Introduction to Deep Learning

Olivier Augereau

augereau@enib.fr

@oaugereau

Basic principles of Convolutional Neural Networks

CNN

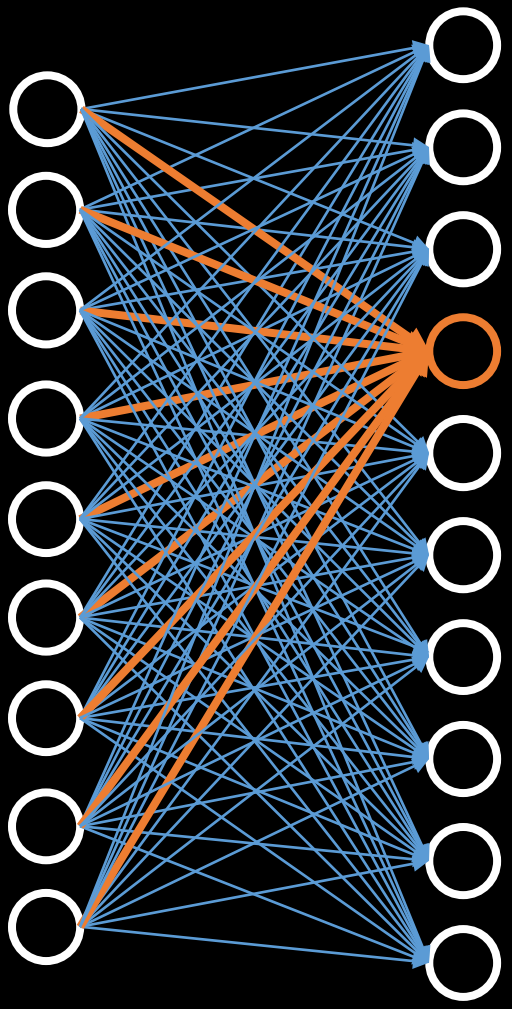
For computer vision

- Not fully connected but locally connected

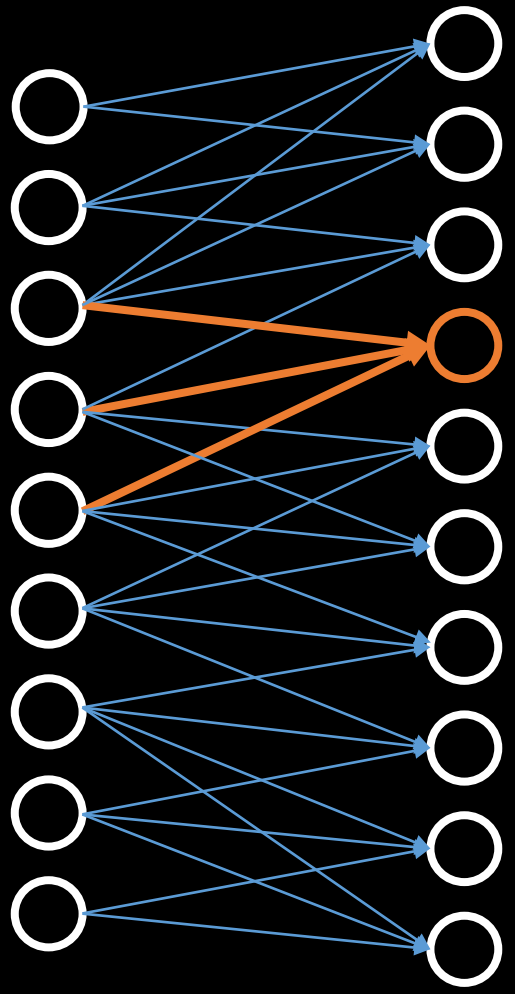
Special layers

- Convolution (extract local features)
- Max-Pooling (reduce the size / down sampling)

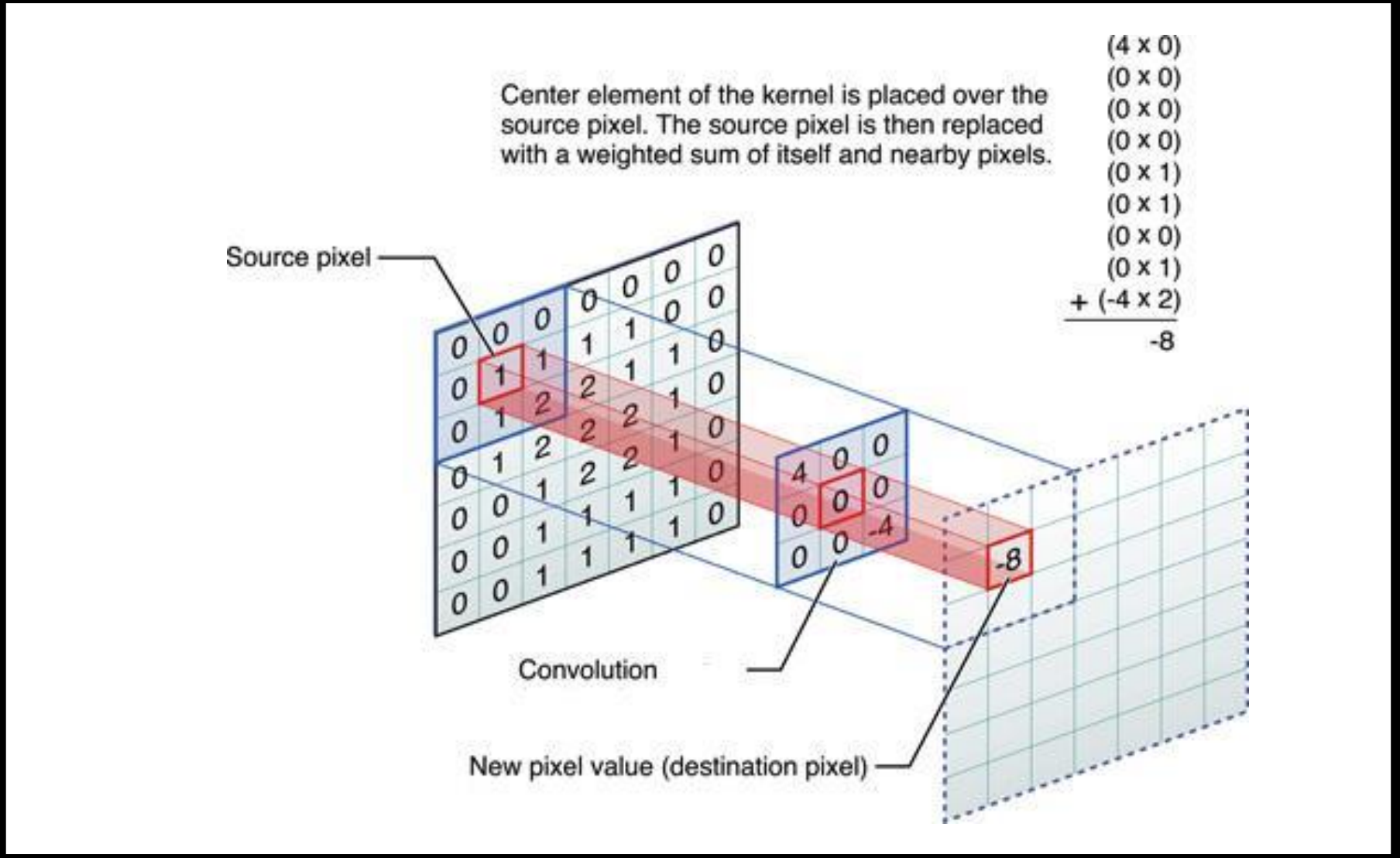
fully connected



locally connected



Convolution



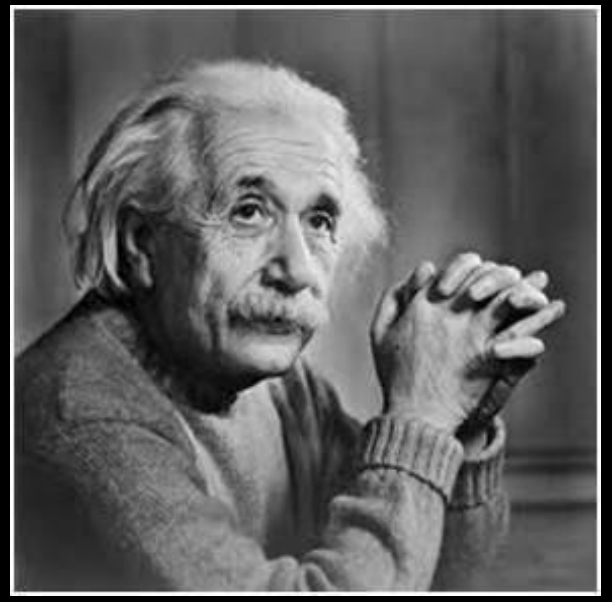
Convolution : Sobel

-1	0	+1
-2	0	+2
-1	0	+1

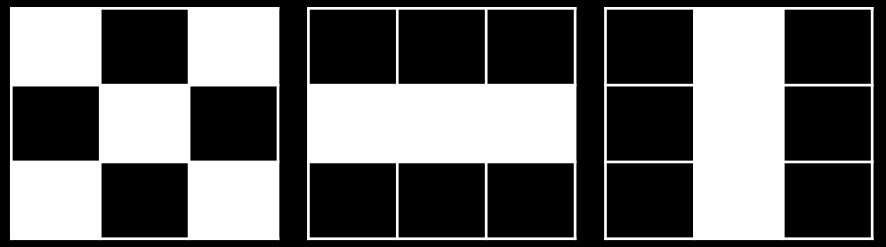
Gx

+1	+2	+1
0	0	0
-1	-2	-1

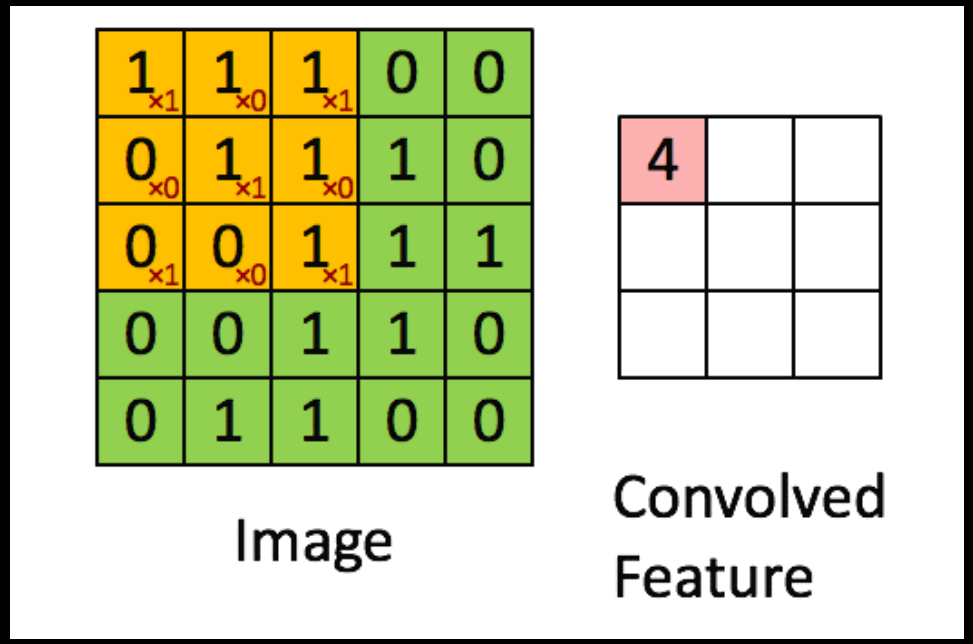
Gy



Convolution



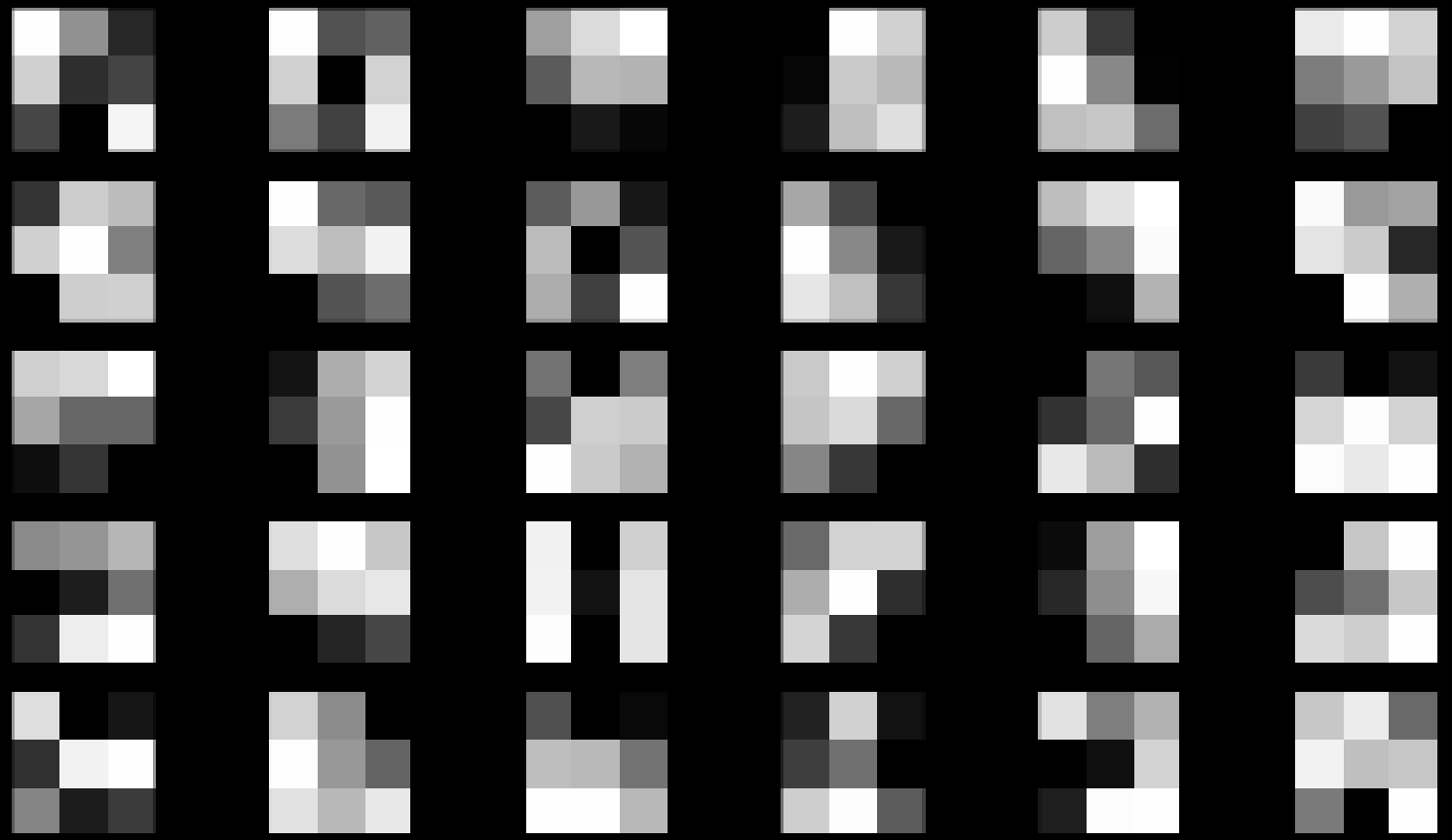
Kernels



Convolution (extract local patterns)

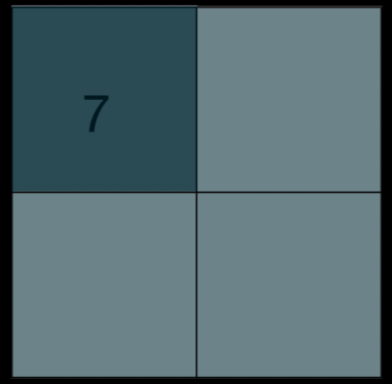
- Local patterns are translation invariant
- Learn spatial hierarchies (edge -> corner -> object parts)

Convolution layer

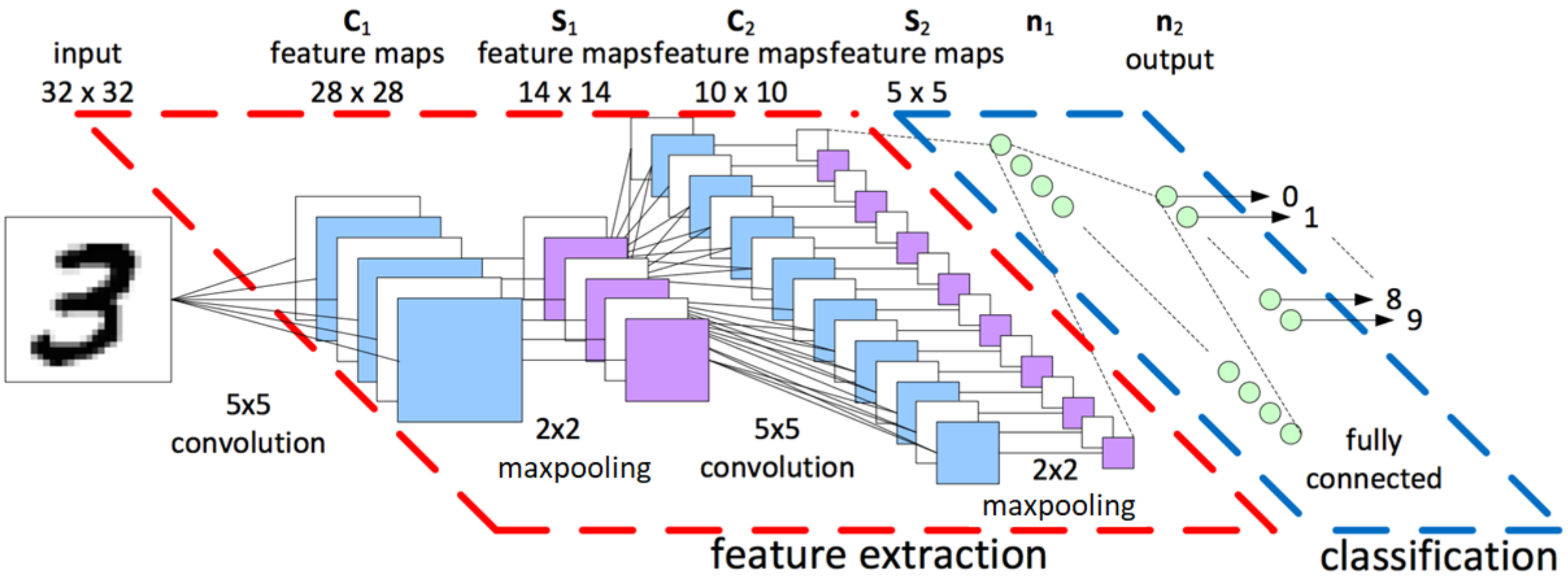


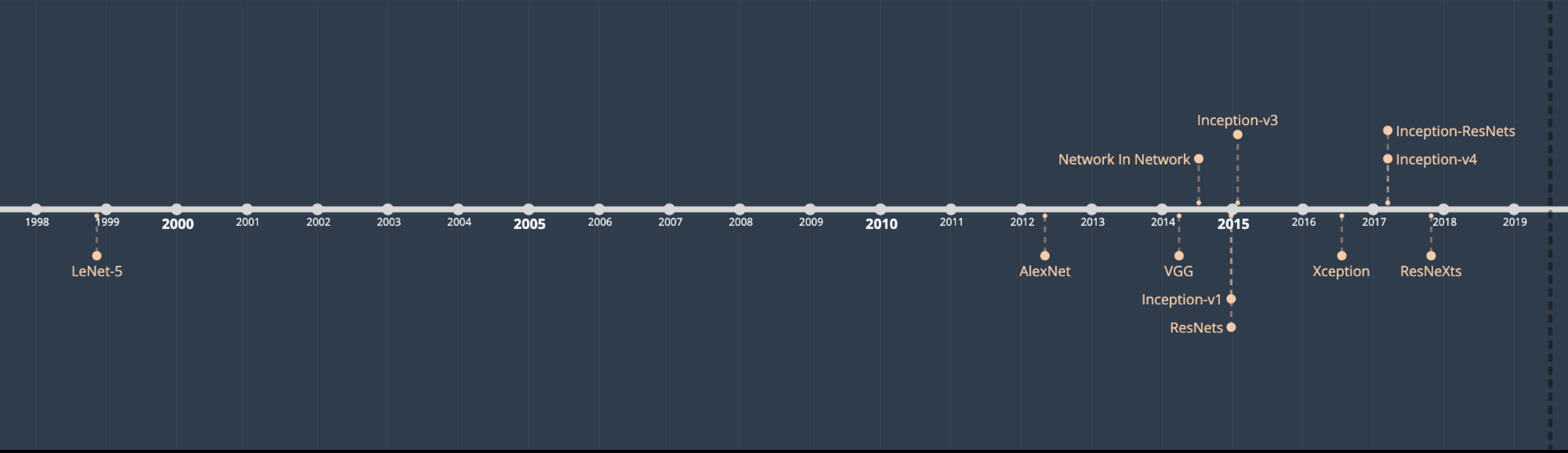
Max-Pooling (reduce the size)

5	5	4	5
7	6	5	3
3	5	6	5
3	4	8	6



CNN basic architecture

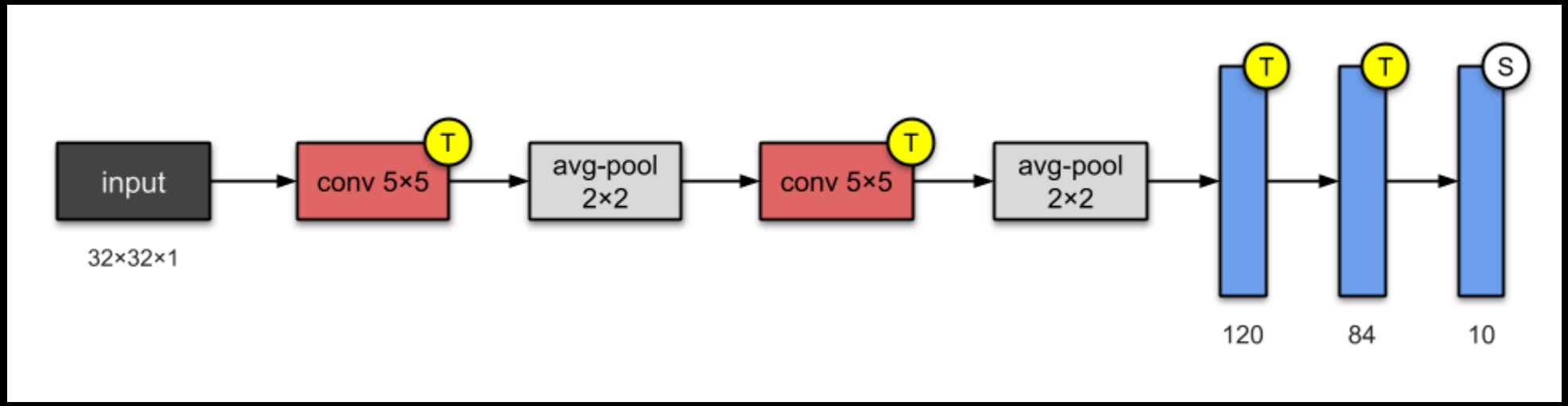




Some famous CNN architechtures :

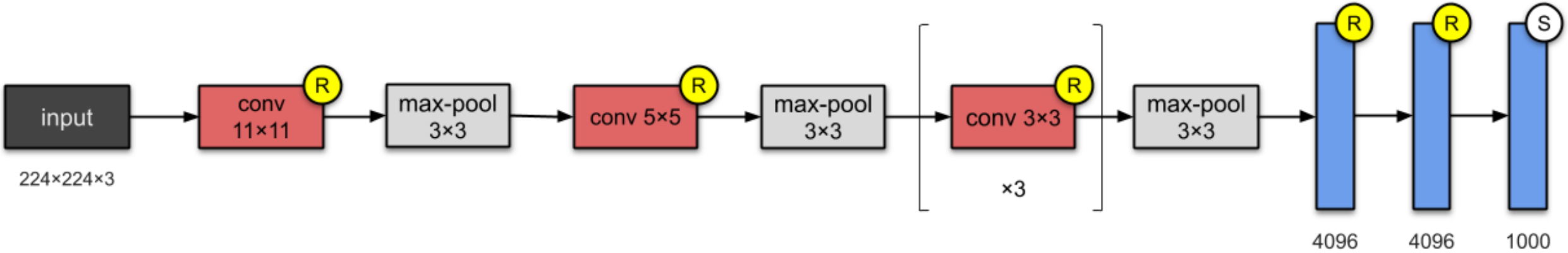
LeNet-5, AlexNet, VGGNet, GoogLeNet, ResNet

LeNet-5 (1998)



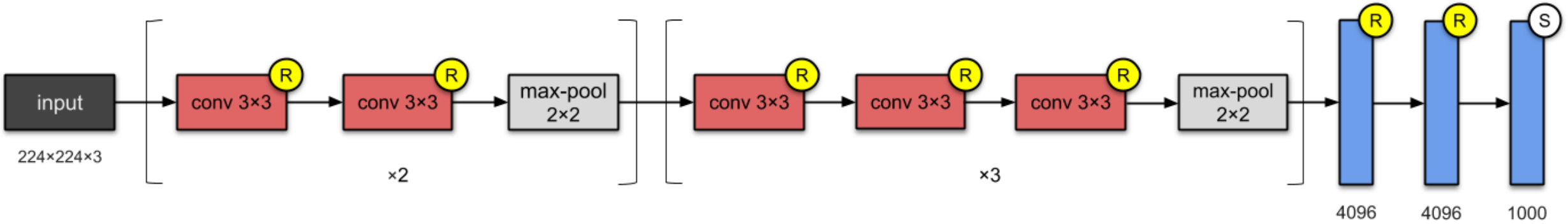
It has 2 convolutional and 3 fully-connected layers (hence “5”)
This architecture has about 60,000 parameters.

AlexNet (2012)



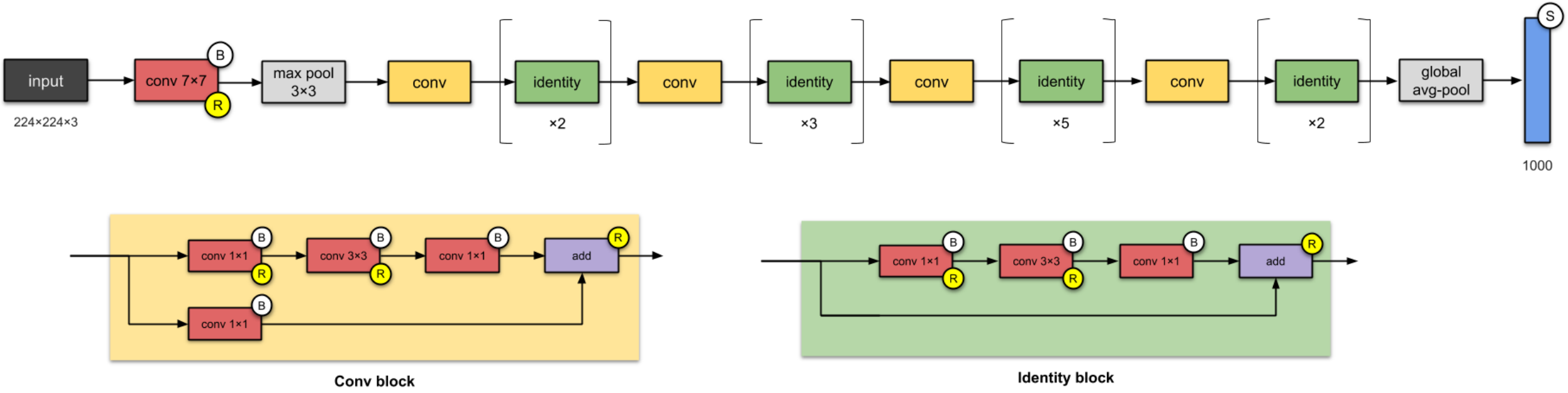
It has 5 convolutional and 3 fully-connected layers
This architecture has about 60M parameters.
Implement ReLUs as activation functions

VGG-16 (2014)



It has 13 convolutional and 3 fully-connected layers
This architecture has about 138M parameters.

ResNet-50 (2015)



skip connections

batch normalisation

This architecture has about 26M parameters.

References

- Lex Friedman, MIT
- François Chollet, Google
- Andrew Ng, Coursera / Standford
- Fei Fei Li, Standford
- Geoffrey Hinton, Yoshua Bengio, Yann Le Cun