

Basic RL.7

Judy Tutorial

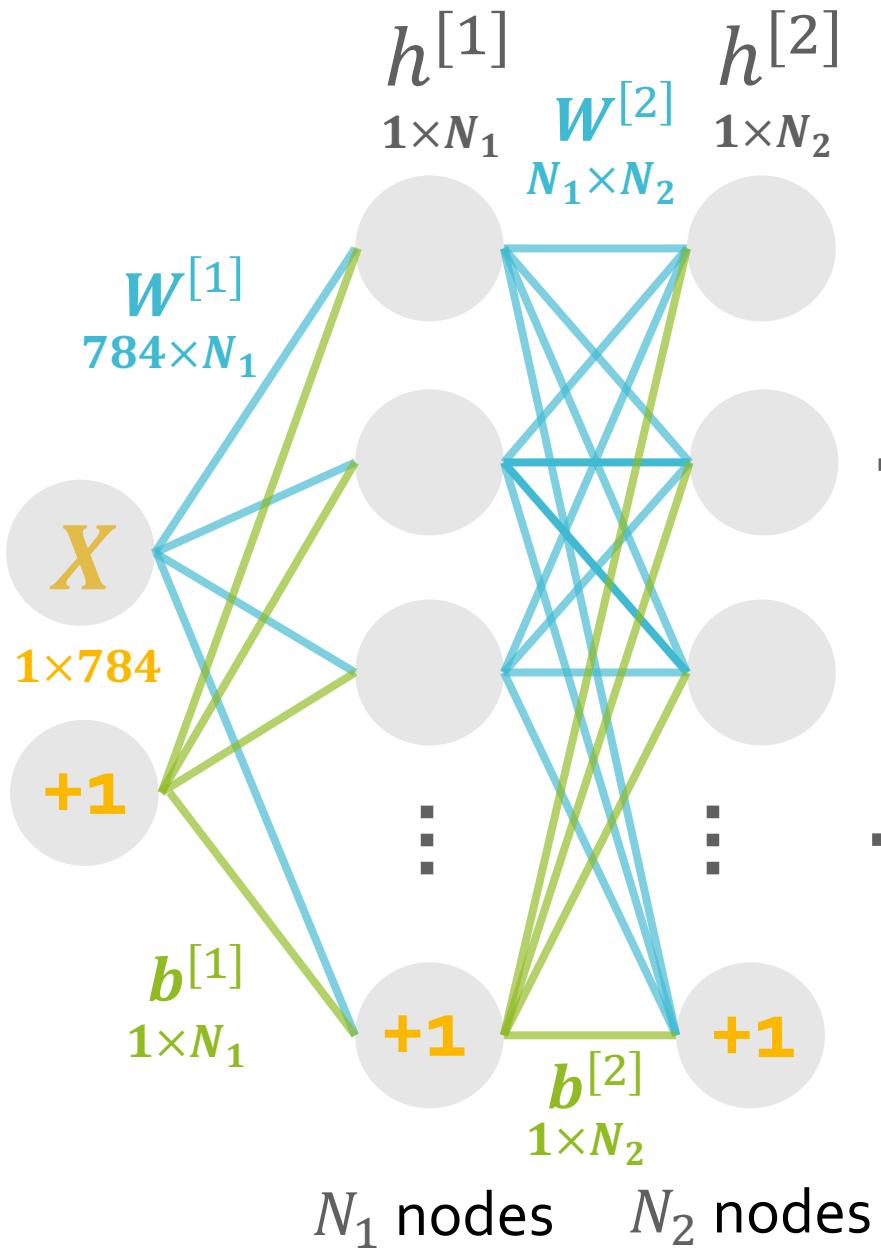
MNIST dataset

[LeCun 1998]

Multiple Labels

$y \in \{0,1,2, \dots, 9\}$

X	- [60000, 28x28 pixels]		$[60000, 1 \text{ int}] - Y$
0	0 0		0
1	1 1		1
2	2 2		2
3	3 3		3
4	4 4		4
5	5 5		5
6	6 6		6
7	7 7		7
8	8 8		8
9	9 9		9

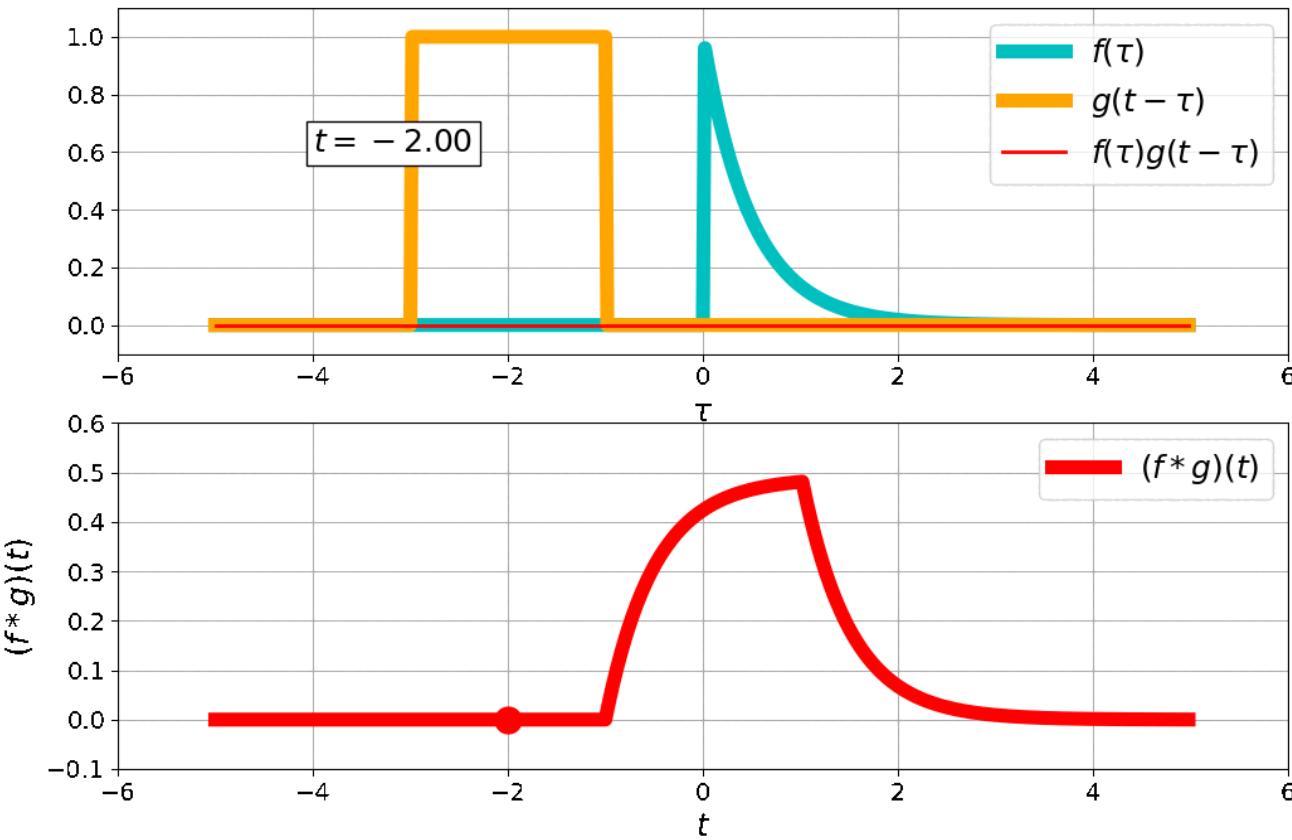


Convolutional Neural Networks

[LeCun 1980s]

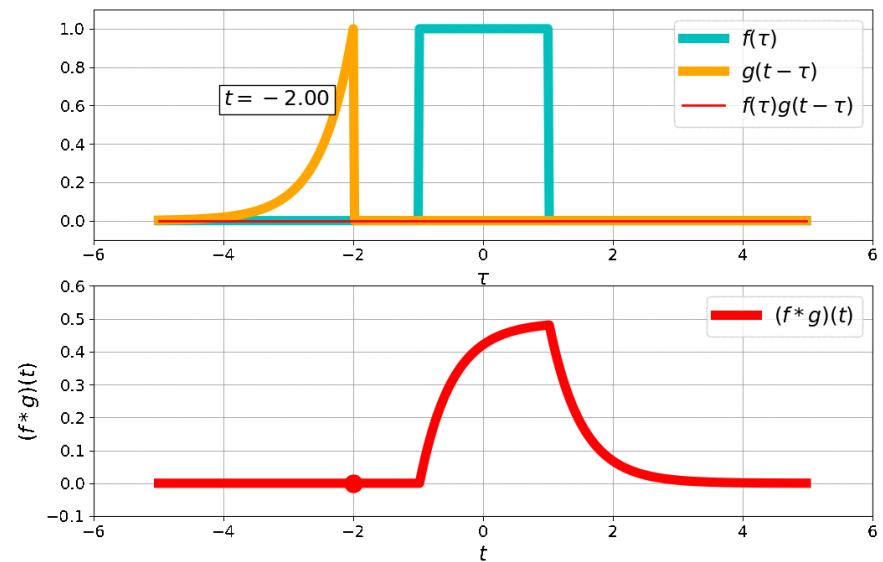
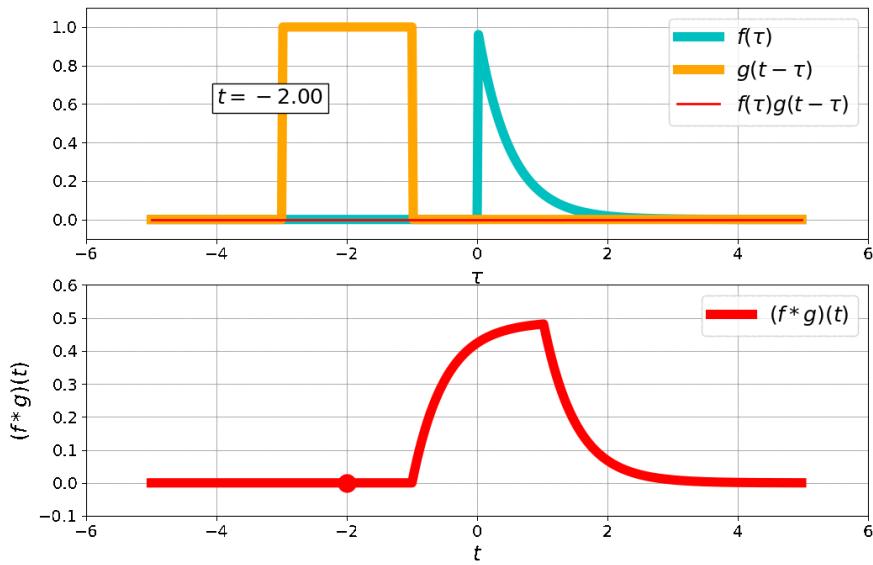
Convolution of two functions

$$(\mathbf{\mathit{f}} * \mathbf{\mathit{g}})(t) \triangleq \int_{-\infty}^{\infty} \mathbf{\mathit{f}}(\tau) \mathbf{\mathit{g}}(t - \tau) d\tau$$



Convolution is Commutative

$$(f * g)(t) = (g * f)(t)$$



In Image Processing [2D]

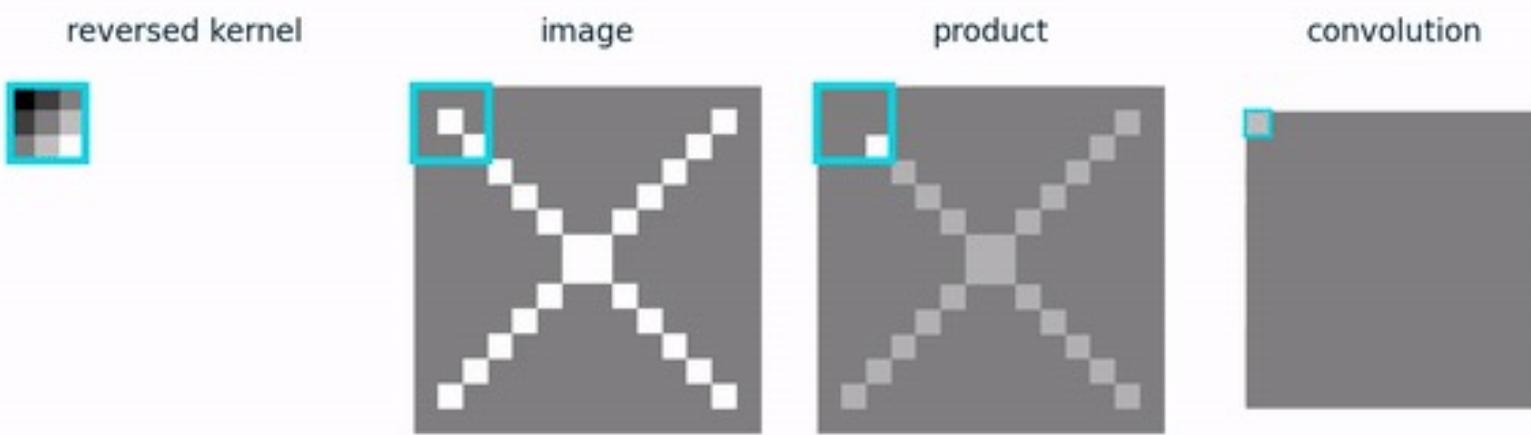
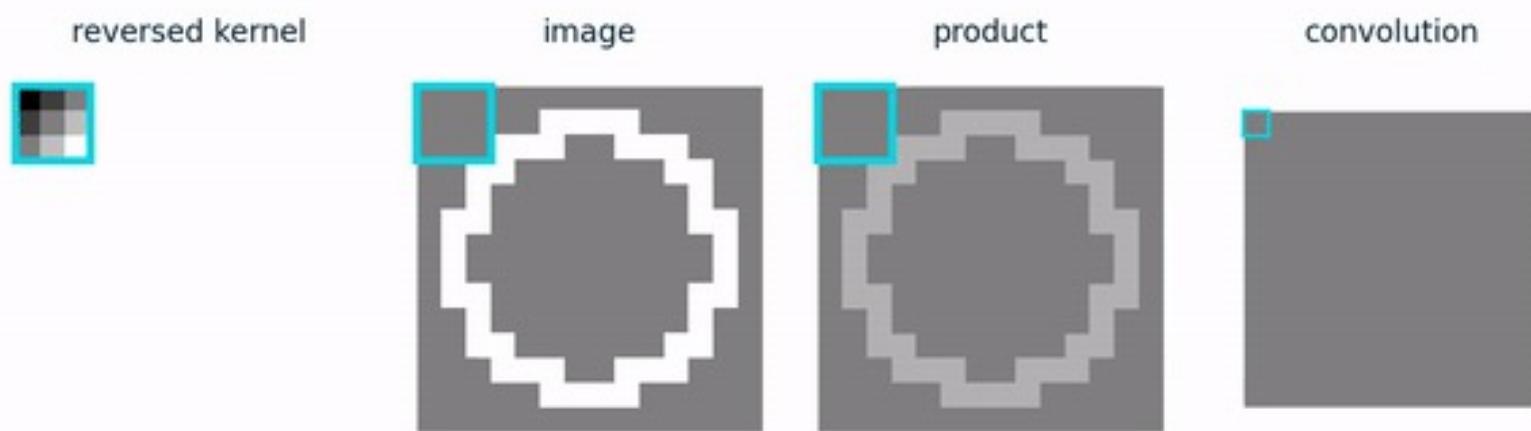
Feature Map

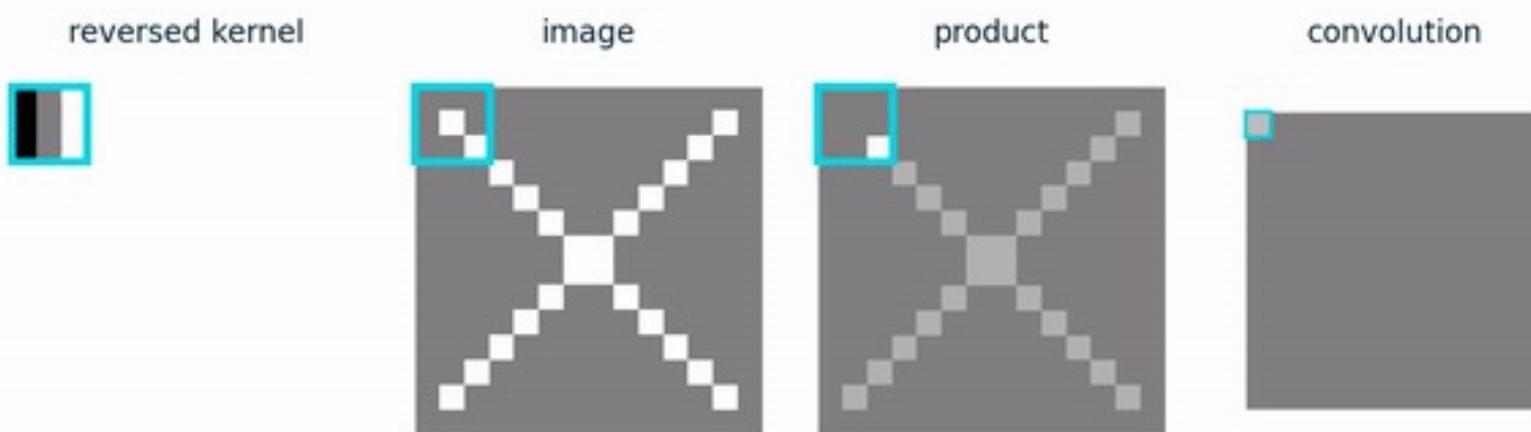
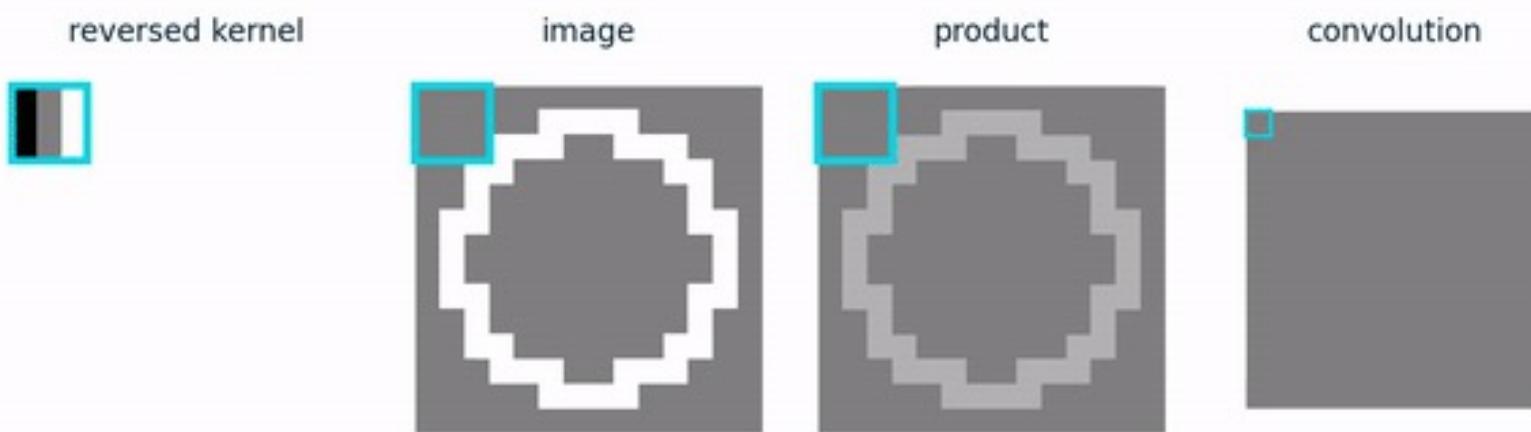
$$FM(i, j)$$

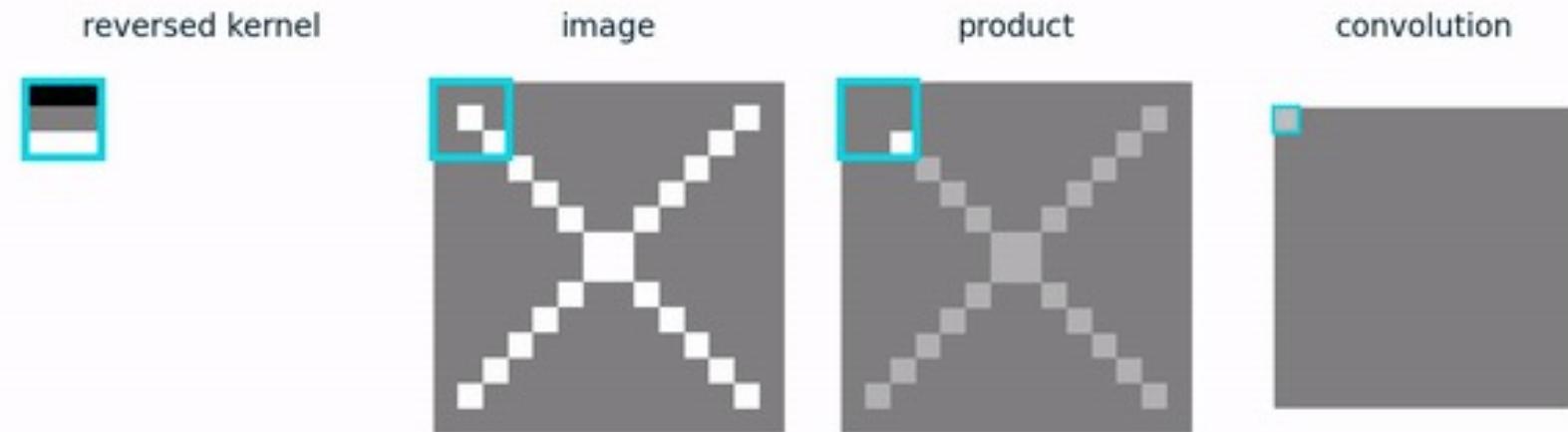
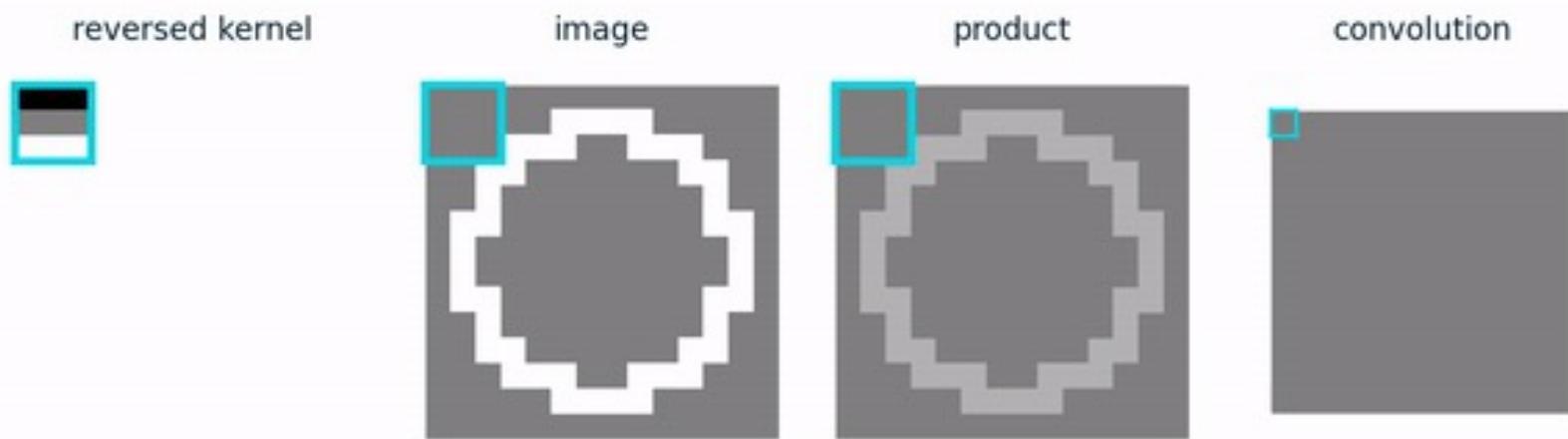
$$= (I * F)(i, j) \sum_m \sum_n I(m, n) F(i - m, i - n)$$

| |

Input Filter/
image kernel

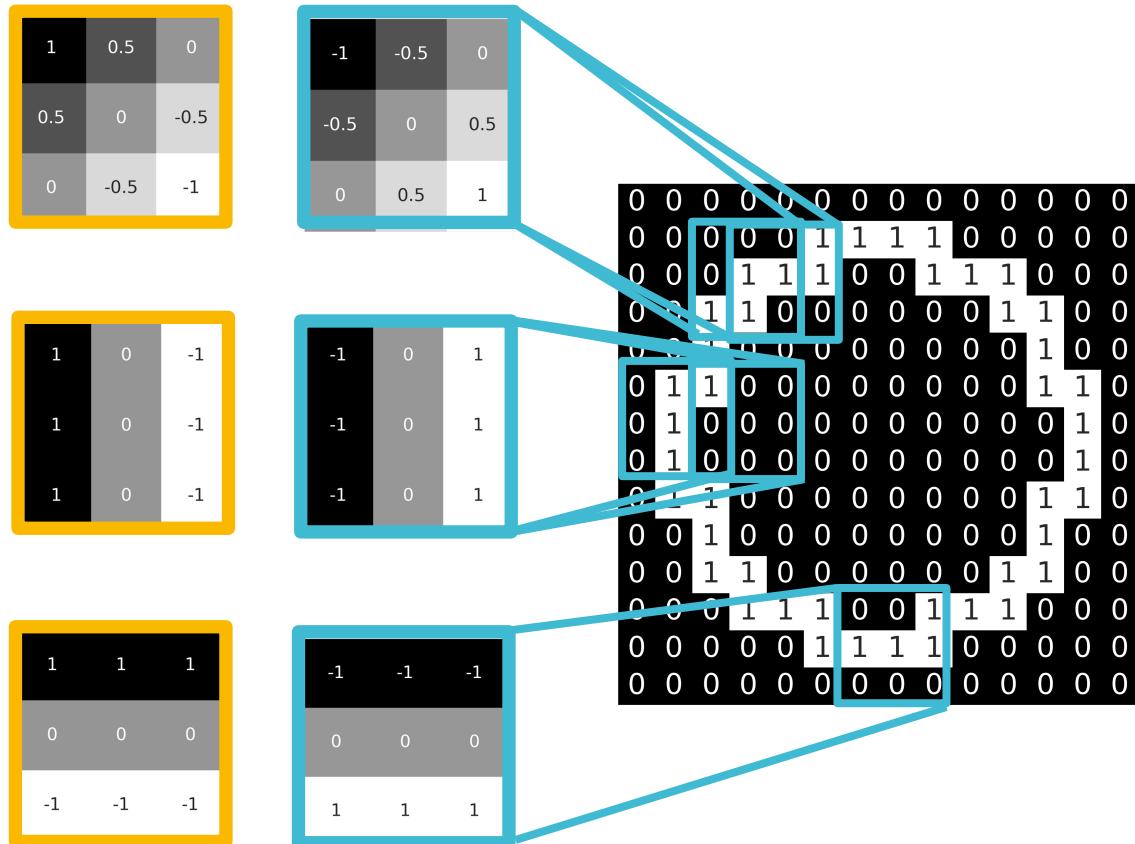






Convolution layer

reverse



F_{size} : $3 \times 3 \times 3$

I_{size} : 14×14×1

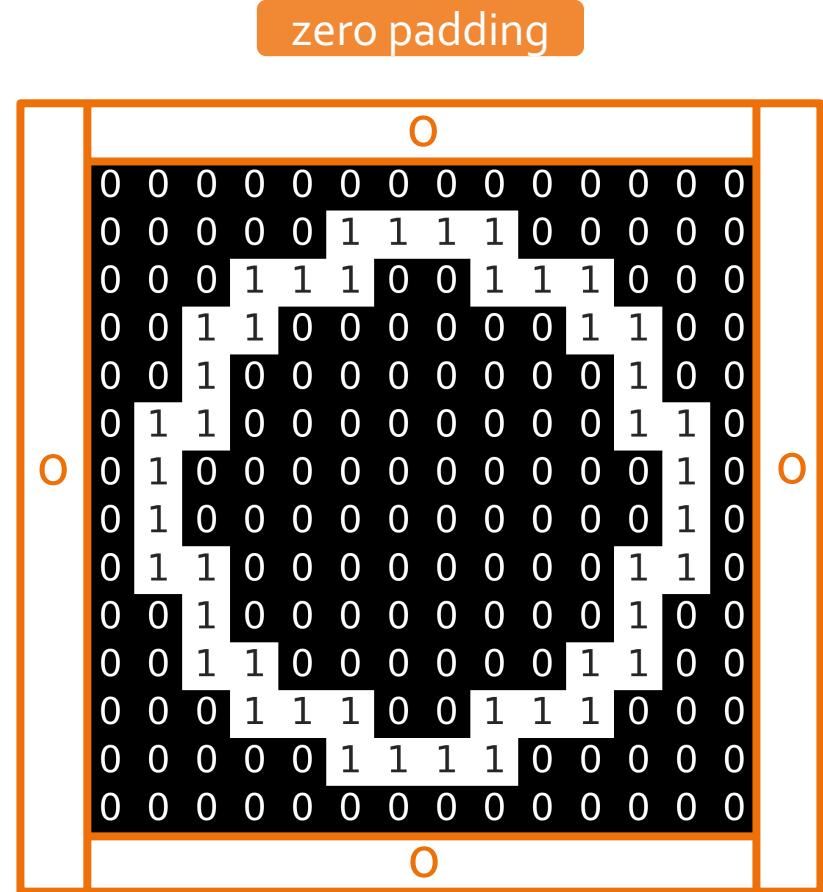
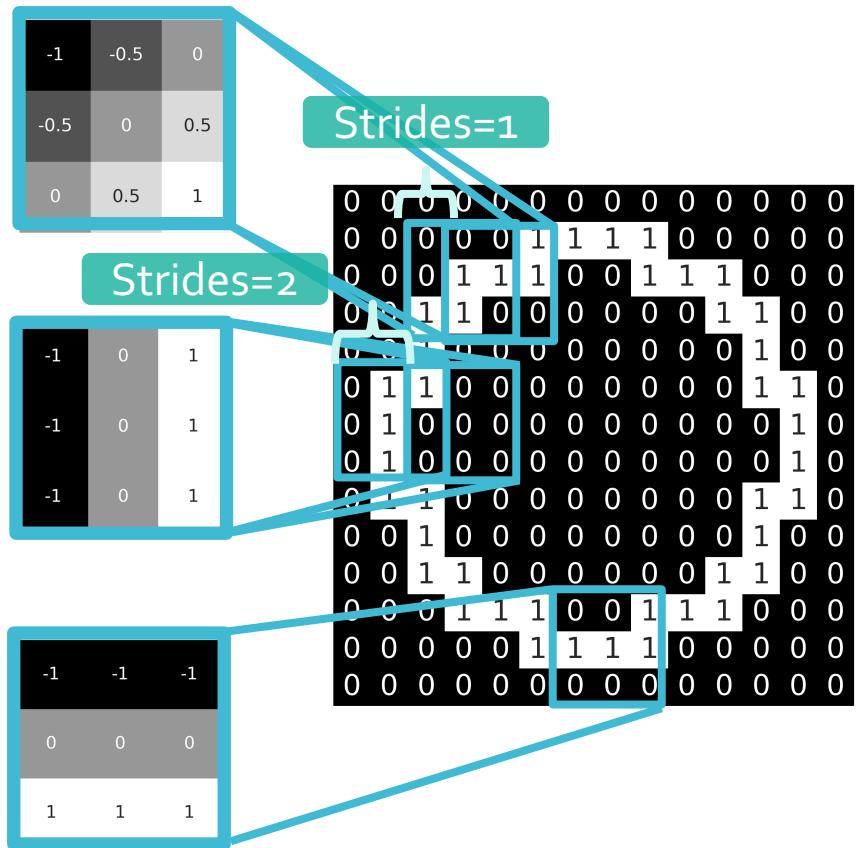
$$FM_{size} = (I_{size} - F_{size} + 2P) / S + 1$$

0	1	1.5	2	1	0	1	1	1	0.5	0	0
1	2	1	0	-1	-2	-1	-1	0	1	0	0
1.5	1	-1	-2	-1.5	-1	0	-0.5	-1	0	-1	-0.5
2	0	-2	-1	0	0	0	0	0	1	0	-1
1	-1	-1.5	0	0	0	0	0	0	0.5	1	-1
0	-2	-1	0	0	0	0	0	0	0	1	1
1	-1	0	0	0	0	0	0	0	1	2	
1	-1	-0.5	0	0	0	0	0	0	1.5	1	-1
1	0	-1	0	0	0	0	0	1	2		-2
0.5	1	0	1	0.5	0	1	1.5	2	1	-1	-1.5
0	0	-1	0	1	1	2	1	0	-1	-2	-1
0	0	-0.5	-1	-1	0	1	-2	-1.5	-1	0	

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

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

Convolution layer



$$FM_{size} = (I_{size} - F_{size} + 2P)/S + 1$$

Convolution layer

0	1	1.5	2	1	0	1	1	1	0.5	0	0
1	2	1	0	-1	-2	-1	-1	0	1	0	0
1.5	1	-1	-2	-1.5	-1	0	-0.5	-1	0	-1	-0.5
2	0	-2	-1	0	0	0	0	0	1	0	-1
1	-1	-1.5	0	0	0	0	0	0	0.5	1	-1
0	-2	-1	0	0	0	0	0	0	0	1	-1
1	-1	0	0	0	0	0	0	0	1	2	0
1	-1	-0.5	0	0	0	0	0	0	1.5	1	-1
1	0	-1	0	0	0	0	0	1	2	0	-2
0.5	1	0	1	0.5	0	1	1.5	2	1	-1	-1.5
0	0	-1	0	1	1	2	1	0	-1	-2	-1
0	0	-0.5	-1	-1	-1	0	-1	-2	-1.5	-1	0



ReLU

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

0	1	1.5	2	1	0	1	1	1	0.5	0
1	2	1	0	-0	-0	-0	-0	0	1	0
1.5	1	-0	-0	-0	-0	0	0	-0	0	-0
2	0	-0	-0	0	0	0	0	0	1	0
1	-0	0	0	0	0	0	0	0.5	1	-0
0	-0	-0	0	0	0	0	0	0	0	1
1	-0	0	0	0	0	0	0	0	1	2
1	-0	-0	0	0	0	0	0	1.5	1	-0
1	0	-0	0	0	0	0	0	1	2	0
0.5	1	0	1	0.5	0	1	1.5	2	1	-0
0	0	-0	-0	-0	-0	0	-0	-0	-0	0
0	0	-0	-0	-0	-0	0	-0	-0	-0	0

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

Pooling layer

0	1	1.5	2	1	0	1	1	1	0.5	0	0	
1	2	1	0	-0	-0	-0	0	0	1	0	0	
1.5	1	-0	-0	-0	-0	0	-0	-0	0	-0	-0	
2	0	-0	-0	0	0	0	0	0	1	0	-0	
1	-0	-0	0	0	0	0	0	0	0.5	1	-0	
0	-0	-0	0	0	0	0	0	0	0	1	-0	
1	-0	0	0	0	0	0	0	0	1	2	0	
1	-0	-0	0	0	0	0	0	0	0	1.5	1	-0
1	0	-0	0	0	0	0	0	0	1	2	0	-0
0.5	1	0	1	0.5	0	1	1.5	2	1	-0	-0	
0	0	-0	0	1	1	2	1	0	-0	-0	-0	
0	0	-0	-0	-0	-0	0	-0	-0	-0	-0	-0	

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

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

[2x2]

[6x6]

2	2	1	1	1	1	0
2	-0	-0	0	1	-0	-0
1	-0	0	0	0.5	1	-0
1	0	0	0	1.5	2	-0
1	1	0.5	1.5	2	0	-0
0	-0	1	2	0	0	-0

2	2	1	1	1	1	0
2	-0	-0	0	1	-0	-0
1	-0	0	0	1.5	2	-0
1	1	1.5	2	0	0	-0
1	1	1.5	2	0	0	-0
1	1	1.5	2	0	0	-0

[2x2]

or

[6x6]

Strides=2

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

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

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

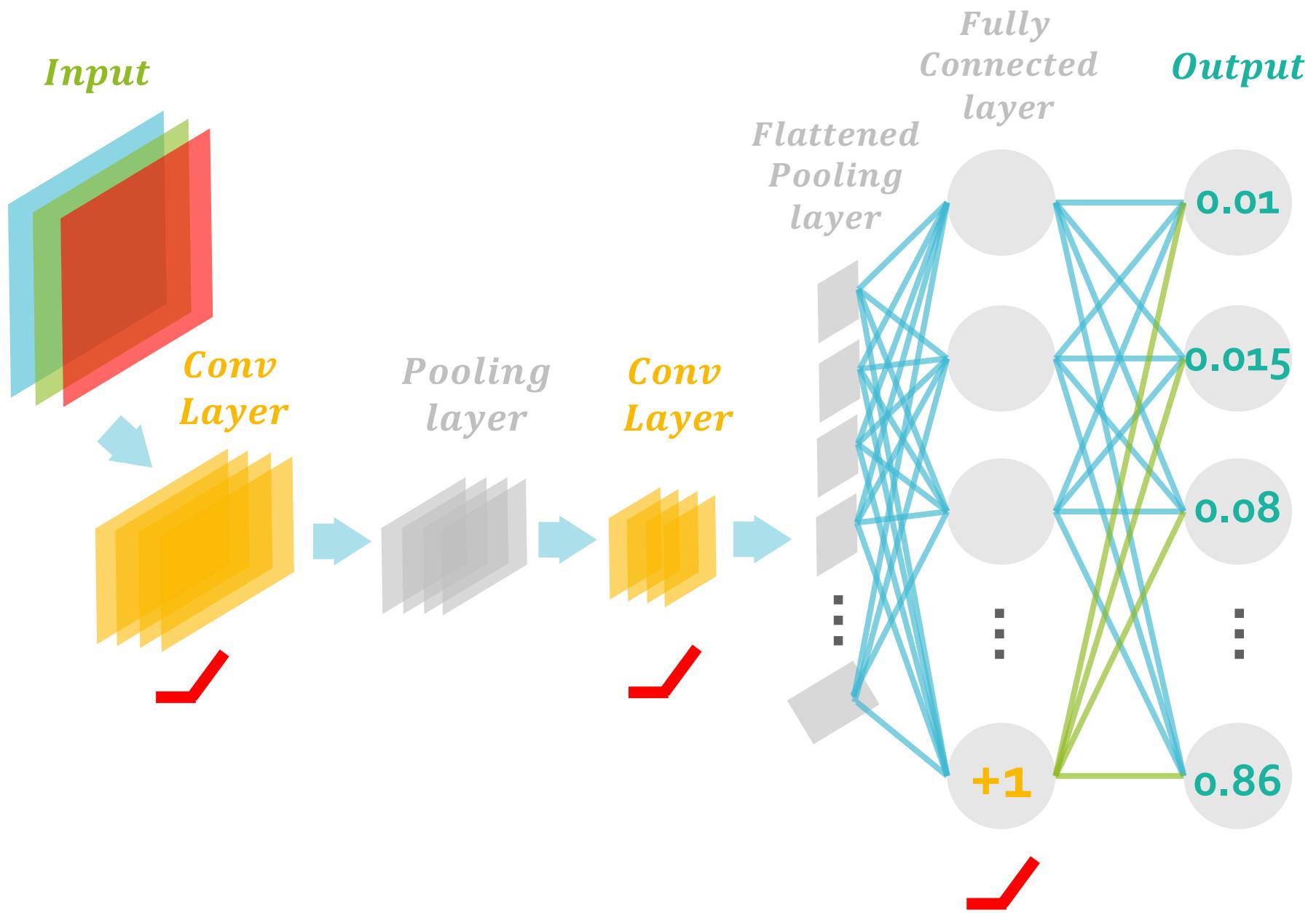
3	3	2	3	3
1	-0	-0	-0	-0
1	1	0	0	1
3	3	2	3	3
3	3	2	3	3

$MP_{size}: 6 \times 6 \times 3$

$MP_{size} = (FM_{size} - MPF_{size})/S + 1$

$4 \times 4 \times 3$

$FM_{size}: 12 \times 12 \times 3$



summary

Convolution layer

- **Sparse interactions:**
 - kernel smaller than the input to lowerize parameter size
 - detect small meaningful features
- **Shared weights:**
 - correlation parts in the same fig
 - Equivariant Representations, i.e. invariant to the change of positions

Pooling layer

- **Downsampling**
 - summarize the responses over a whole neighborhood
 - invariant to small translations of the input

MNIST dataset

[LeCun 1998]

Multiple Labels

$y \in \{0,1,2, \dots, 9\}$

in python keras

```
cnn=Sequential()
cnn.add(Conv2D(filters=16,kernel_size=(3,3),activation='relu',input_shape=(28,28,1)))
cnn.add(MaxPooling2D((2,2)))
cnn.add(Conv2D(filters=16,kernel_size=(3,3),activation='relu'))
cnn.add(MaxPooling2D((2,2)))
cnn.add(Flatten())
cnn.add(Dense(32,activation='relu'))
cnn.add(Dense(10,activation='softmax'))
cnn.summary()

cnn.compile(loss='categorical_crossentropy',
            optimizer='adam',
            metrics=['categorical_accuracy'])

traj=cnn.fit(x_train,y_train,epochs=20,batch_size=32,verbose=1,
              shuffle=True,validation_data=(x_test, y_test))
```

Build
model

} Training

x_train: (60000, 28, 28, 1)

y_train: (60000, 10) - one hot

x_test: (10000, 28, 28, 1)

y_test: (10000, 10) – one hot

Summary

Model: "sequential_17"

	Layer (type)	Output Shape	Param #
<hr/>			
layer.0	conv2d_61 (Conv2D)	(None, 26, 26, 16)	160
layer.1	max_pooling2d_32 (MaxPooling2D)	(None, 13, 13, 16)	0
layer.2	conv2d_62 (Conv2D)	(None, 11, 11, 16)	2320
layer.3	max_pooling2d_33 (MaxPooling2D)	(None, 5, 5, 16)	0
	flatten_16 (Flatten)	(None, 400)	0
	dense_32 (Dense)	(None, 32)	12832
	dense_33 (Dense)	(None, 10)	330
<hr/>			

Total params: 15,642

Trainable params: 15,642

Non-trainable params: 0

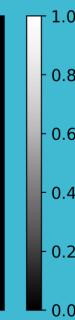
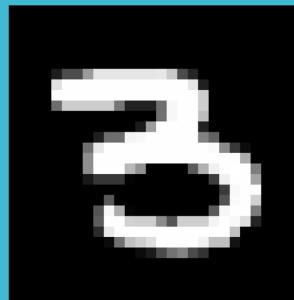
Filter Weights [layer.o]

0.011	0.31	-0.083	-0.72	-0.18	0.19	0.14	0.5	-0.14	0.032	0.0093	0.033
-0.23	0.33	0.035	-0.54	-0.12	0.53	-0.34	0.34	0.078	-0.0029	0.013	-0.048
-0.041	0.14	0.066	0.17	0.3	0.26	-0.32	-0.00066	0.4	-0.15	0.18	-0.079
0.084	0.42	-0.021	0.62	-0.59	-0.55	-0.24	0.23	0.13	-0.6	-0.21	0.4
0.023	0.29	-0.11	-0.56	-1.3	0.42	-0.13	0.52	0.15	0.32	0.41	0.28
-0.055	0.2	0.15	-1.1	0.35	0.65	0.078	0.42	-0.4	0.33	-0.31	-0.62
0.057	0.085	0.35	0.13	0.099	0.11	0.6	0.51	0.4	0.3	0.26	0.12
0.57	0.48	0.17	-0.1	0.019	-0.044	-0.13	-0.062	-0.14	0.43	0.01	-0.26
-0.21	-0.75	-0.57	-0.12	0.28	-0.15	-0.63	-0.62	-0.17	0.29	0.41	-0.59
-0.53	-0.86	-0.58	-0.13	0.058	0.069	0.1	0.15	0.54	0.19	0.54	0.44
0.26	0.31	0.24	-0.19	0.18	0.39	-0.2	0.31	0.41	-0.57	-0.011	0.61
0.59	0.34	0.27	0.026	0.025	-0.0023	-0.93	-0.55	0.18	-0.37	-0.76	-0.54

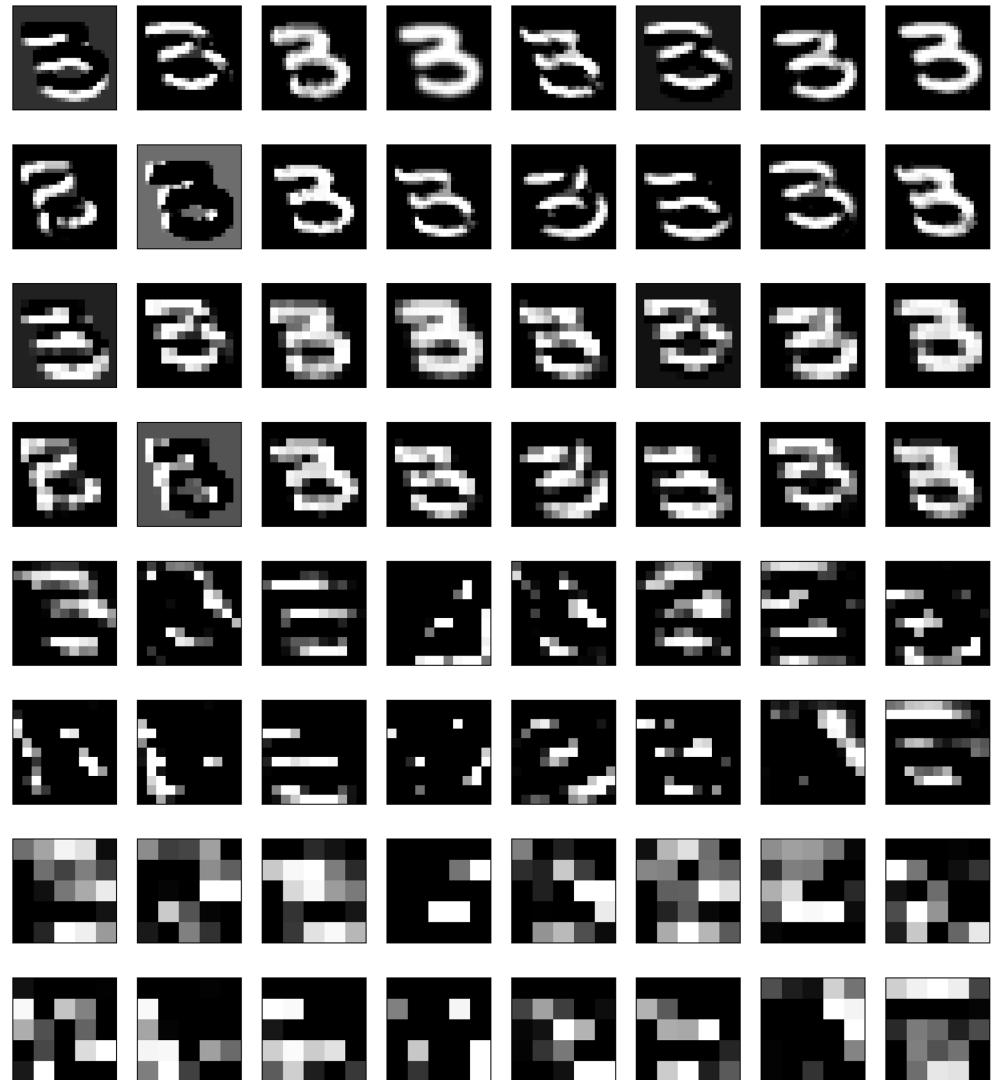
Filter Weights [layer.2]

0.28	0.12	-0.27	0.019	0.019	-0.36	-0.09	0.41	-0.019	-0.066	0.14	0.35
0.17	-0.17	-0.24	0.17	0.075	0.14	0.006	0.099	-0.089	-0.069	0.024	-0.29
-0.11	-0.58	0.18	-0.037	0.093	0.064	-0.06	0.0071	-0.054	-0.15	-0.14	-0.21
0.049	-0.19	0.15	-0.39	0.18	0.29	0.01	0.17	0.013	-0.7	-0.35	-0.17
0.16	-0.14	0.14	-0.38	0.032	0.28	0.07	0.25	0.041	-0.11	-0.11	-0.11
-0.45	-0.45	-0.17	-0.097	-0.048	0.072	-0.36	0.0023	0.16	0.22	0.36	0.24
0.025	-0.11	0.26	0.41	0.096	-0.47	0.01	-0.16	-0.13	0.021	0.00076	0.019
-0.1	-0.17	0.1	0.33	-0.017	-0.21	0.084	0.056	-0.22	-0.094	-0.1	-0.12
-0.21	0.21	0.14	0.028	-0.072	-0.15	0.04	0.13	0.027	0.059	0.053	0.089
-0.13	-0.00079	0.11	0.11	-0.11	0.0084	-0.094	0.057	0.07	-0.13	0.24	0.34
0.019	-0.3	0.35	0.0029	0.16	0.099	0.06	0.21	-0.24	-0.18	-0.1	-0.28
-0.31	-0.26	0.48	0.2	0.3	0.057	-0.14	-0.16	-0.11	-0.17	-0.38	-0.15

Activations of each layer



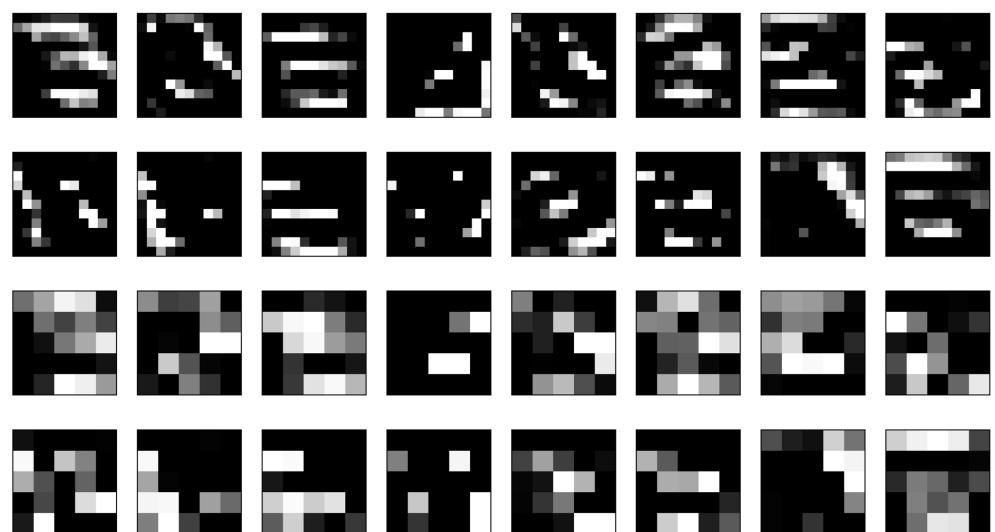
layer.0
 26×26



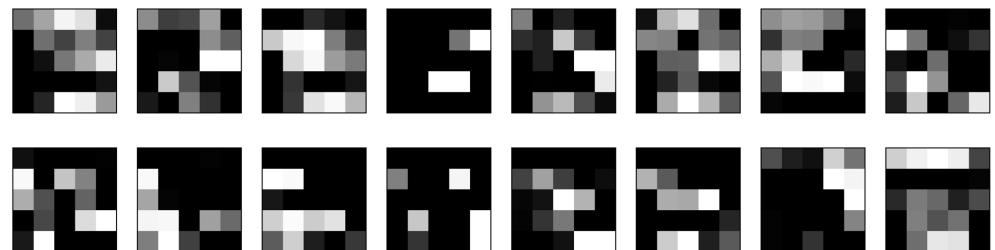
layer.1
 13×13



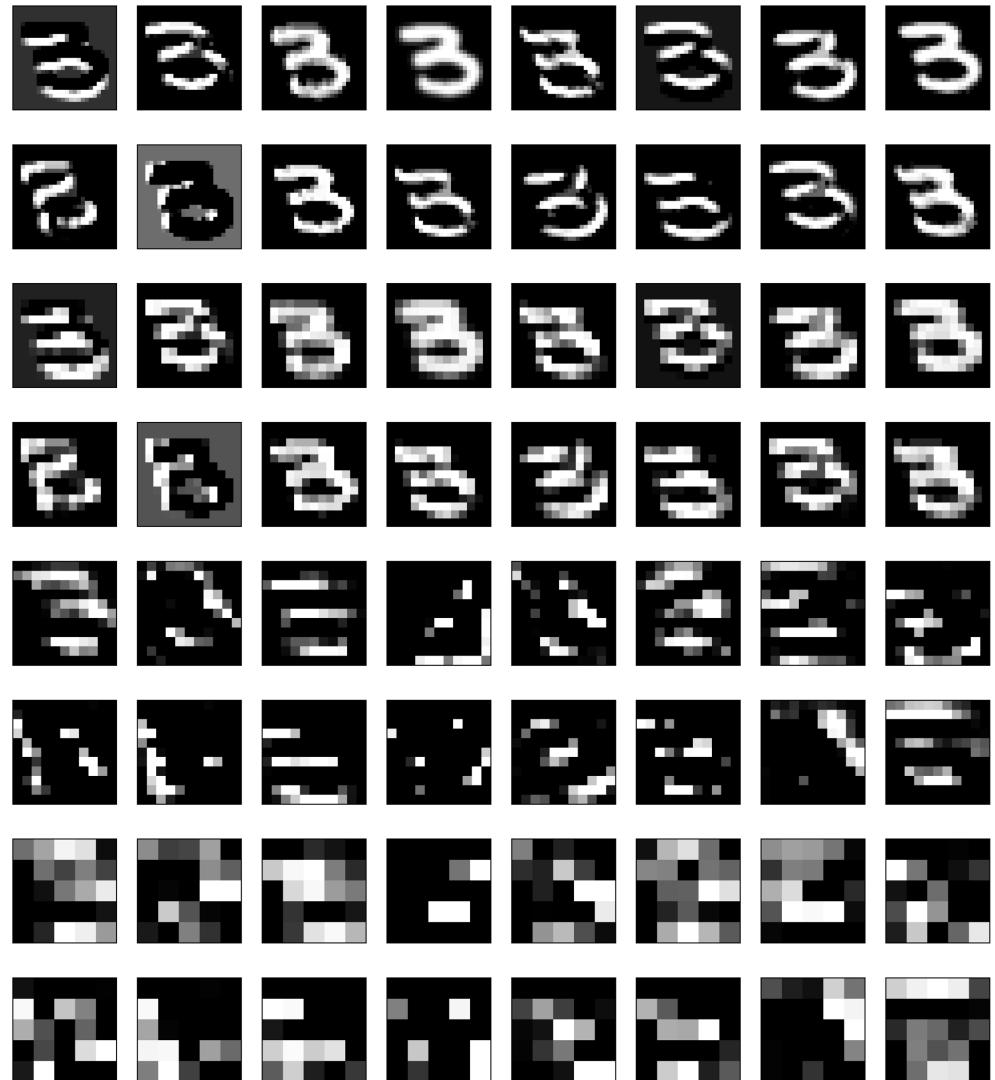
layer.2
 11×11



layer.3
 5×5



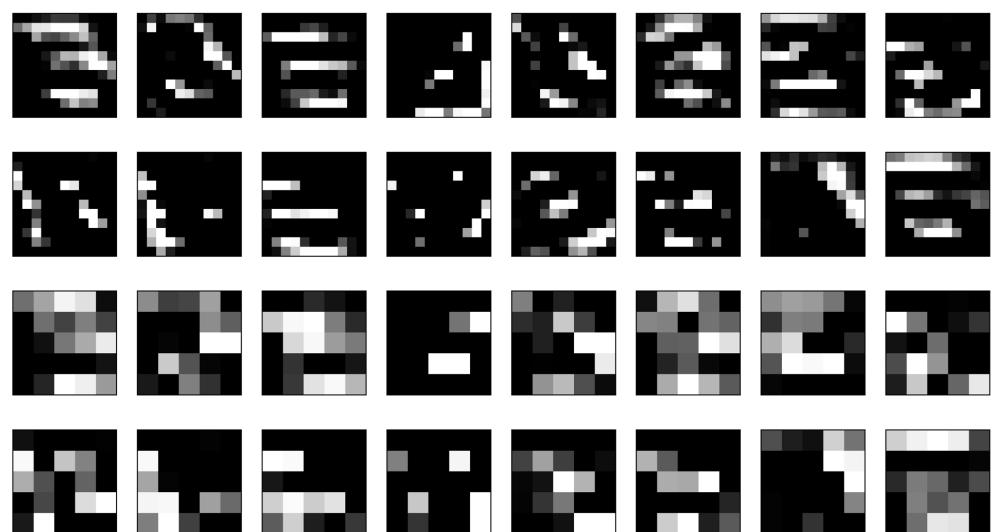
layer.0
 26×26



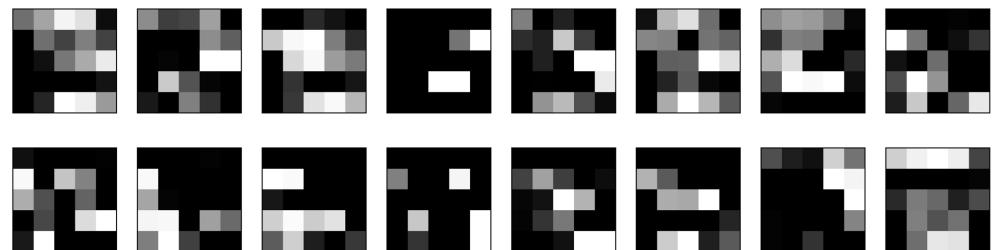
layer.1
 13×13



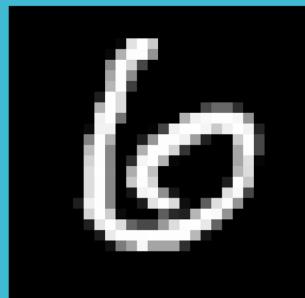
layer.2
 11×11



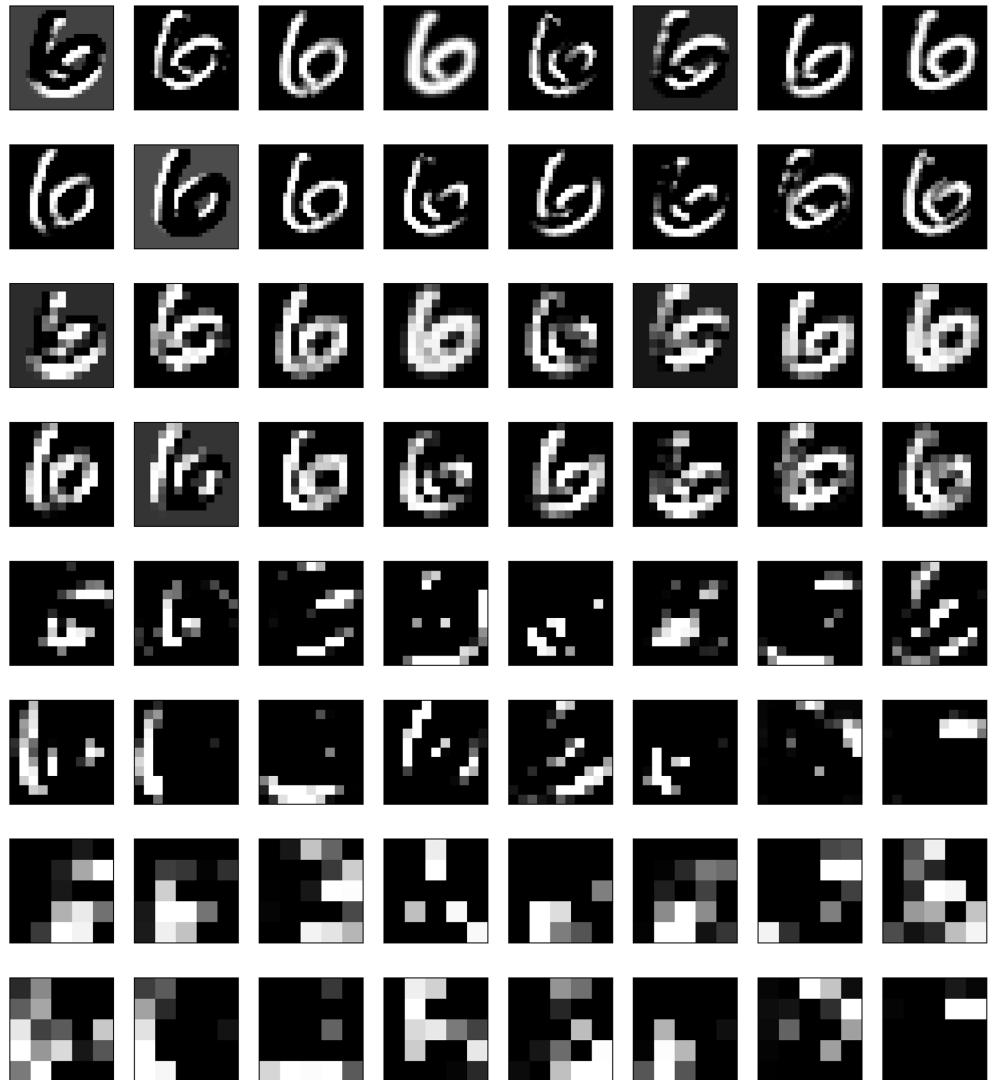
layer.3
 5×5



Activations of each layer



layer.0
 26×26



layer.1
 13×13

layer.2
 11×11

layer.3
 5×5

Activations of each layer



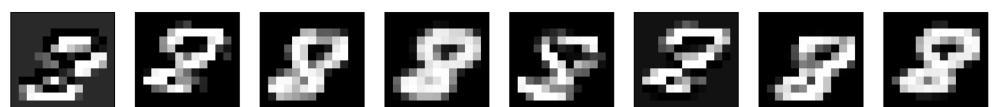
layer.0
26x26



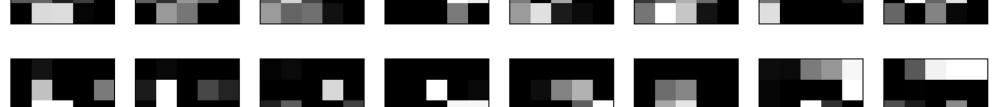
layer.1
13x13



layer.2
11x11

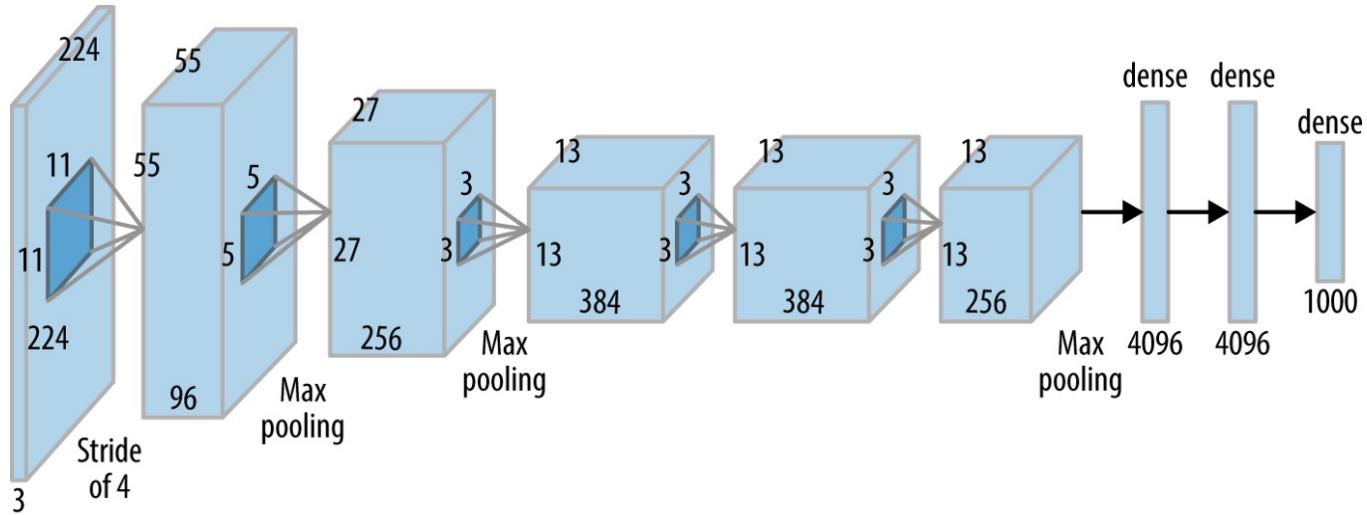


layer.3
5x5



Alexnet

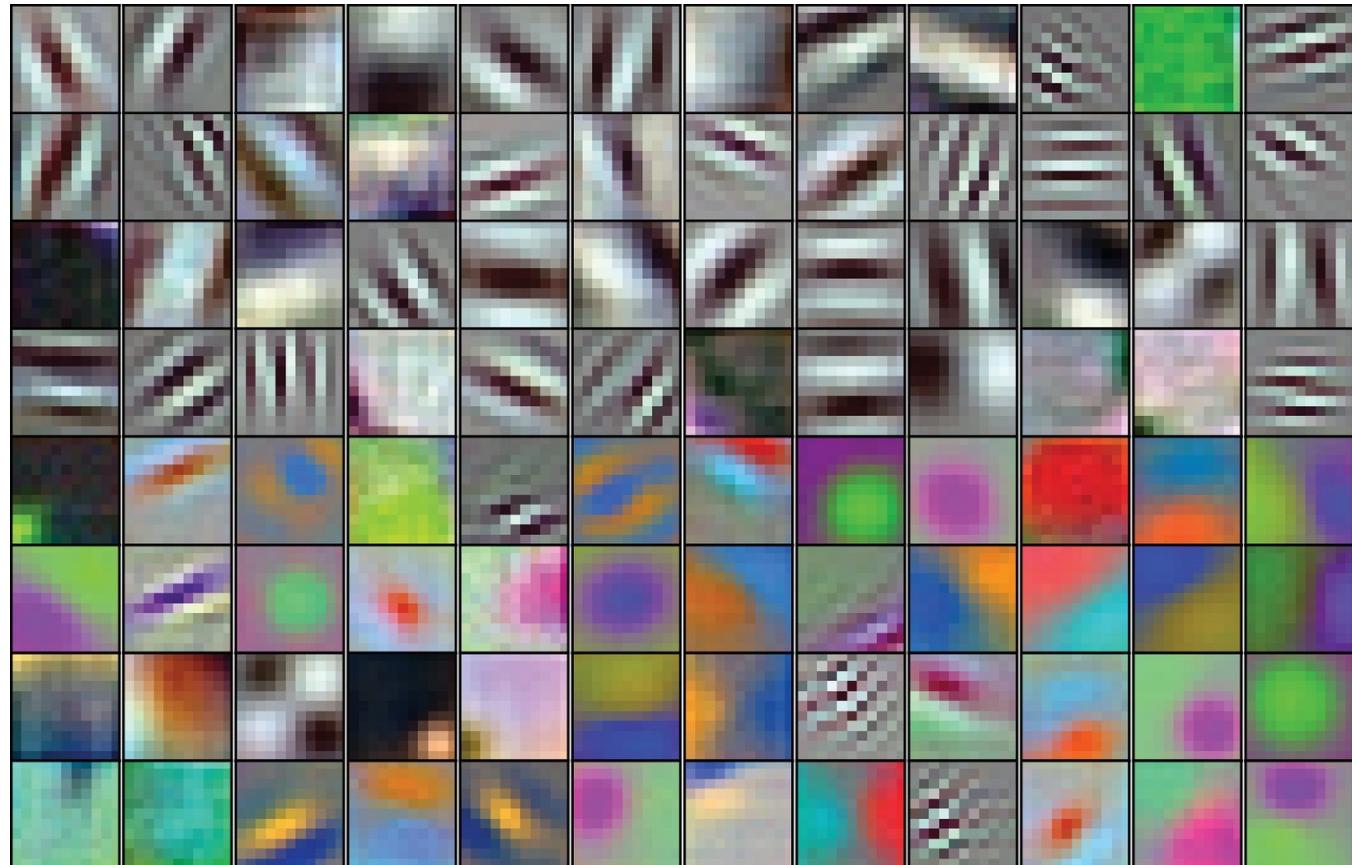
[Krizhevsky et al. 2012]



[ImageNet Large Scale Visual Recognition Challenge 2012]

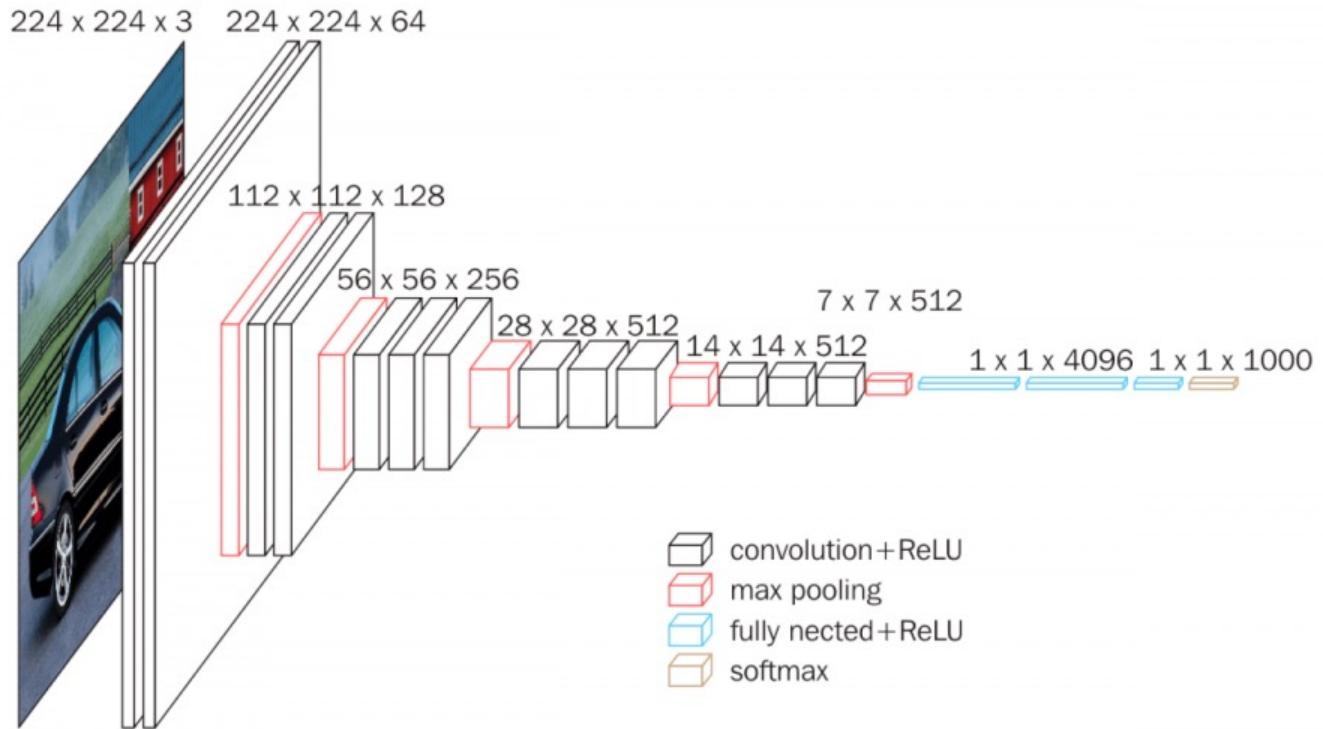
Alexnet

[Krizhevsky
et al. 2012]



VGG

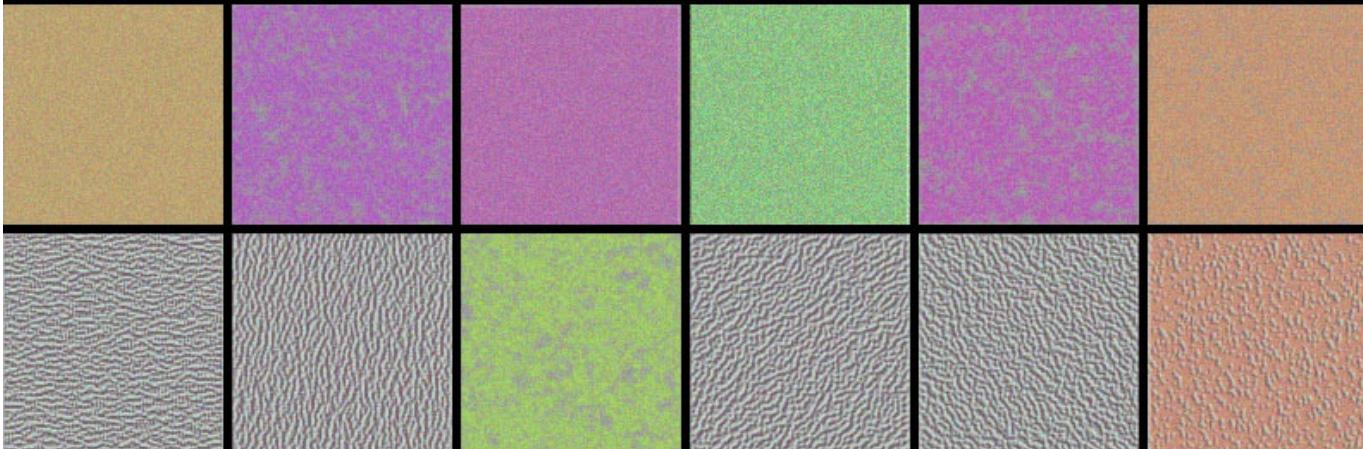
[Simonyan et al. 2014]



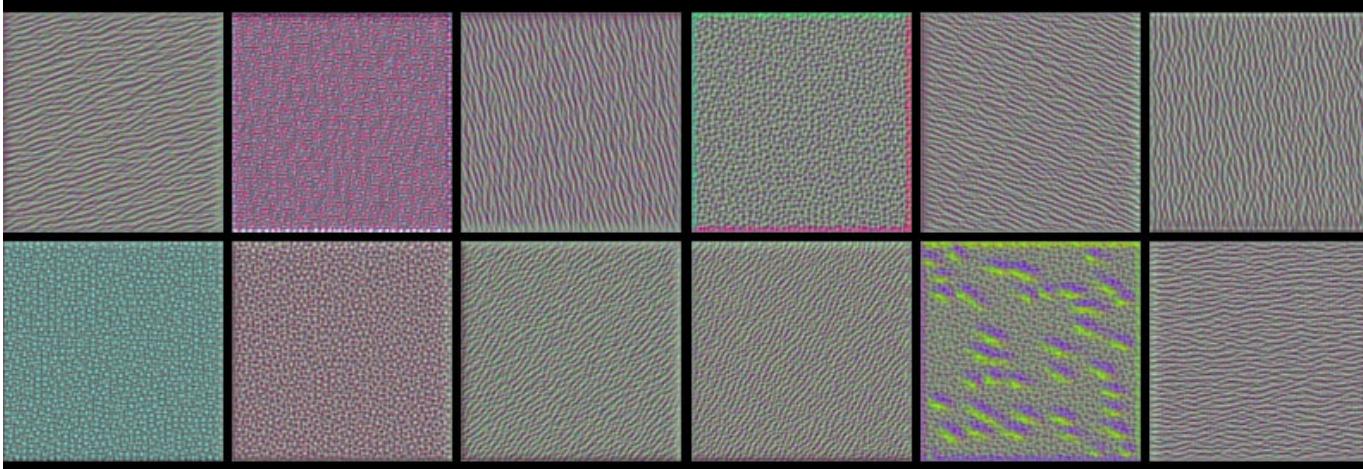
VGG

[Simonyan et
al. 2014]

conv1_1: a few of the 64 filters



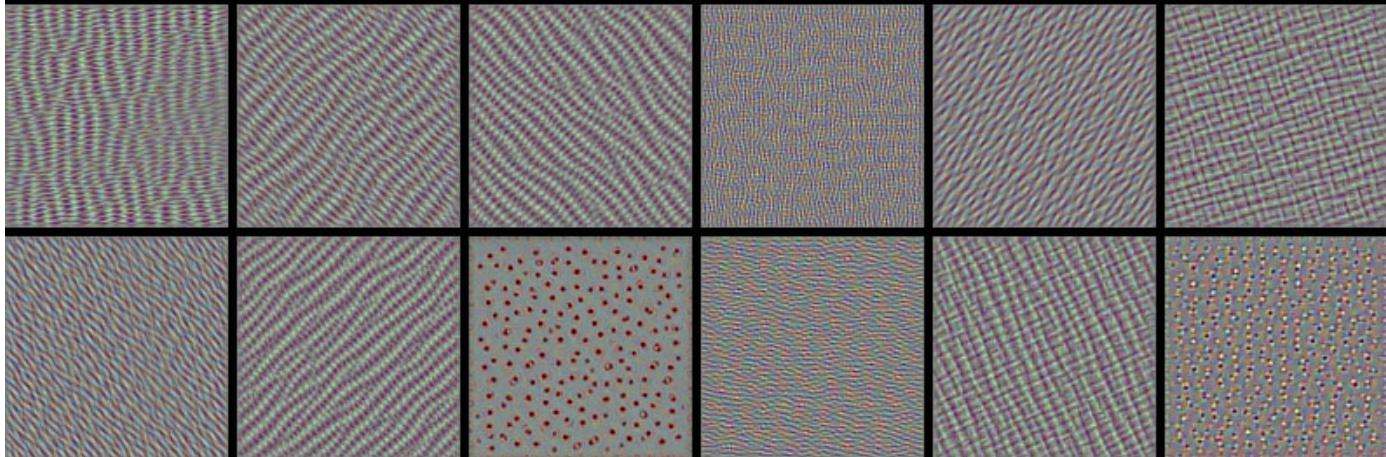
conv2_1: a few of the 128 filters



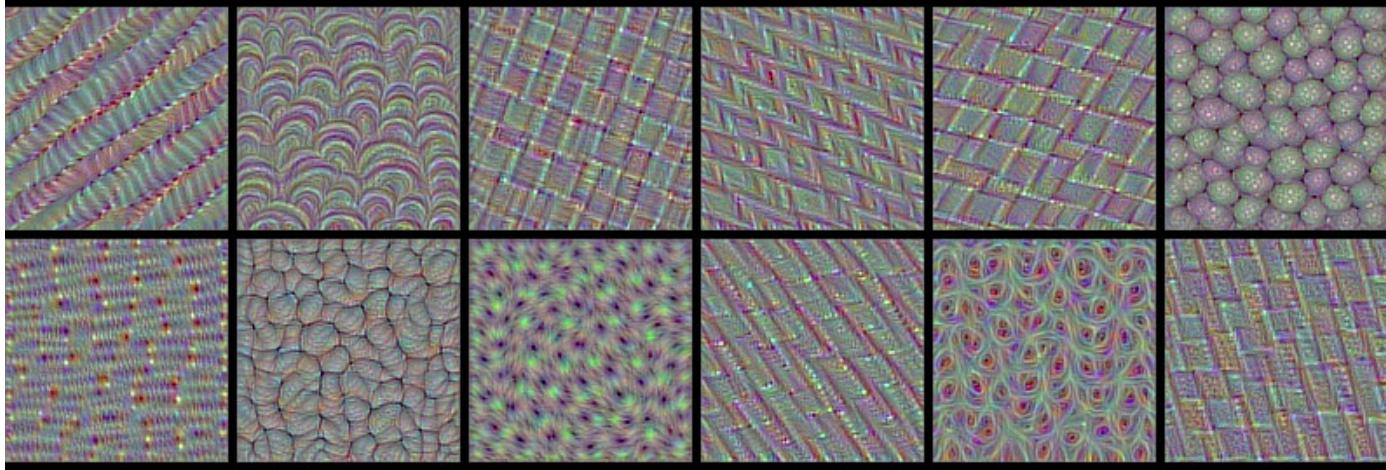
VGG

[Simonyan et
al. 2014]

conv3_1: a few of the 256 filters



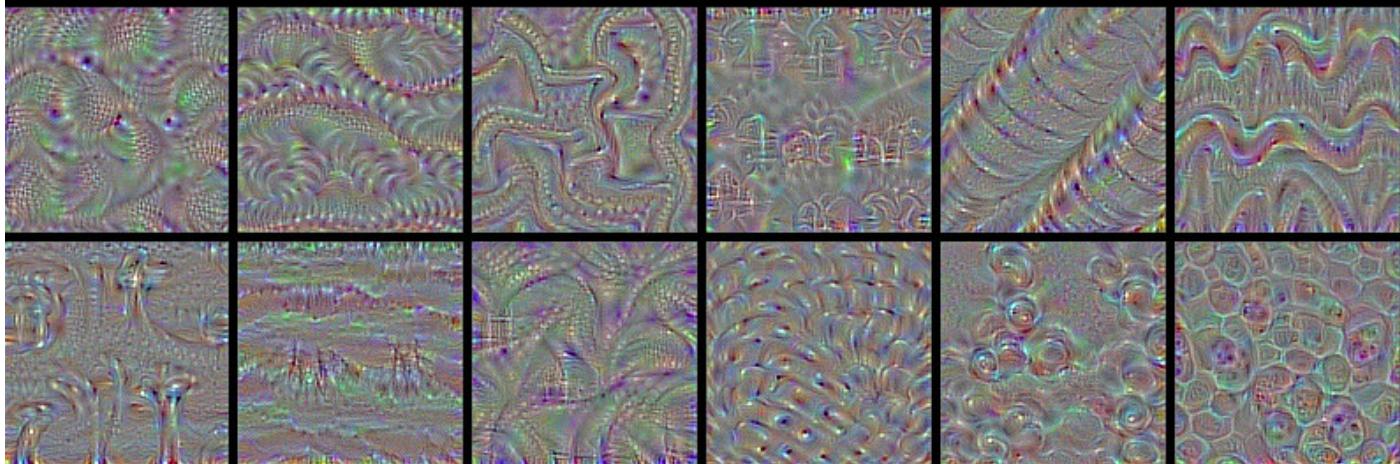
conv4_1: a few of the 512 filters



VGG

[Simonyan et
al. 2014]

conv5_1: a few of the 512 filters



Interesting extensions

- Inceptionism: Going Deeper into Neural Networks

<https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>

- Understanding Neural Networks Through Deep Visualization

<https://yosinski.com/deepvis>

- Feature Visualization

<https://distill.pub/2017/feature-visualization/>