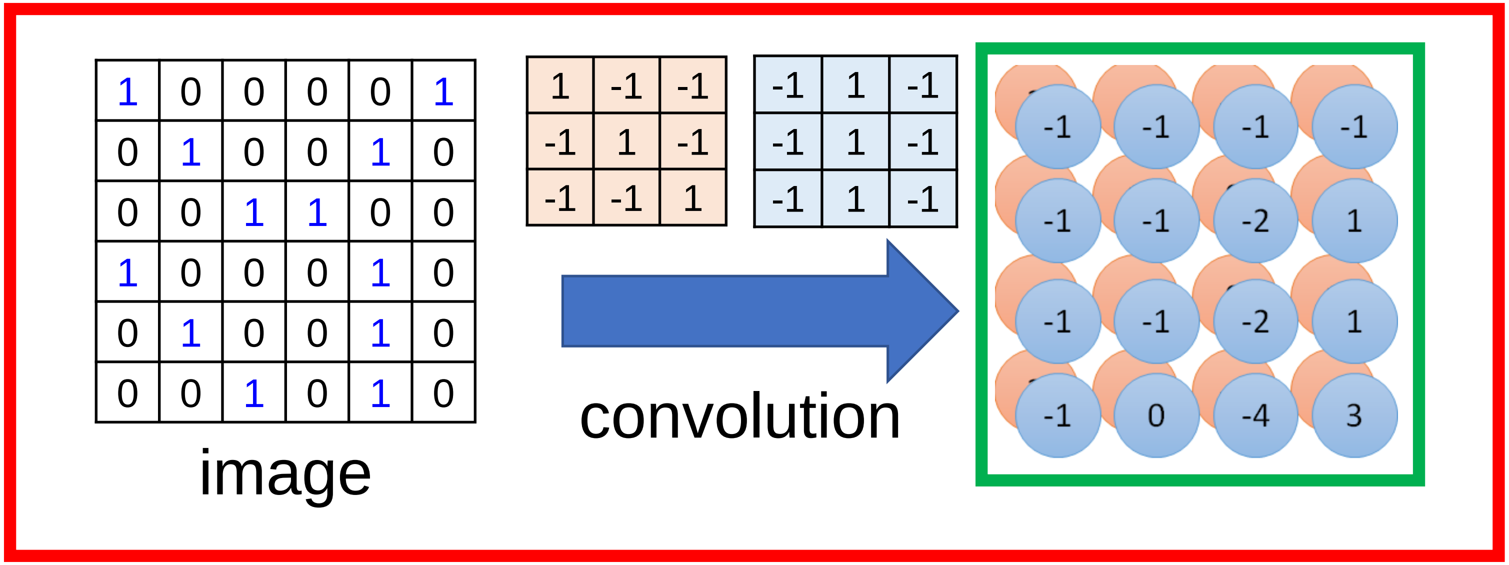
4 x 4 image



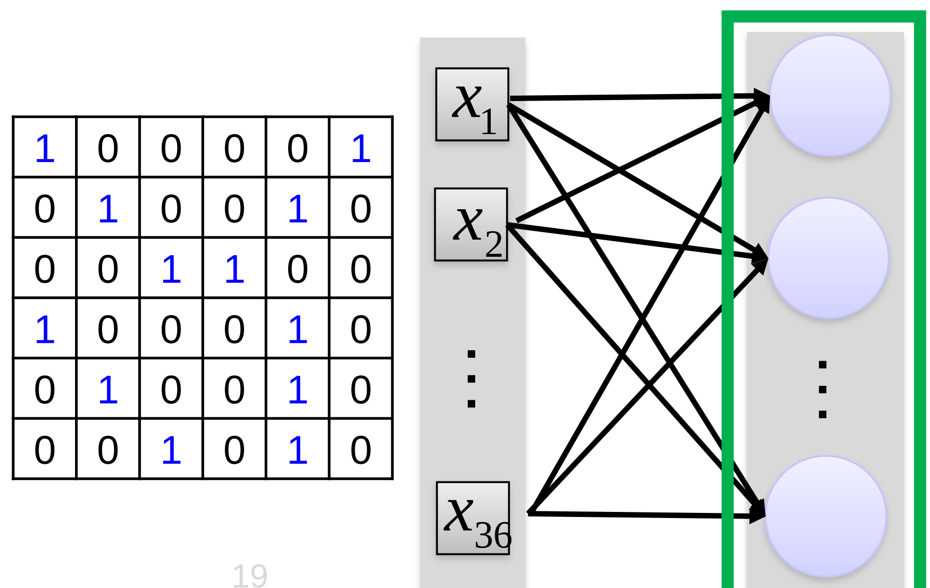
Implementation as Matrix Multiplication. Note that the convolution operation essentially performs dot products between the filters and local regions of the input. A common implementation pattern of the CONV layer is to take advantage of this fact and formulate the forward pass of a convolutional layer as one big matrix multiply as follows:

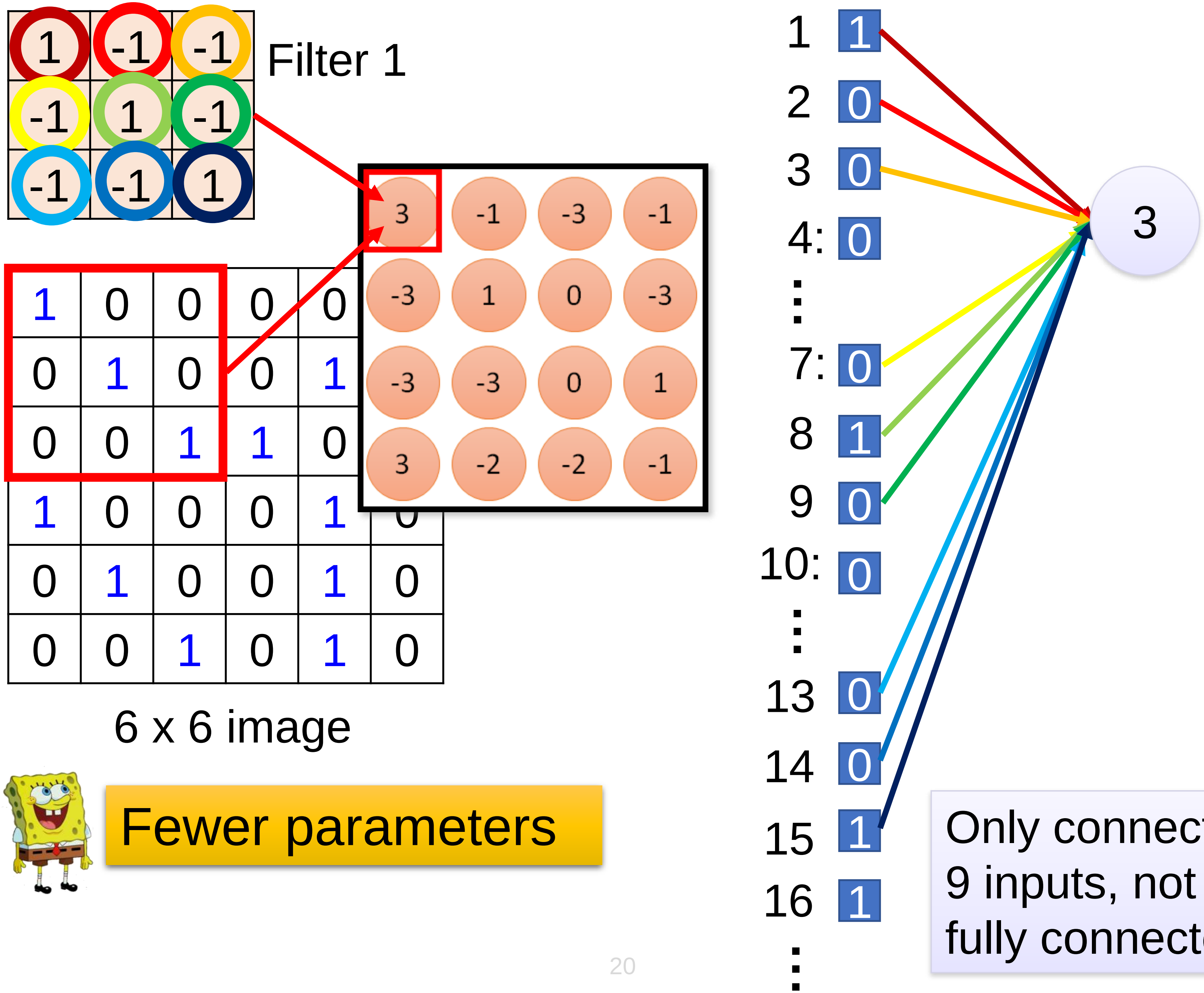
https://cs231n.github.io/convolutional-networks/

Convolution v.s. Fully Connected



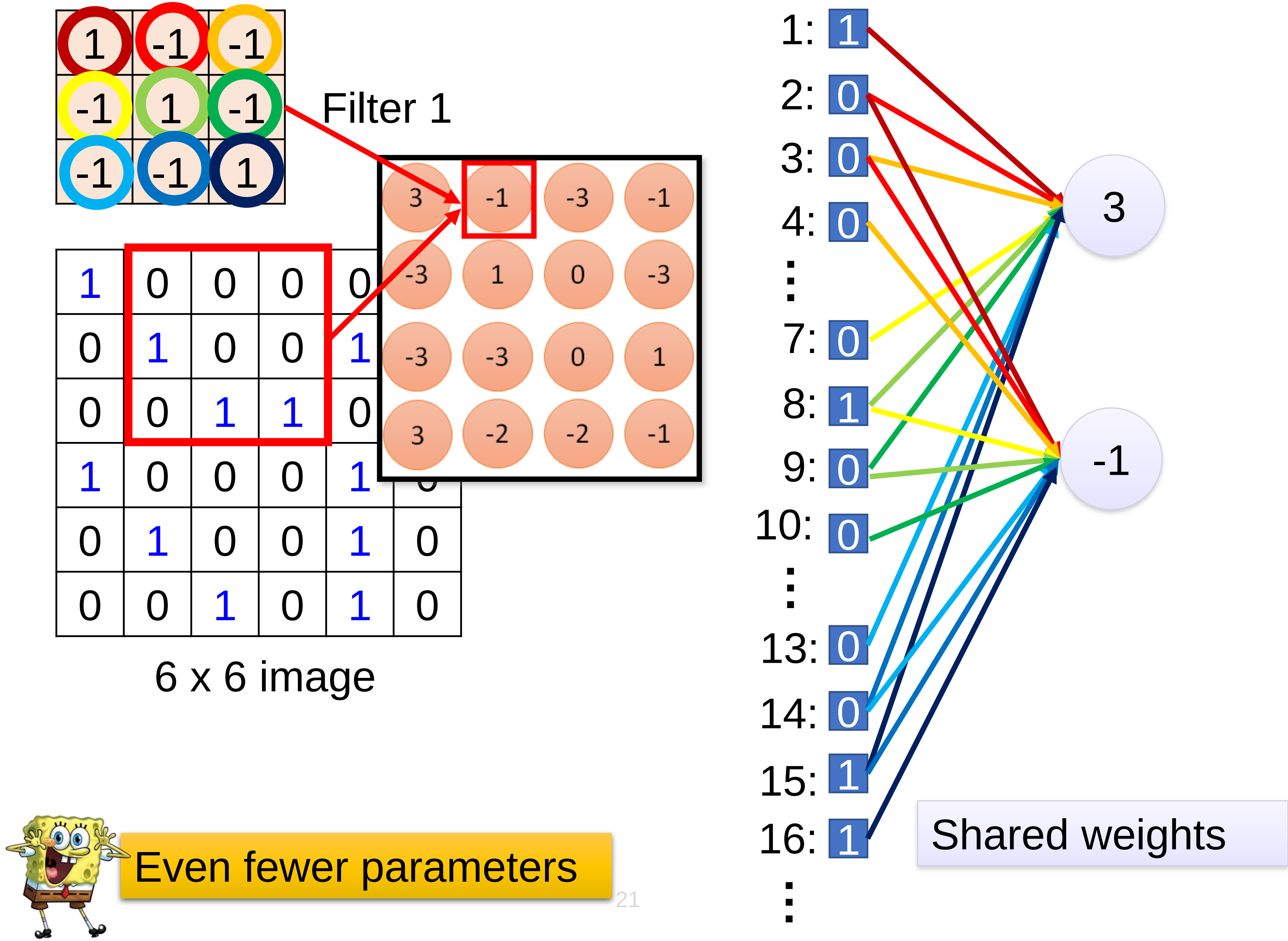
Fully-connected





Fewer parameters

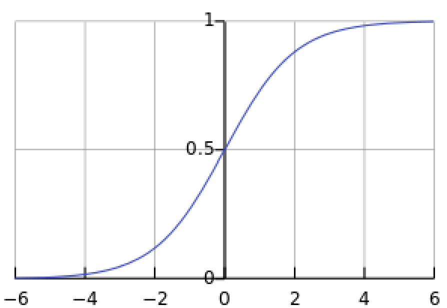
Only connect to 9 inputs, not fully connected

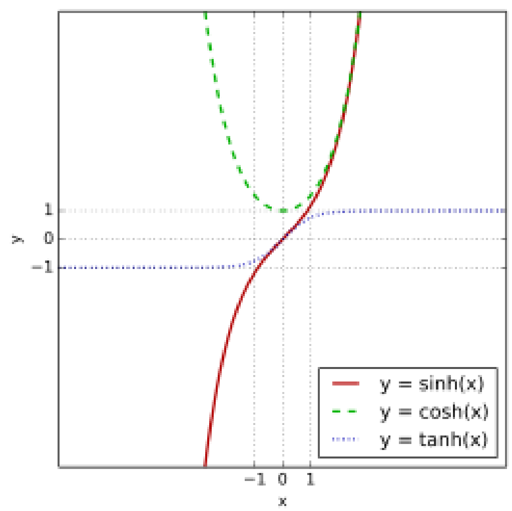


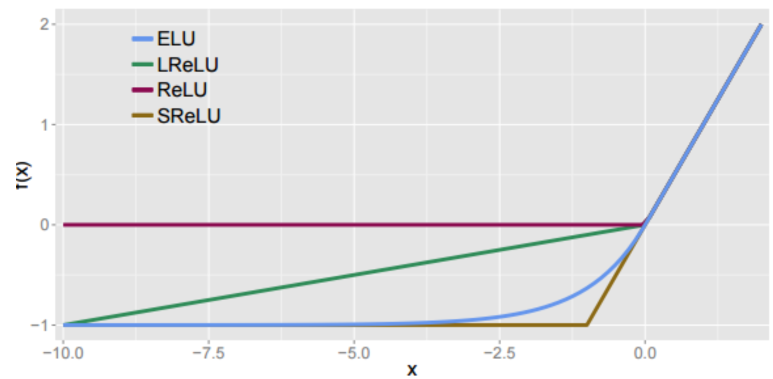
Filter 1
1 -1 -1
-1 1 -1
-1 -1 1
6 x 6 image
1 0 0 0 0 1
0 1 0 0 1 0
0 0 1 1 0 0
1 0 0 0 1 0
0 1 0 0 1 0
0 0 1 0 1 0
3 -1 -3 -1
-3 1 0 -3
-3 -3 0 1
3 -2 -2 -1
1: 1
2: 0
3: 0
4: 0
7: 0
8: 1
9: 0
10: 0
13: 0
14: 0
15: 1
16: 1
3
-1
Shared weights
Even fewer parameters

Nonlinear Activations

- Why activation? Nonlinearity
  - Sigmoid
  - tanh
  - ReLU family

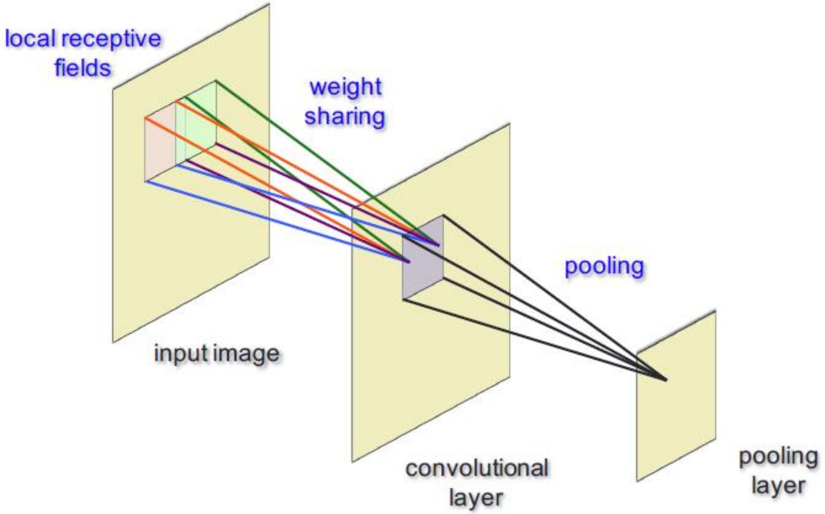






Pooling

- Common pooling operations:
  - Max pooling: reports the maximum output within a rectangular neighborhood.
  - Average pooling: reports the average output of a rectangular neighborhood (possibly weighted by the distance from the central pixel).



Pooling Example (Summing or averaging)

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

Convolved feature

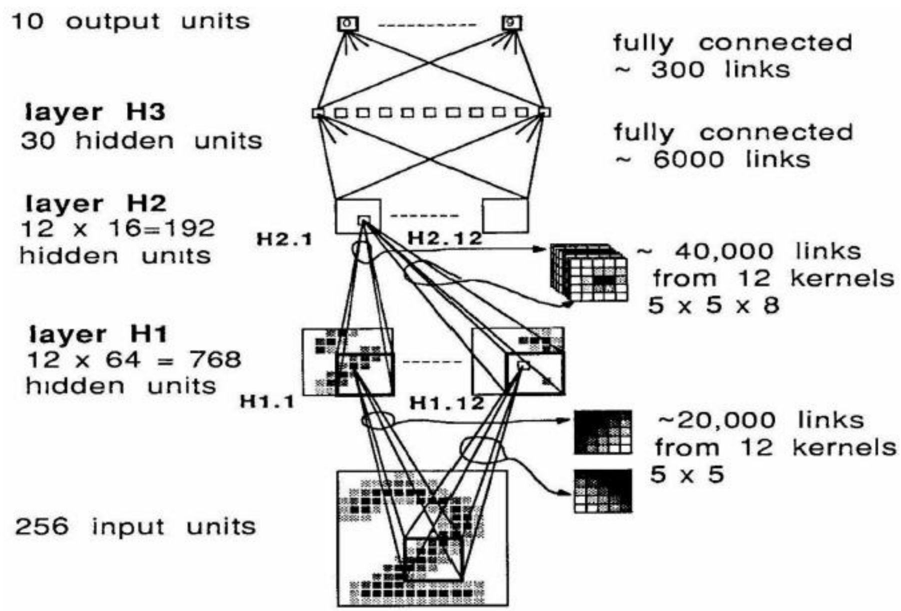
1	

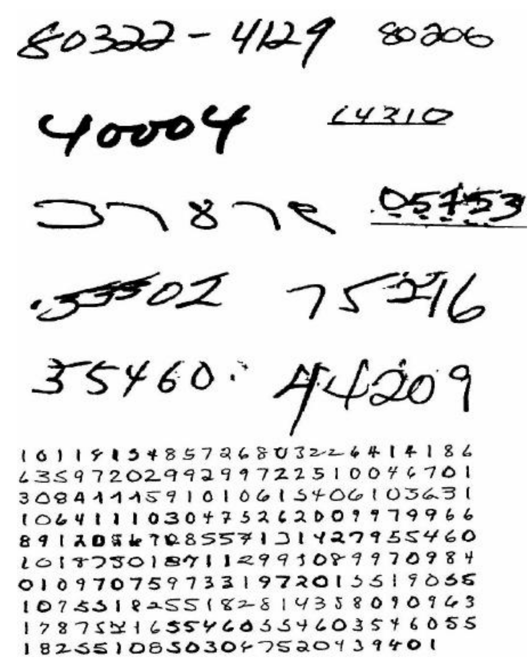
Pooled feature

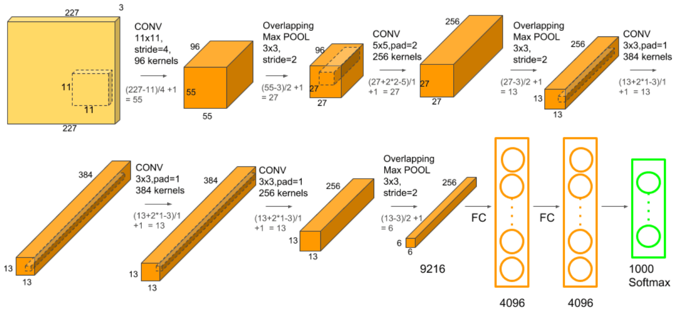
Fewer parameters to characterize the image

First CNNs [LeCun et al. 89]

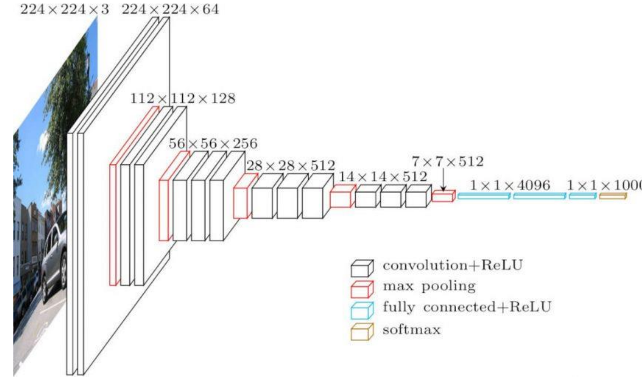
- Trained with Backpropagation
- USPS Zipcode digits: 7300 training, 2000 test
- Convolution with stride. No separate pooling



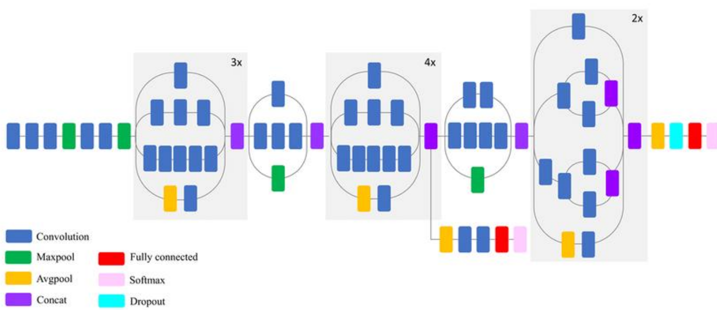




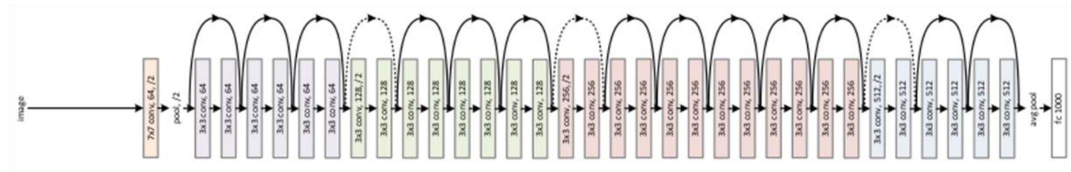
AlexNet (NIPS 2012)



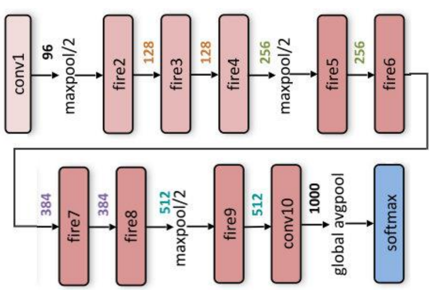
VGG-16 (ICLR 2015)



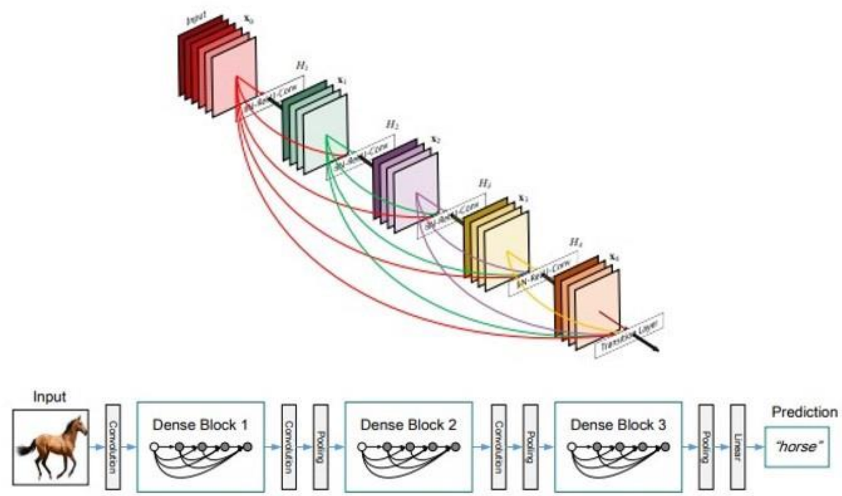
Inception v3 (CVPR 2016)



ResNet50 (CVPR 2015)



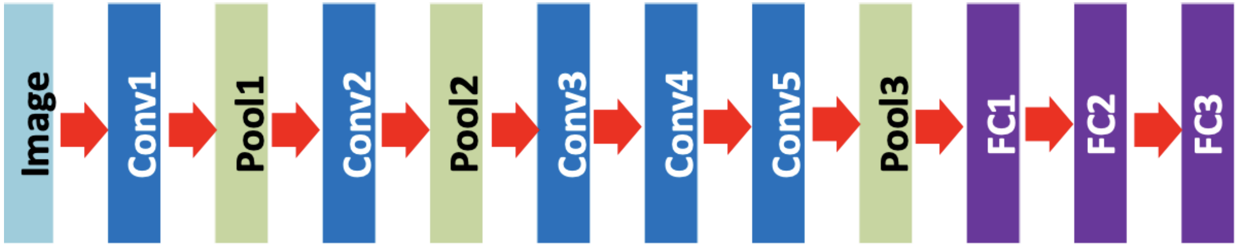
SqeezeNet (2016)



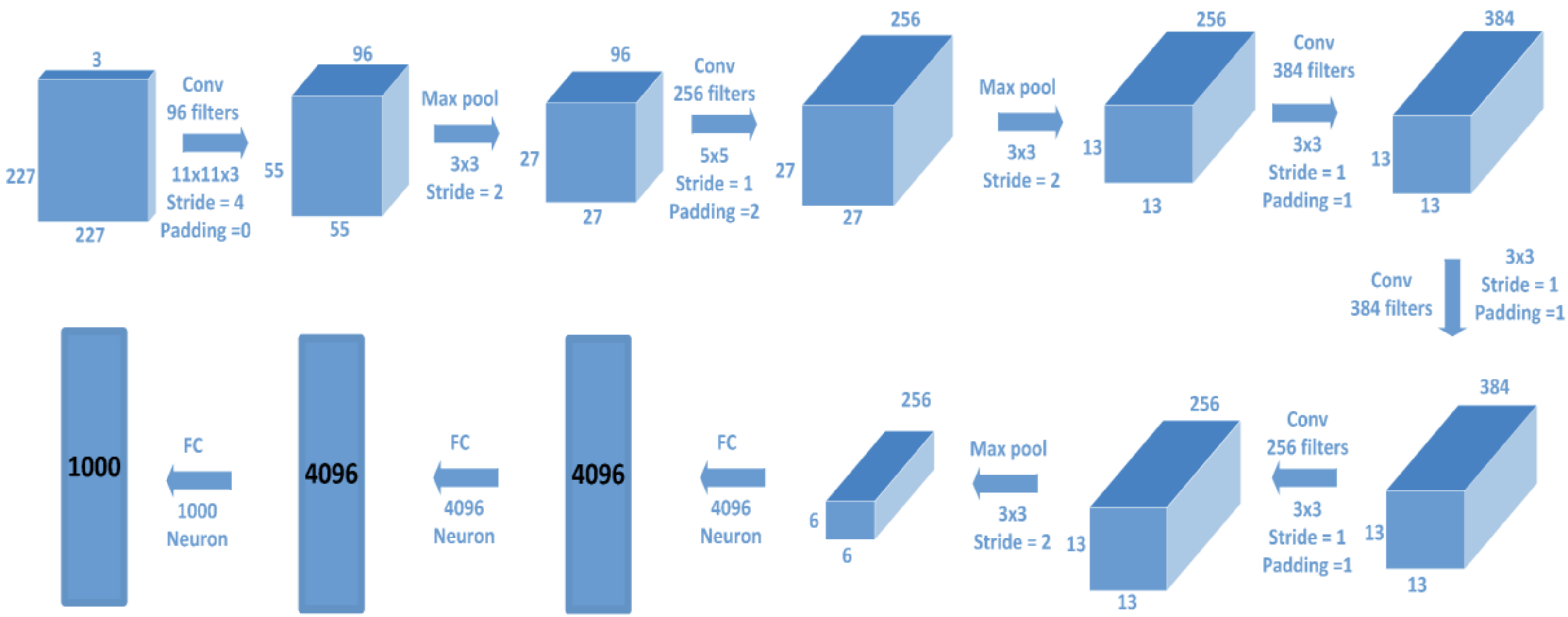
Densenet121 (CVPR 2017)

AlexNet (2012)

- AlexNet achieves on ILSVRC 2012 competition 15.3% Top-5 error rate compare to 26.2% achieved by the second best entry.
- AlexNet has 8 layers without counting pooling layers.
- AlexNet trained on two GTX 580 GPUs for five to six days



AlexNet (2012)



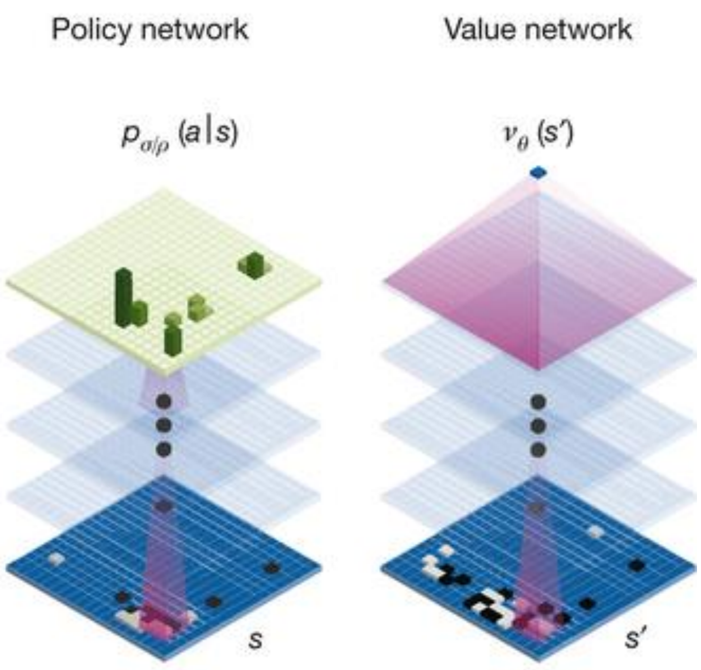
Total (label and softmax not included)
Memory: 2.24 million
Weights: 62.37 million

(Figure from Dr. Mohamed Loey)

AlexNet (2012)

- ReLU
- Norm layers
- Data augmentation
- Dropout 0.5
- Batch size is 128
- SGD Momentum 0.9
- Learning rate 1e-2

Deep CNN in AlphaGO



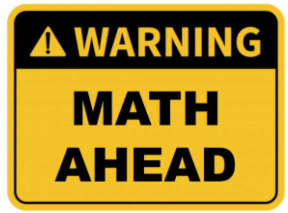
(Silver et al, 2016)

Policy network:

- Input: 19x19, 48 input channels
- Layer 1: 5x5 kernel, 192 filters
- Layer 2 to 12: 3x3 kernel, 192 filters
- Layer 13: 1x1 kernel, 1 filter

Value network has similar architecture to policy network

How to backpropagate with convolution?

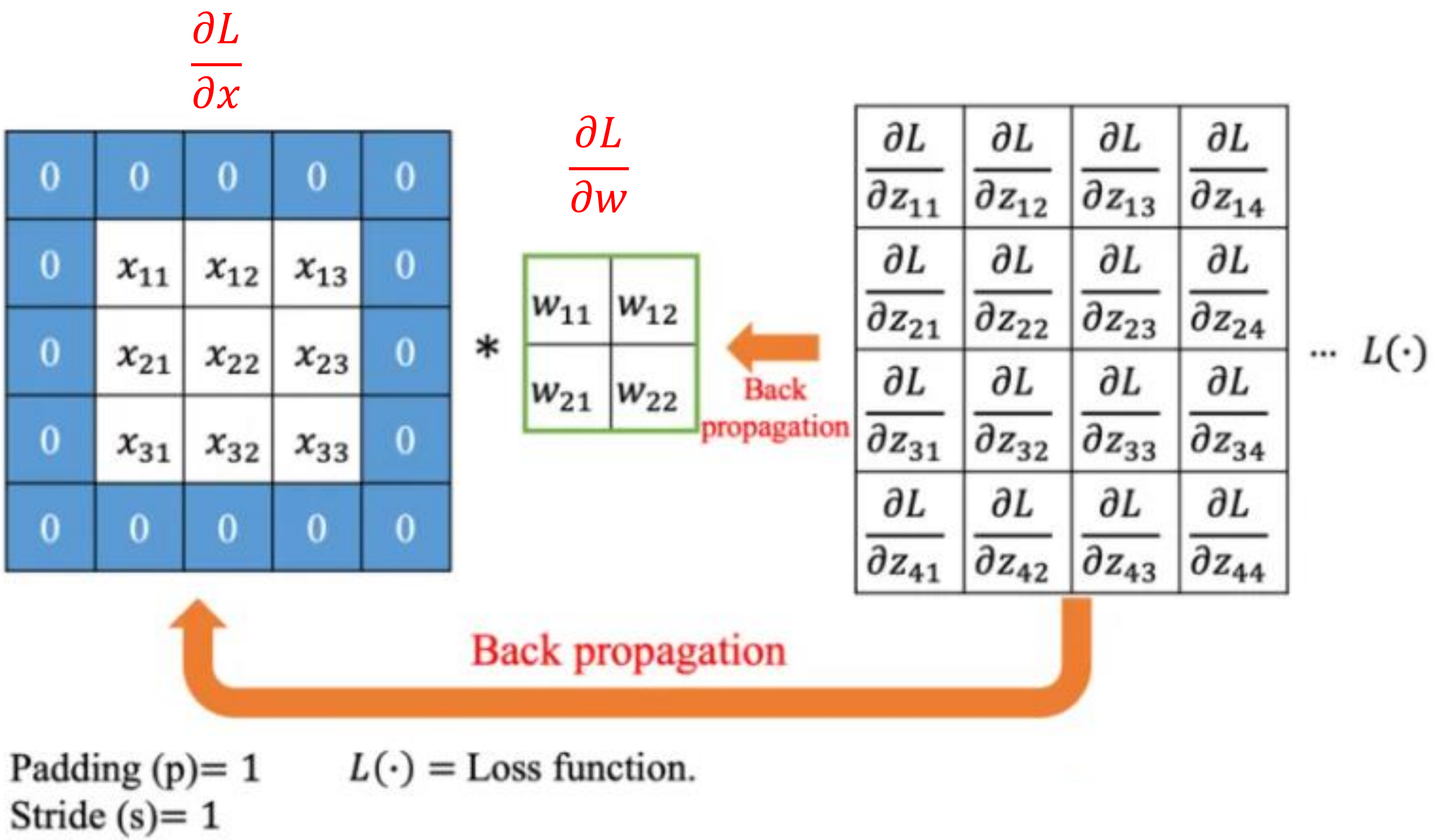


$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} * \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ z_{31} & z_{32} & z_{33} & z_{34} \\ z_{41} & z_{42} & z_{43} & z_{44} \end{bmatrix} \cdots \; L(\cdot)$$

Padding (p)= 1 $\quad$ $L(\cdot) =$ Loss function.

Stride (s)= 1

Ref: https://www.brilliantcode.net/1670/convolutional-neural-networks-4-backpropagation-in-kernels-of-cnns/?cli_action=1604504837.339



Ref: https://www.brilliantcode.net/1670/convolutional-neural-networks-4-backpropagation-in-kernels-of-cnns/?cli_action=1604504837.339

$$z_{11} = 0w_{11} + 0w_{12} + 0w_{21} + x_{11}w_{22}$$
$$z_{12} = 0w_{11} + 0w_{12} + x_{11}w_{21} + x_{12}w_{22}$$
$$z_{13} = 0w_{11} + 0w_{12} + x_{12}w_{21} + x_{13}w_{22}$$
$$z_{14} = 0w_{11} + 0w_{12} + x_{13}w_{21} + 0w_{22}$$

$$z_{21} = 0w_{11} + x_{11}w_{12} + 0w_{21} + x_{21}w_{22}$$
$$z_{22} = x_{11}w_{11} + x_{12}w_{12} + x_{21}w_{21} + x_{22}w_{22}$$
$$z_{23} = x_{12}w_{11} + x_{13}w_{12} + x_{22}w_{21} + x_{23}w_{22}$$
$$z_{24} = x_{13}w_{11} + 0w_{12} + x_{23}w_{21} + 0w_{22}$$

$$z_{31} = 0w_{11} + x_{21}w_{12} + 0w_{21} + x_{31}w_{22}$$
$$z_{32} = x_{21}w_{11} + x_{22}w_{12} + x_{31}w_{21} + x_{32}w_{22}$$
$$z_{33} = x_{22}w_{11} + x_{23}w_{12} + x_{32}w_{21} + x_{33}w_{22}$$
$$z_{34} = x_{23}w_{11} + 0w_{12} + x_{33}w_{21} + 0w_{22}$$

$$z_{41} = 0w_{11} + x_{31}w_{12} + 0w_{21} + 0w_{22}$$
$$z_{42} = x_{31}w_{11} + x_{32}w_{12} + 0w_{21} + 0w_{22}$$
$$z_{43} = x_{32}w_{11} + x_{33}w_{12} + 0w_{21} + 0w_{22}$$
$$z_{44} = x_{33}w_{11} + 0w_{12} + 0w_{21} + 0w_{22}$$

Ref: https://www.brilliantcode.net/1670/convolutional-neural-networks-4-backpropagation-in-kernels-of-cnns/?cli_action=1604504837.339

$$
\begin{aligned}
z_{11} &= 0w_{11} + 0w_{12} + 0w_{21} + x_{11}w_{22} \\
z_{12} &= 0w_{11} + 0w_{12} + x_{11}w_{21} + x_{12}w_{22} \\
z_{13} &= 0w_{11} + 0w_{12} + x_{12}w_{21} + x_{13}w_{22} \\
z_{14} &= 0w_{11} + 0w_{12} + x_{13}w_{21} + 0w_{22}
\end{aligned}
$$

$$
\begin{aligned}
z_{21} &= 0w_{11} + x_{11}w_{12} + 0w_{21} + x_{21}w_{22} \\
z_{22} &= \boxed{x_{11}w_{11}} + x_{12}w_{12} + x_{21}w_{21} + x_{22}w_{22} \\
z_{23} &= \boxed{x_{12}w_{11}} + x_{13}w_{12} + x_{22}w_{21} + x_{23}w_{22} \\
z_{24} &= \boxed{x_{13}w_{11}} + 0w_{12} + x_{23}w_{21} + 0w_{22}
\end{aligned}
$$

$\frac{\partial L}{\partial w_{11}}$?

$$
\begin{aligned}
z_{31} &= 0w_{11} + x_{21}w_{12} + 0w_{21} + x_{31}w_{22} \\
z_{32} &= \boxed{x_{21}w_{11}} + x_{22}w_{12} + x_{31}w_{21} + x_{32}w_{22} \\
z_{33} &= \boxed{x_{22}w_{11}} + x_{23}w_{12} + x_{32}w_{21} + x_{33}w_{22} \\
z_{34} &= \boxed{x_{23}w_{11}} + 0w_{12} + x_{33}w_{21} + 0w_{22}
\end{aligned}
$$

$$
\begin{aligned}
z_{41} &= 0w_{11} + x_{31}w_{12} + 0w_{21} + 0w_{22} \\
z_{42} &= \boxed{x_{31}w_{11}} + x_{32}w_{12} + 0w_{21} + 0w_{22} \\
z_{43} &= \boxed{x_{32}w_{11}} + x_{33}w_{12} + 0w_{21} + 0w_{22} \\
z_{44} &= \boxed{x_{33}w_{11}} + 0w_{12} + 0w_{21} + 0w_{22}
\end{aligned}
$$

Ref: https://www.brilliantcode.net/1670/convolutional-neural-networks-4-backpropagation-in-kernels-of-cnns/?cli_action=1604504837.339

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial w}$$

$$\begin{aligned}
\frac{\partial L}{\partial w_{11}} &= \frac{\partial L}{\partial z_{22}}\boxed{\frac{\partial z_{22}}{\partial w_{11}}} + \frac{\partial L}{\partial z_{23}}\frac{\partial z_{23}}{\partial w_{11}} + \frac{\partial L}{\partial z_{24}}\frac{\partial z_{24}}{\partial w_{11}} \\
&+ \frac{\partial L}{\partial z_{32}}\frac{\partial z_{32}}{\partial w_{11}} + \frac{\partial L}{\partial z_{33}}\frac{\partial z_{33}}{\partial w_{11}} + \frac{\partial L}{\partial z_{34}}\frac{\partial z_{34}}{\partial w_{11}} \\
&+ \frac{\partial L}{\partial z_{42}}\frac{\partial z_{42}}{\partial w_{11}} + \frac{\partial L}{\partial z_{43}}\frac{\partial z_{43}}{\partial w_{11}} + \frac{\partial L}{\partial z_{44}}\frac{\partial z_{44}}{\partial w_{11}} \\
&= \frac{\partial L}{\partial z_{22}}\boxed{x_{11}} + \frac{\partial L}{\partial z_{23}}x_{12} + \frac{\partial L}{\partial z_{24}}x_{13} \\
&+ \frac{\partial L}{\partial z_{32}}x_{21} + \frac{\partial L}{\partial z_{33}}x_{22} + \frac{\partial L}{\partial z_{34}}x_{23} \\
&+ \frac{\partial L}{\partial z_{42}}x_{31} + \frac{\partial L}{\partial z_{43}}x_{32} + \frac{\partial L}{\partial z_{44}}x_{33}
\end{aligned}$$

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial z_{22}}x_{11} + \frac{\partial L}{\partial z_{23}}x_{12} + \frac{\partial L}{\partial z_{24}}x_{13}$$
$$+\frac{\partial L}{\partial z_{32}}x_{12} + \frac{\partial L}{\partial z_{33}}x_{22} + \frac{\partial L}{\partial z_{34}}x_{23}$$
$$+\frac{\partial L}{\partial z_{42}}x_{31} + \frac{\partial L}{\partial z_{43}}x_{32} + \frac{\partial L}{\partial z_{44}}x_{33}$$

$$\frac{\partial L}{\partial w_{12}} = \frac{\partial L}{\partial z_{21}}x_{11} + \frac{\partial L}{\partial z_{22}}x_{12} + \frac{\partial L}{\partial z_{23}}x_{13}+$$
$$+\frac{\partial L}{\partial z_{31}}x_{21} + \frac{\partial L}{\partial z_{32}}x_{22} + \frac{\partial L}{\partial z_{33}}x_{23}$$
$$+\frac{\partial L}{\partial z_{41}}x_{31} + \frac{\partial L}{\partial z_{42}}x_{32} + \frac{\partial L}{\partial z_{43}}x_{33}$$

$$\frac{\partial L}{\partial w_{21}} = \frac{\partial L}{\partial z_{12}}x_{11} + \frac{\partial L}{\partial z_{13}}x_{12} + \frac{\partial L}{\partial z_{14}}x_{13}$$
$$+\frac{\partial L}{\partial z_{22}}x_{21} + \frac{\partial L}{\partial z_{23}}x_{22} + \frac{\partial L}{\partial z_{24}}x_{23}$$
$$+\frac{\partial L}{\partial z_{32}}x_{31} + \frac{\partial L}{\partial z_{33}}x_{32} + \frac{\partial L}{\partial z_{34}}x_{33}$$

$$\frac{\partial L}{\partial w_{22}} = \frac{\partial L}{\partial z_{11}}x_{11} + \frac{\partial L}{\partial z_{12}}x_{12} + \frac{\partial L}{\partial z_{13}}x_{13}$$
$$+\frac{\partial L}{\partial z_{21}}x_{21} + \frac{\partial L}{\partial z_{22}}x_{22} + \frac{\partial L}{\partial z_{23}}x_{23}$$
$$+\frac{\partial L}{\partial z_{31}}x_{31} + \frac{\partial L}{\partial z_{32}}x_{32} + \frac{\partial L}{\partial z_{33}}x_{33}$$

$$z_{11} = 0w_{11} + 0w_{12} + 0w_{21} + \boxed{x_{11}}w_{22}$$
$$z_{12} = 0w_{11} + 0w_{12} + \boxed{x_{11}}w_{21} + x_{12}w_{22}$$
$$z_{13} = 0w_{11} + 0w_{12} + x_{12}w_{21} + x_{13}w_{22}$$
$$z_{14} = 0w_{11} + 0w_{12} + x_{13}w_{21} + 0w_{22}$$

$$z_{21} = 0w_{11} + \boxed{x_{11}}w_{12} + 0w_{21} + x_{21}w_{22}$$
$$z_{22} = \boxed{x_{11}}w_{11} + x_{12}w_{12} + x_{21}w_{21} + x_{22}w_{22}$$
$$z_{23} = x_{12}w_{11} + x_{13}w_{12} + x_{22}w_{21} + x_{23}w_{22}$$
$$z_{24} = x_{13}w_{11} + 0w_{12} + x_{23}w_{21} + 0w_{22}$$

$$\frac{\partial L}{\partial x_{11}}$$?

$$z_{31} = 0w_{11} + x_{21}w_{12} + 0w_{21} + x_{31}w_{22}$$
$$z_{32} = x_{21}w_{11} + x_{22}w_{12} + x_{31}w_{21} + x_{32}w_{22}$$
$$z_{33} = x_{22}w_{11} + x_{23}w_{12} + x_{32}w_{21} + x_{33}w_{22}$$
$$z_{34} = x_{23}w_{11} + 0w_{12} + x_{33}w_{21} + 0w_{22}$$

$$z_{41} = 0w_{11} + x_{31}w_{12} + 0w_{21} + 0w_{22}$$
$$z_{42} = x_{31}w_{11} + x_{32}w_{12} + 0w_{21} + 0w_{22}$$
$$z_{43} = x_{32}w_{11} + x_{33}w_{12} + 0w_{21} + 0w_{22}$$
$$z_{44} = x_{33}w_{11} + 0w_{12} + 0w_{21} + 0w_{22}$$

Ref: https://www.brilliantcode.net/1670/convolutional-neural-networks-4-backpropagation-in-kernels-of-cnns/?cli_action=1604504837.339

$$\boxed{\frac{\partial L}{\partial x} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial x}}$$

$$\frac{\partial L}{\partial x_{11}} = \frac{\partial L}{\partial z_{11}}\frac{\partial z_{11}}{\partial x_{11}} + \frac{\partial L}{\partial z_{12}}\frac{\partial z_{12}}{\partial x_{11}} + \frac{\partial L}{\partial z_{21}}\frac{\partial z_{21}}{\partial x_{11}} + \frac{\partial L}{\partial z_{22}}\frac{\partial z_{22}}{\partial x_{11}}$$

$$= \frac{\partial L}{\partial z_{11}} w_{22} + \frac{\partial L}{\partial z_{12}} w_{21} + \frac{\partial L}{\partial z_{21}} w_{12} + \frac{\partial L}{\partial z_{22}} w_{11}$$

$$\frac{\partial L}{\partial x_{22}} = \frac{\partial L}{\partial z_{22}}\frac{\partial z_{22}}{\partial x_{22}} + \frac{\partial L}{\partial z_{23}}\frac{\partial z_{23}}{\partial x_{22}} + \frac{\partial L}{\partial z_{32}}\frac{\partial z_{32}}{\partial x_{22}} + \frac{\partial L}{\partial z_{33}}\frac{\partial z_{33}}{\partial x_{22}}$$

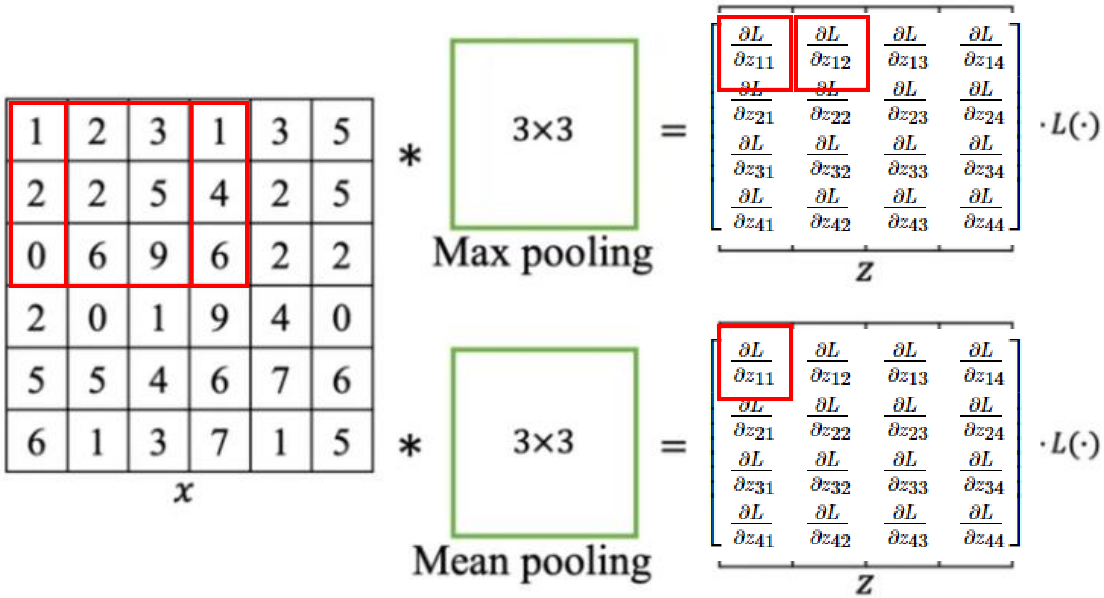
$$= \frac{\partial L}{\partial z_{22}} w_{22} + \frac{\partial L}{\partial z_{23}} w_{21} + \frac{\partial L}{\partial z_{32}} w_{12} + \frac{\partial L}{\partial z_{33}} w_{11}$$

$$\vdots$$

$$
\frac{\partial L}{\partial w} =
\begin{array}{|l|l|}
\hline
\begin{aligned} &= \frac{\partial L}{\partial z_{22}} x_{11} + \frac{\partial L}{\partial z_{23}} x_{12} + \frac{\partial L}{\partial z_{24}} x_{13} \\ &+ \frac{\partial L}{\partial z_{32}} x_{12} + \frac{\partial L}{\partial z_{33}} x_{22} + \frac{\partial L}{\partial z_{34}} x_{23} \\ &+ \frac{\partial L}{\partial z_{42}} x_{31} + \frac{\partial L}{\partial z_{43}} x_{32} + \frac{\partial L}{\partial z_{44}} x_{33} \end{aligned} &
\begin{aligned} &= \frac{\partial L}{\partial z_{21}} x_{11} + \frac{\partial L}{\partial z_{22}} x_{12} + \frac{\partial L}{\partial z_{23}} x_{13} + \\ &+ \frac{\partial L}{\partial z_{31}} x_{21} + \frac{\partial L}{\partial z_{32}} x_{22} + \frac{\partial L}{\partial z_{33}} x_{23} \\ &+ \frac{\partial L}{\partial z_{41}} x_{31} + \frac{\partial L}{\partial z_{42}} x_{32} + \frac{\partial L}{\partial z_{43}} x_{33} \end{aligned} \\
\hline
\begin{aligned} &= \frac{\partial L}{\partial z_{12}} x_{11} + \frac{\partial L}{\partial z_{13}} x_{12} + \frac{\partial L}{\partial z_{14}} x_{13} \\ &+ \frac{\partial L}{\partial z_{22}} x_{21} + \frac{\partial L}{\partial z_{23}} x_{22} + \frac{\partial L}{\partial z_{24}} x_{23} \\ &+ \frac{\partial L}{\partial z_{32}} x_{31} + \frac{\partial L}{\partial z_{33}} x_{32} + \frac{\partial L}{\partial z_{34}} x_{33} \end{aligned} &
\begin{aligned} &= \frac{\partial L}{\partial z_{11}} x_{11} + \frac{\partial L}{\partial z_{12}} x_{12} + \frac{\partial L}{\partial z_{13}} x_{13} \\ &+ \frac{\partial L}{\partial z_{21}} x_{21} + \frac{\partial L}{\partial z_{22}} x_{22} + \frac{\partial L}{\partial z_{23}} x_{23} \\ &+ \frac{\partial L}{\partial z_{31}} x_{31} + \frac{\partial L}{\partial z_{32}} x_{32} + \frac{\partial L}{\partial z_{33}} x_{33} \end{aligned} \\
\hline
\end{array}
$$

$$
\frac{\partial L}{\partial x} =
\begin{array}{|c|c|c|}
\hline
\begin{gathered} \frac{\partial L}{\partial z_{11}} w_{22} + \frac{\partial L}{\partial z_{12}} w_{21} + \\ \frac{\partial L}{\partial z_{21}} w_{12} + \frac{\partial L}{\partial z_{22}} w_{11} \end{gathered} &
\begin{gathered} \frac{\partial L}{\partial z_{12}} w_{22} + \frac{\partial L}{\partial z_{13}} w_{21} + \\ \frac{\partial L}{\partial z_{22}} w_{12} + \frac{\partial L}{\partial z_{23}} w_{11} \end{gathered} &
\begin{gathered} \frac{\partial L}{\partial z_{13}} w_{22} + \frac{\partial L}{\partial z_{14}} w_{21} + \\ \frac{\partial L}{\partial z_{23}} w_{12} + \frac{\partial L}{\partial z_{24}} w_{11} \end{gathered} \\
\hline
\begin{gathered} \frac{\partial L}{\partial z_{21}} w_{22} + \frac{\partial L}{\partial z_{22}} w_{21} + \\ \frac{\partial L}{\partial z_{31}} w_{12} + \frac{\partial L}{\partial z_{32}} w_{11} \end{gathered} &
\begin{gathered} \frac{\partial L}{\partial z_{22}} w_{22} + \frac{\partial L}{\partial z_{23}} w_{21} + \\ \frac{\partial L}{\partial z_{32}} w_{12} + \frac{\partial L}{\partial z_{33}} w_{11} \end{gathered} &
\begin{gathered} \frac{\partial L}{\partial z_{23}} w_{22} + \frac{\partial L}{\partial z_{24}} w_{21} + \\ \frac{\partial L}{\partial z_{33}} w_{12} + \frac{\partial L}{\partial z_{34}} w_{11} \end{gathered} \\
\hline
\begin{gathered} \frac{\partial L}{\partial z_{31}} w_{22} + \frac{\partial L}{\partial z_{32}} w_{21} + \\ \frac{\partial L}{\partial z_{41}} w_{12} + \frac{\partial L}{\partial z_{42}} w_{11} \end{gathered} &
\begin{gathered} \frac{\partial L}{\partial z_{32}} w_{22} + \frac{\partial L}{\partial z_{33}} w_{21} + \\ \frac{\partial L}{\partial z_{42}} w_{12} + \frac{\partial L}{\partial z_{43}} w_{11} \end{gathered} &
\begin{gathered} \frac{\partial L}{\partial z_{33}} w_{22} + \frac{\partial L}{\partial z_{34}} w_{21} + \\ \frac{\partial L}{\partial z_{43}} w_{12} + \frac{\partial L}{\partial z_{44}} w_{11} \end{gathered} \\
\hline
\end{array}
$$

How about Pooling layers?



Pooling kernel size= (3×3), Stride (s)= 1, $L(\cdot)$ = Loss function.

Max:
$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial L}{\partial z_{11}} + \frac{\partial L}{\partial z_{12}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

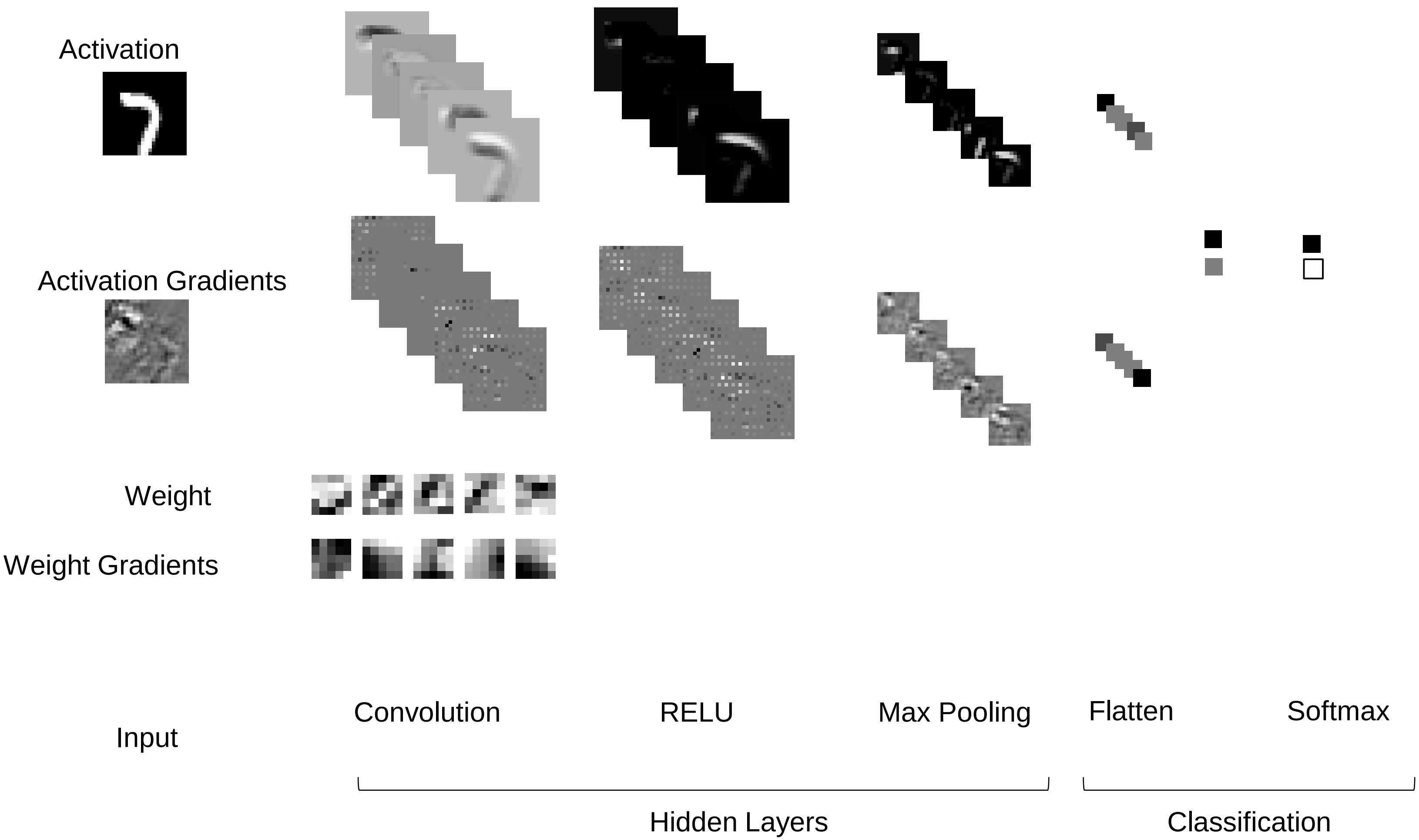
Mean:
$$\begin{bmatrix} (\frac{\partial L}{\partial z_{11}})/9 & (\frac{\partial L}{\partial z_{11}})/9 & (\frac{\partial L}{\partial z_{11}})/9 & 0 & 0 & 0 \\ (\frac{\partial L}{\partial z_{11}})/9 & (\frac{\partial L}{\partial z_{11}})/9 & (\frac{\partial L}{\partial z_{11}})/9 & 0 & 0 & 0 \\ (\frac{\partial L}{\partial z_{11}})/9 & (\frac{\partial L}{\partial z_{11}})/9 & (\frac{\partial L}{\partial z_{11}})/9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

CNNs are good for

- Signals that comes to you in the form of (multidimensional) arrays.
- Signals that have strong local correlations
- Signals where features can appear anywhere
- Signals in which objects are invariant to translations.
- 1D CNNs: sequential signals, text
  - Text, music, audio, speech, time series.
- 2D CNNs: images, time-frequency representations (speech and audio)
  - Object detection, localization, recognition
- 3D CNNs: video, volumetric images, tomography images
  - Video recognition / understanding
  - Biomedical image analysis

Model Visualization

http://cs.stanford.edu/people/karpathy/convnetjs/



Questions?

How to confuse your ConvNets?