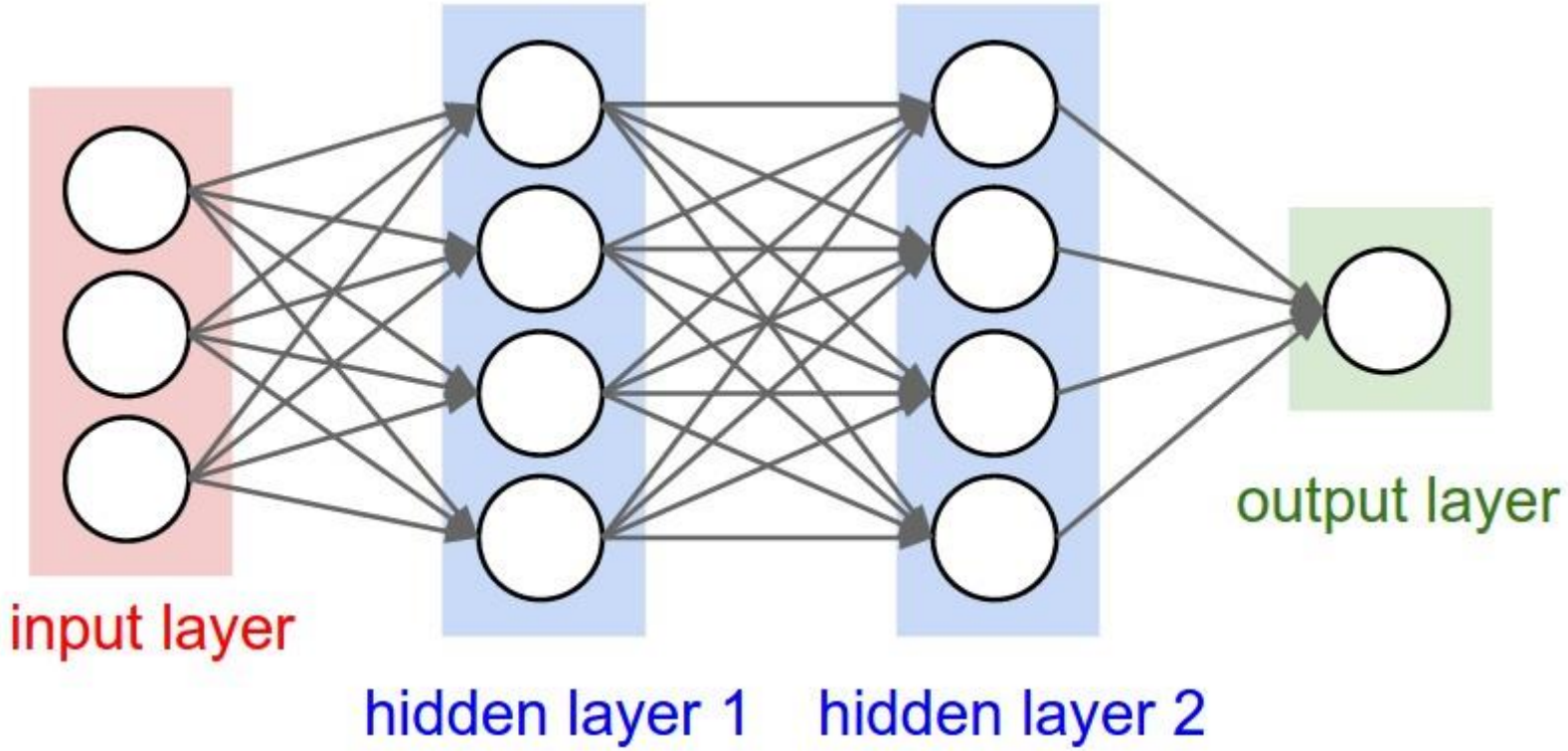
Convolutional Neural Networks Introduction

Piotr Mazur

Agenda

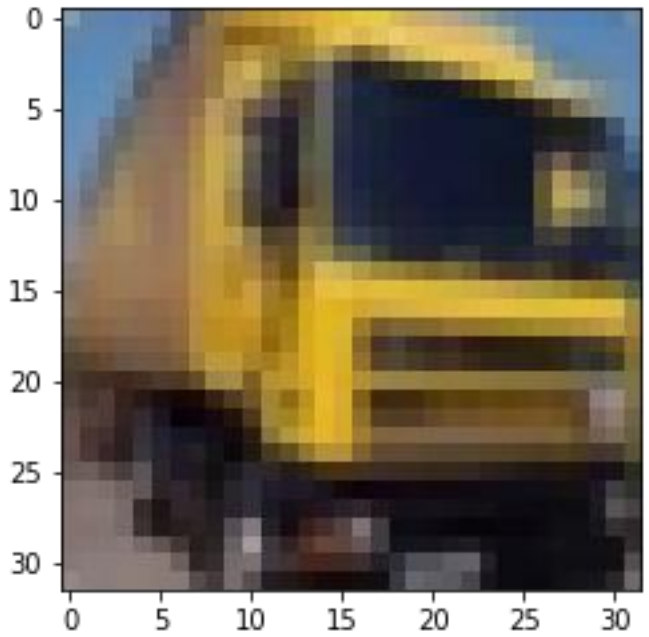
- Introduction
- Basic CNN architecture
- Convolutional Network Layers
  - Convolutional Layer
  - Pooling Layer
- Layer sizing
- Important CNN architectures
- Coding session

Fully-connected Neural Network



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

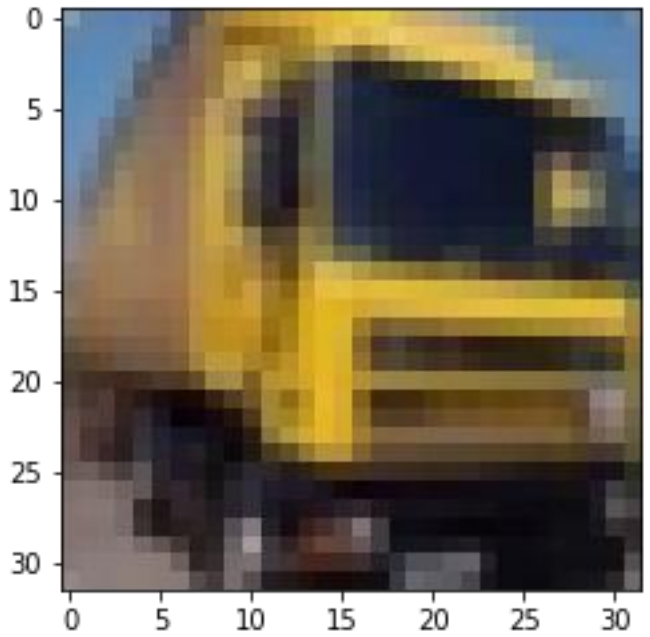
”Big” Data



Source: CIFAR-10

32x32x3

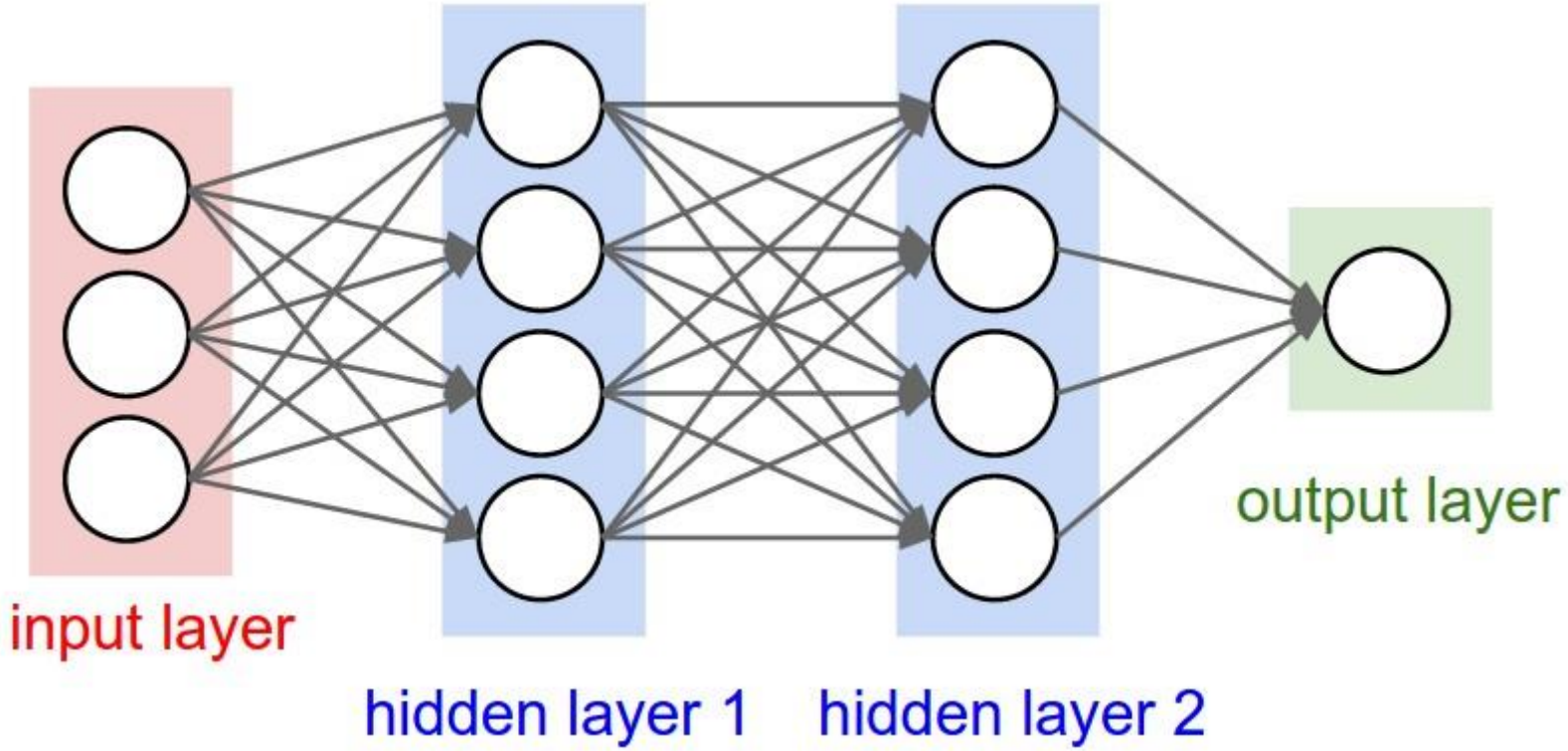
”Big” Data



Source: CIFAR-10

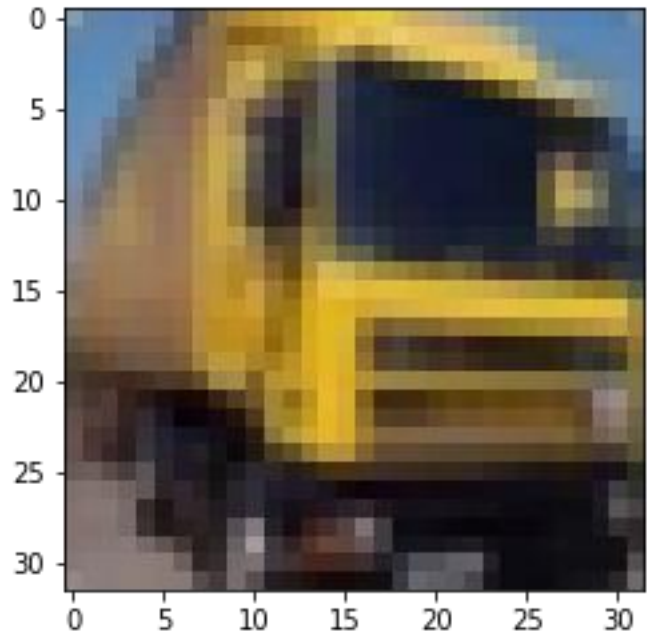
3.072 features

Fully-connected Neural Network



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

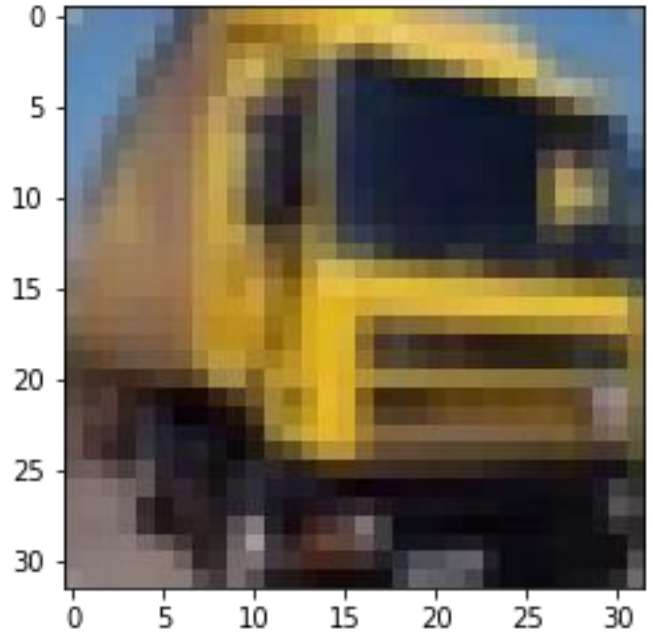
”Big” Data



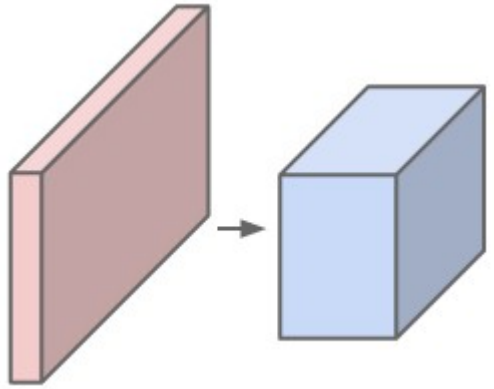
Source: CIFAR-10

12.288 parameters

Data arrangement

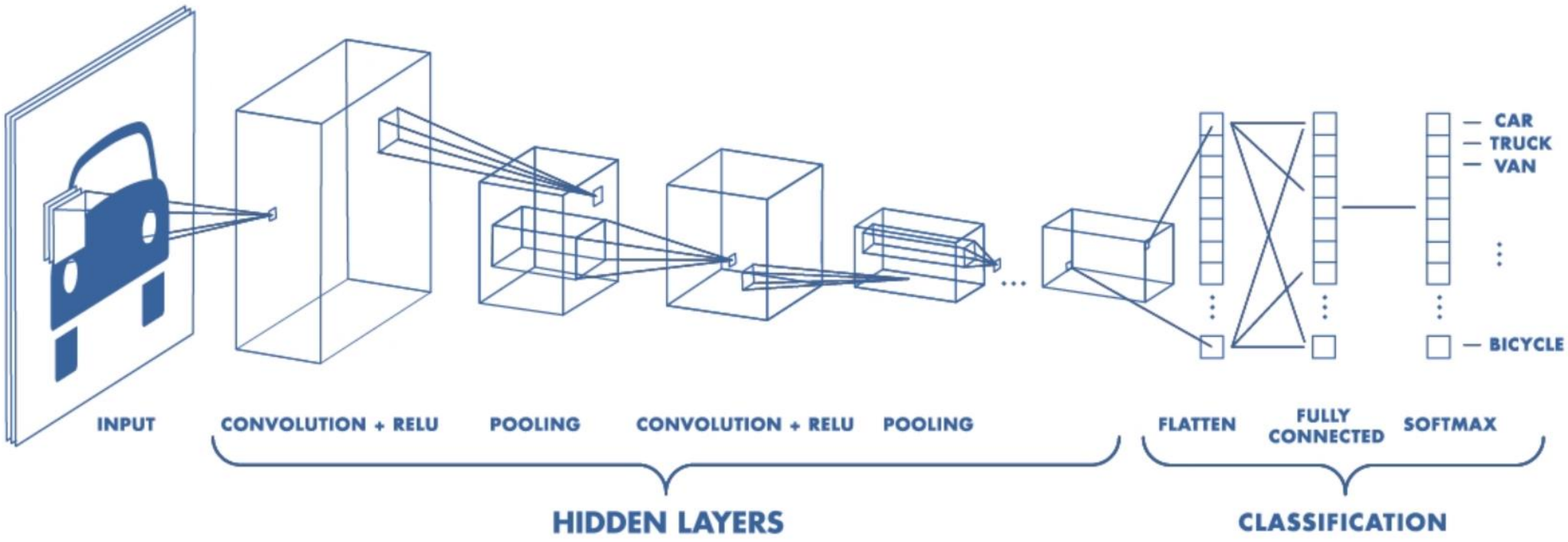


Source: CIFAR-10



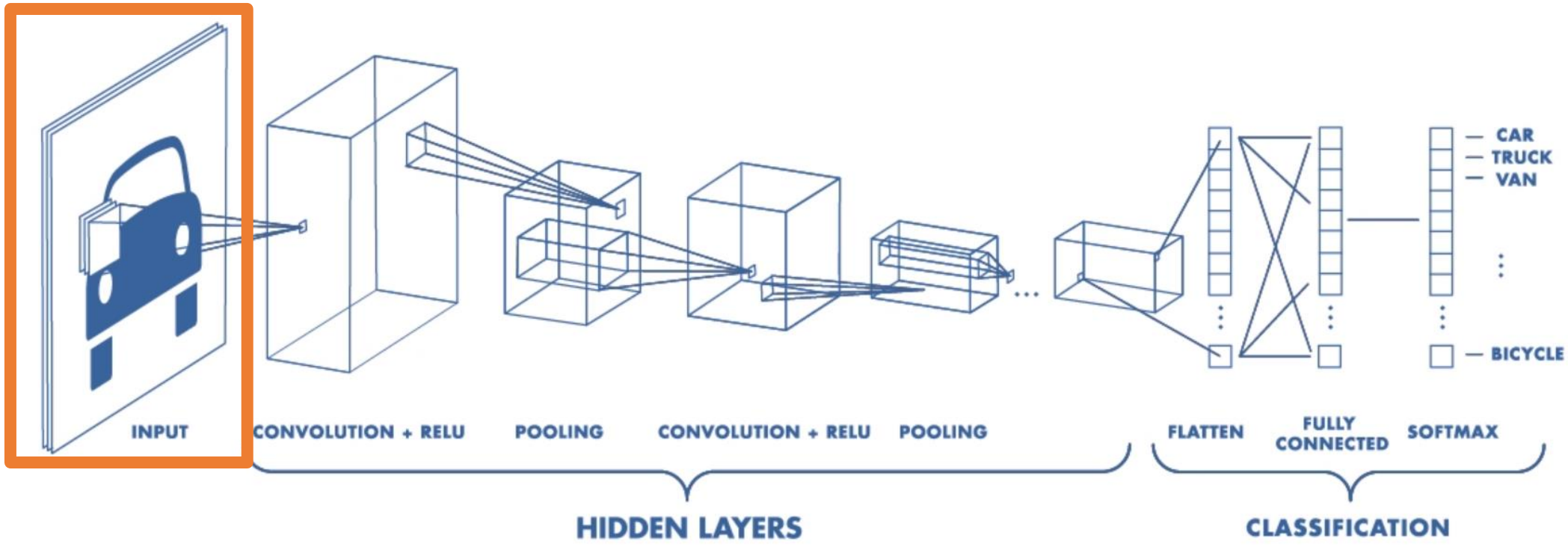
Source: https://cs231n.github.io/convolutional-networks

Convolutional Neural Network (ConvNet)



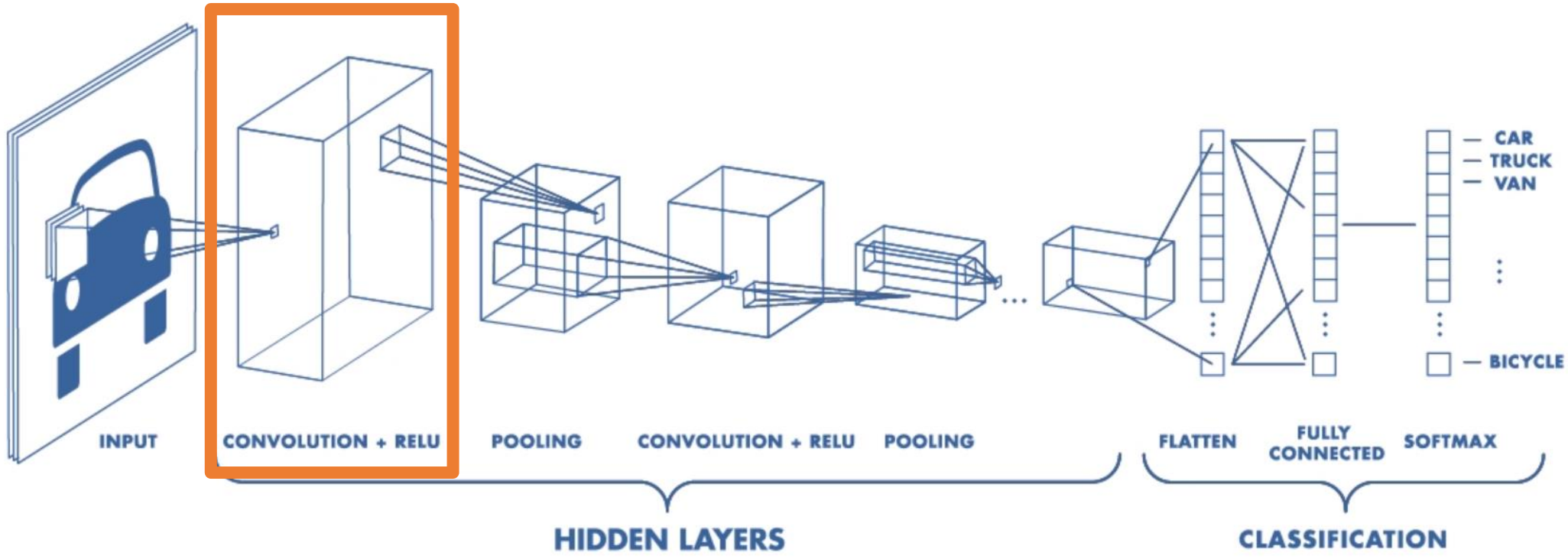
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Neural Network (ConvNet)



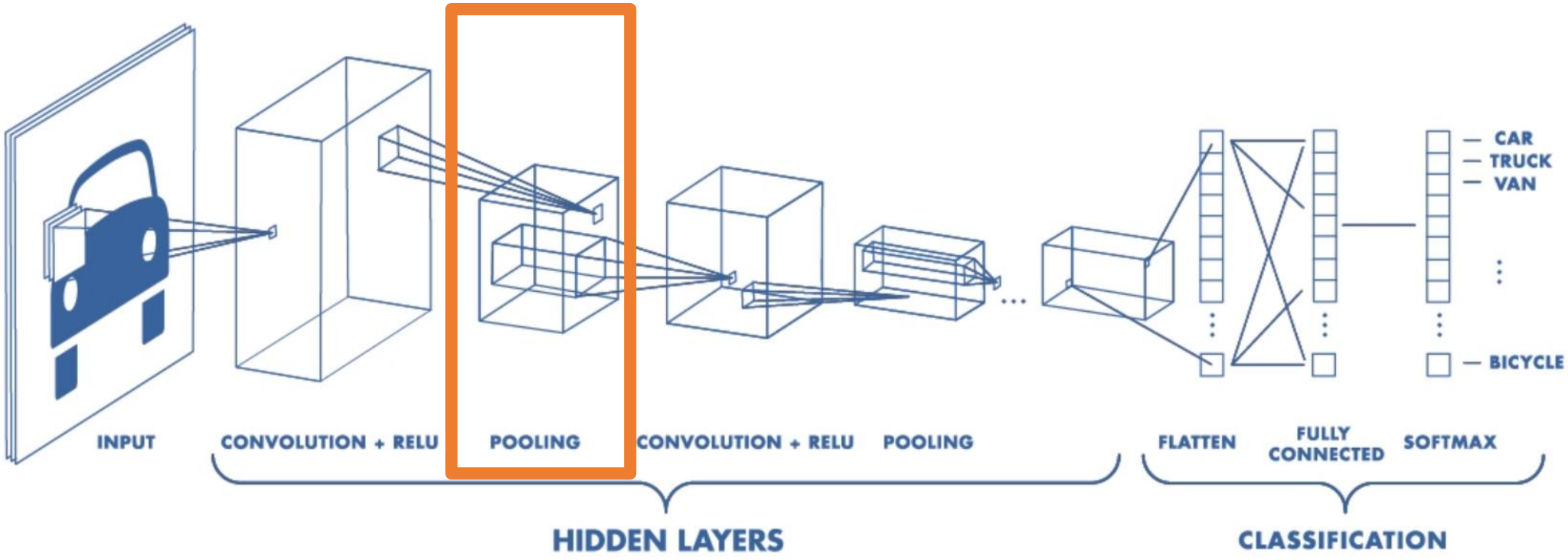
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Neural Network (ConvNet)



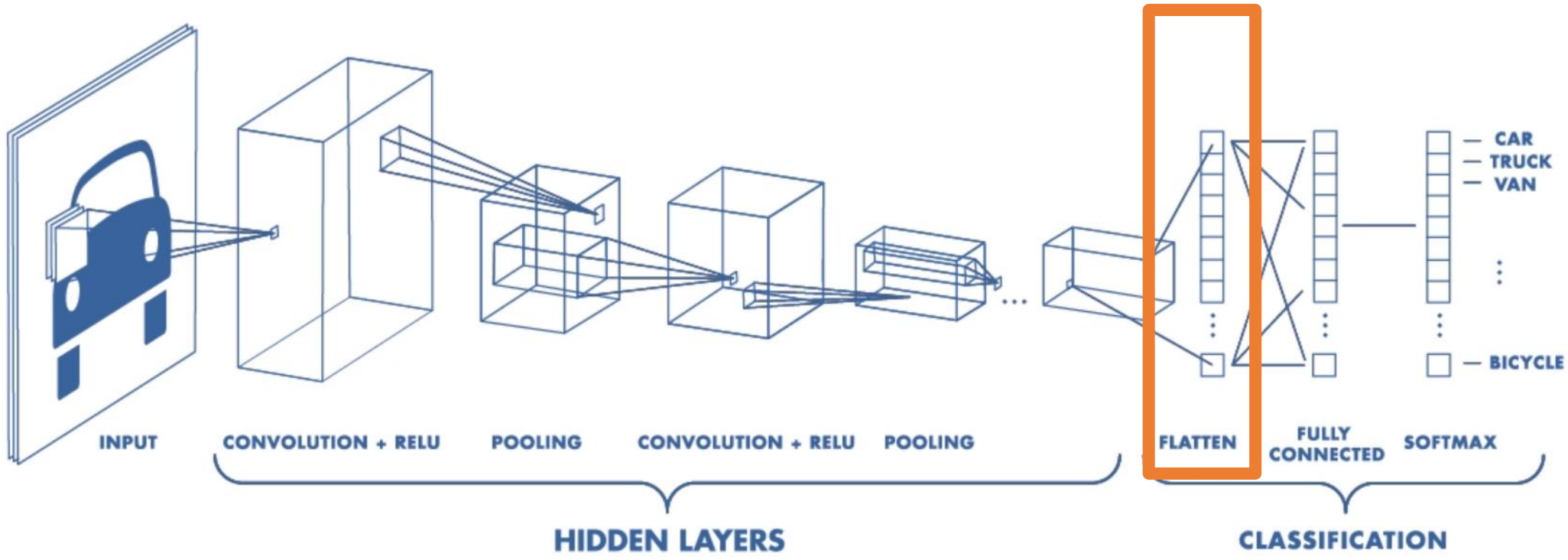
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Neural Network (ConvNet)



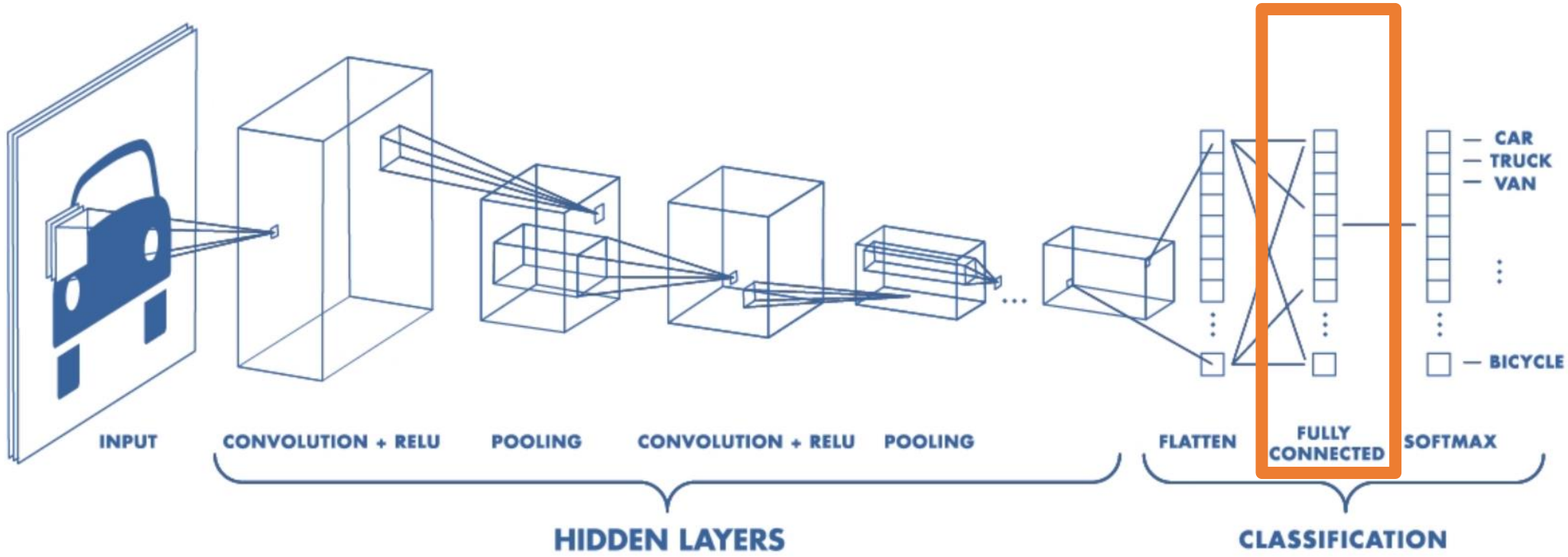
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Neural Network (ConvNet)



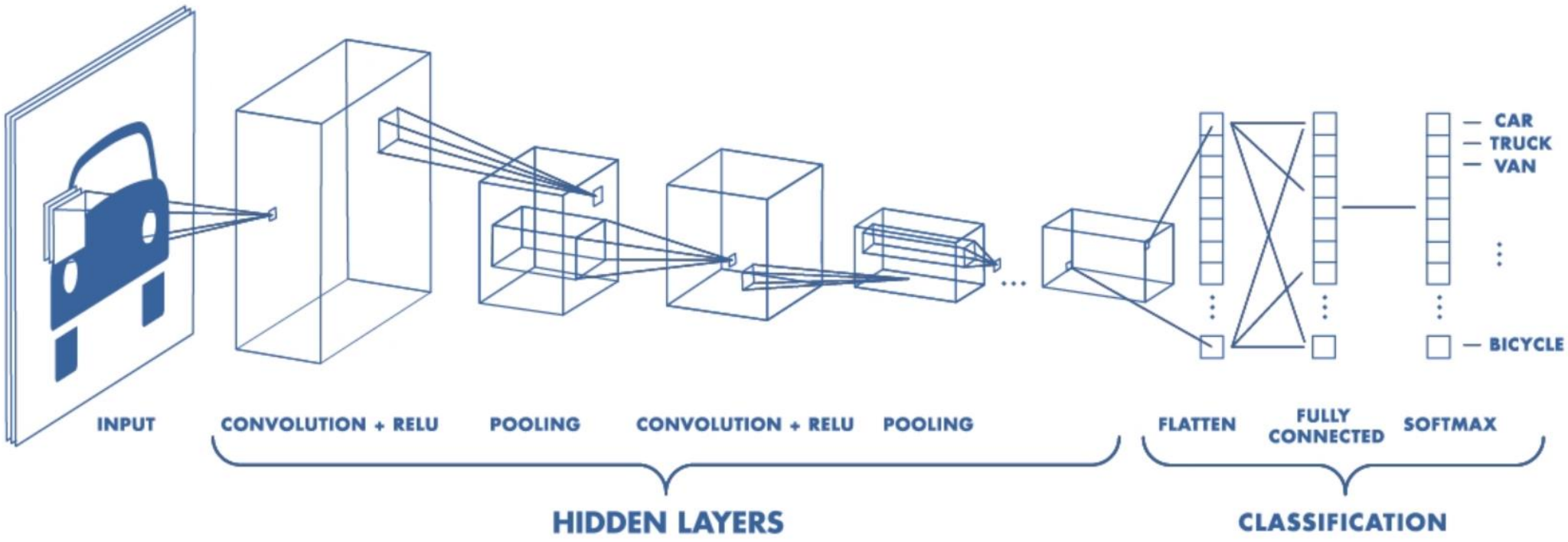
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Neural Network (ConvNet)



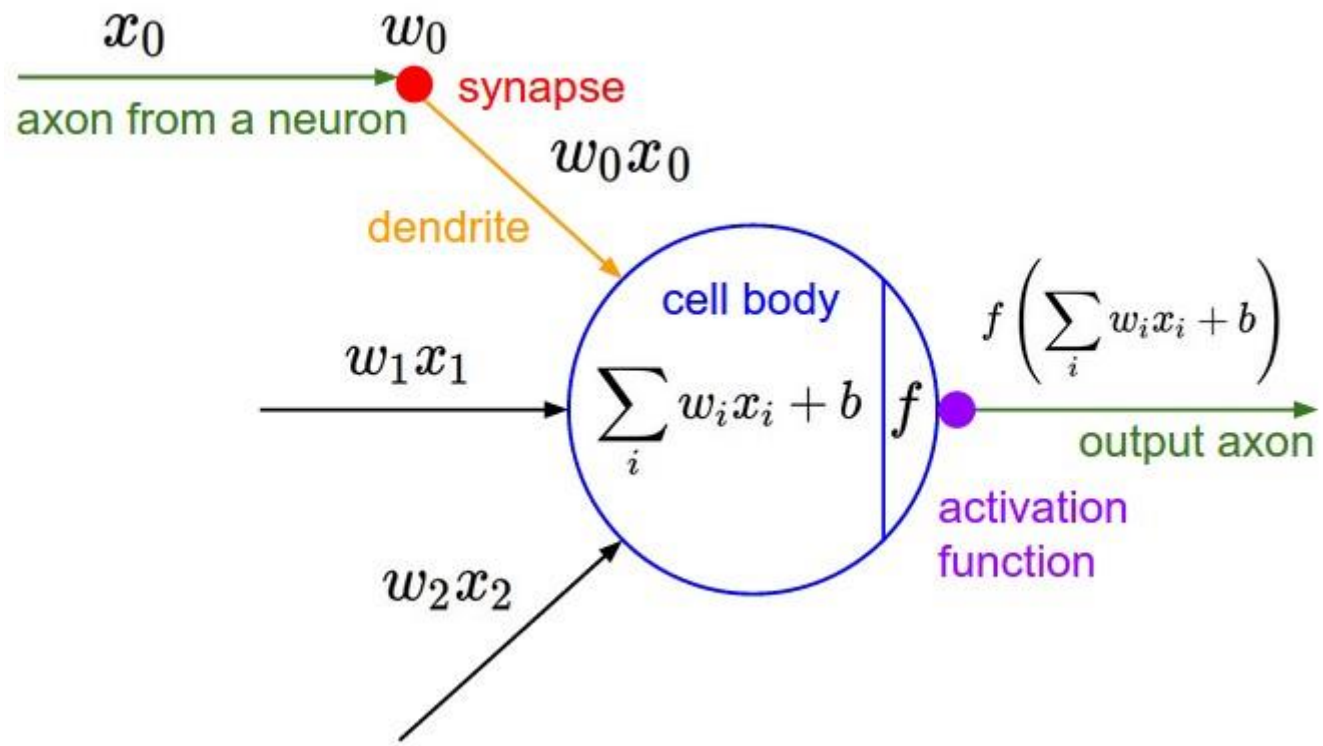
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Neural Network (ConvNet)



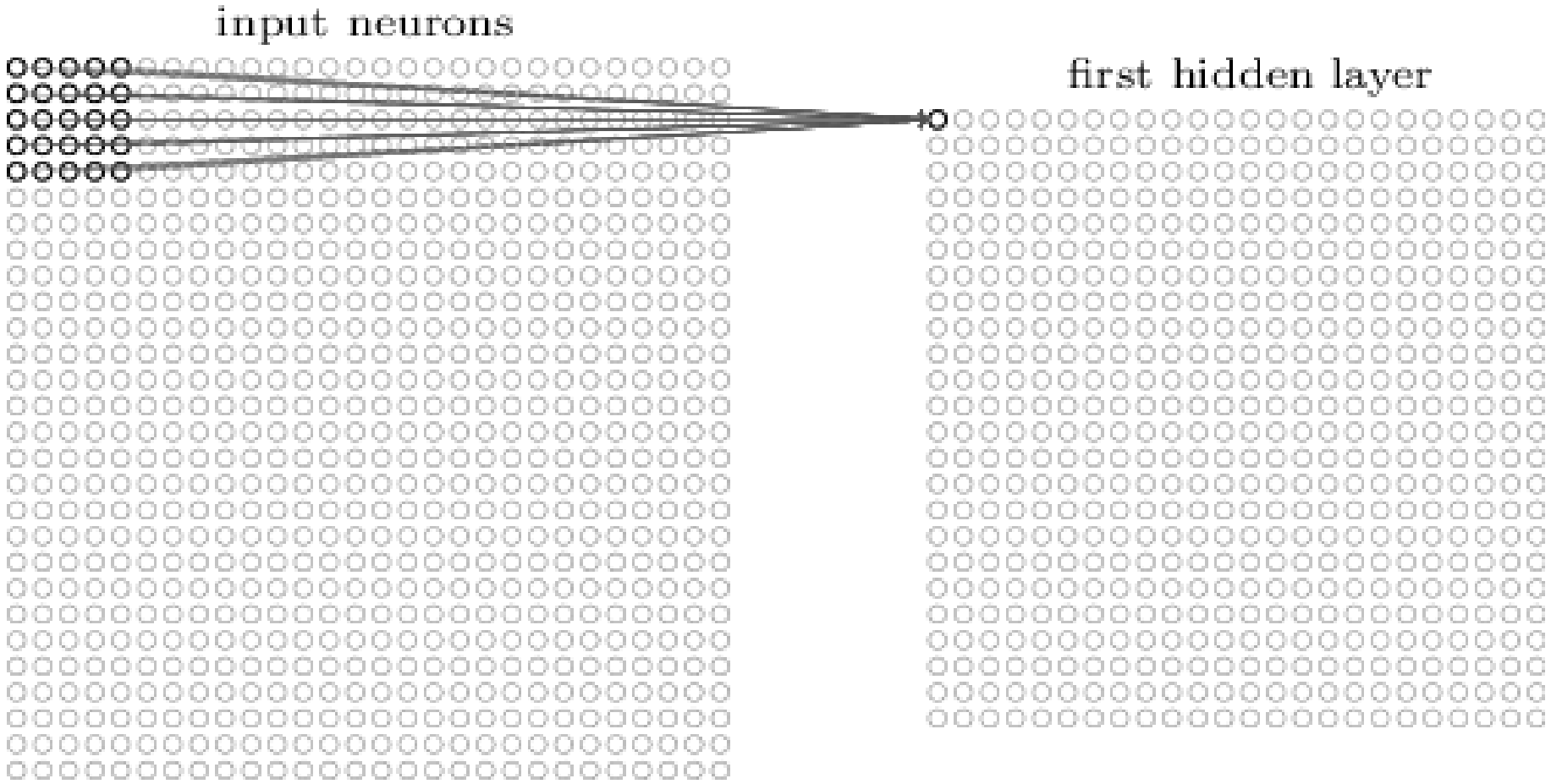
Source: https://mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Convolutional Layer



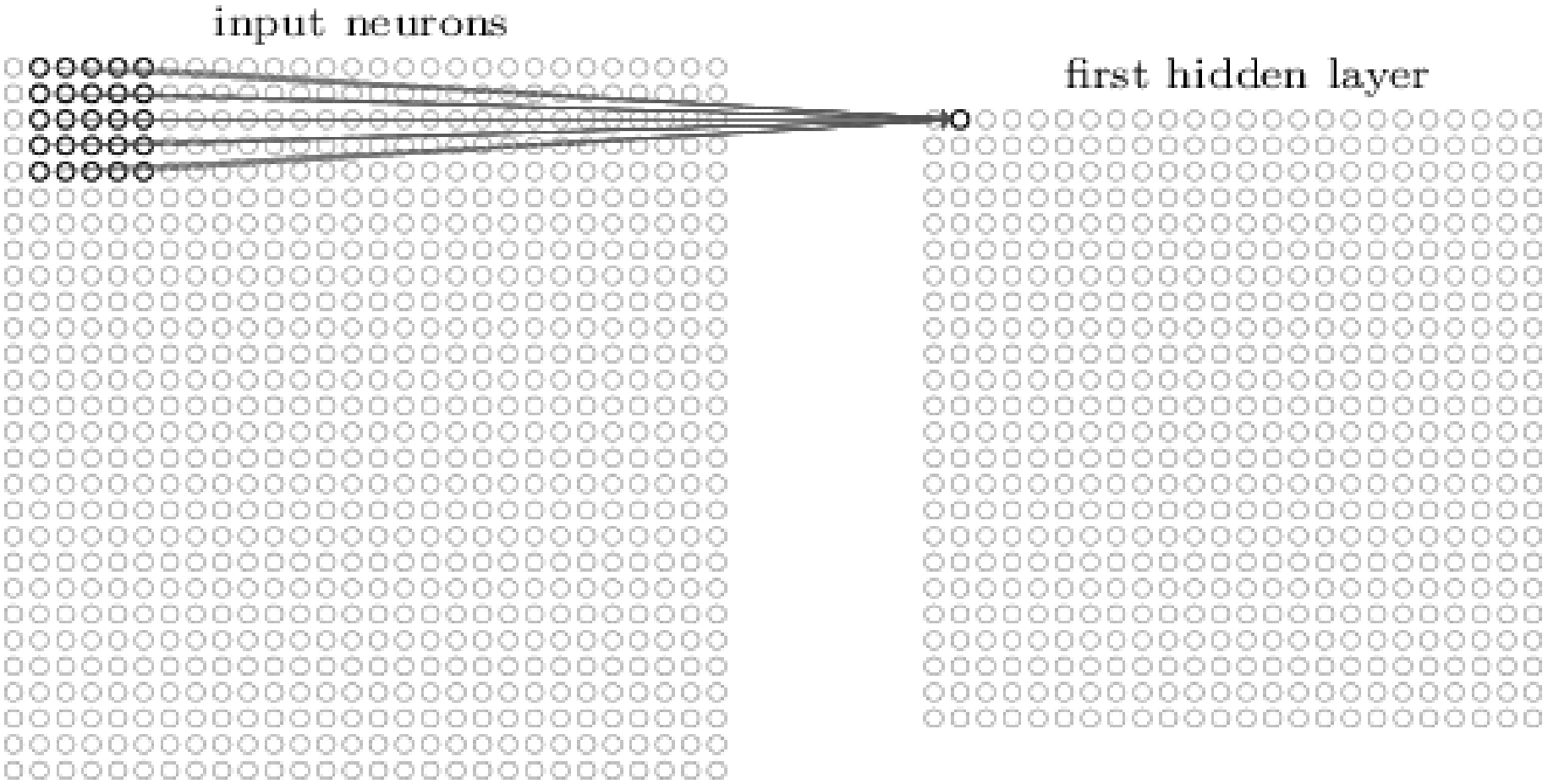
Source: https://cs231n.github.io/convolutional-networks

Convolutional Layer – single kernel



Source: http://neuralnetworksanddeeplearning.com

Convolutional Layer – single kernel

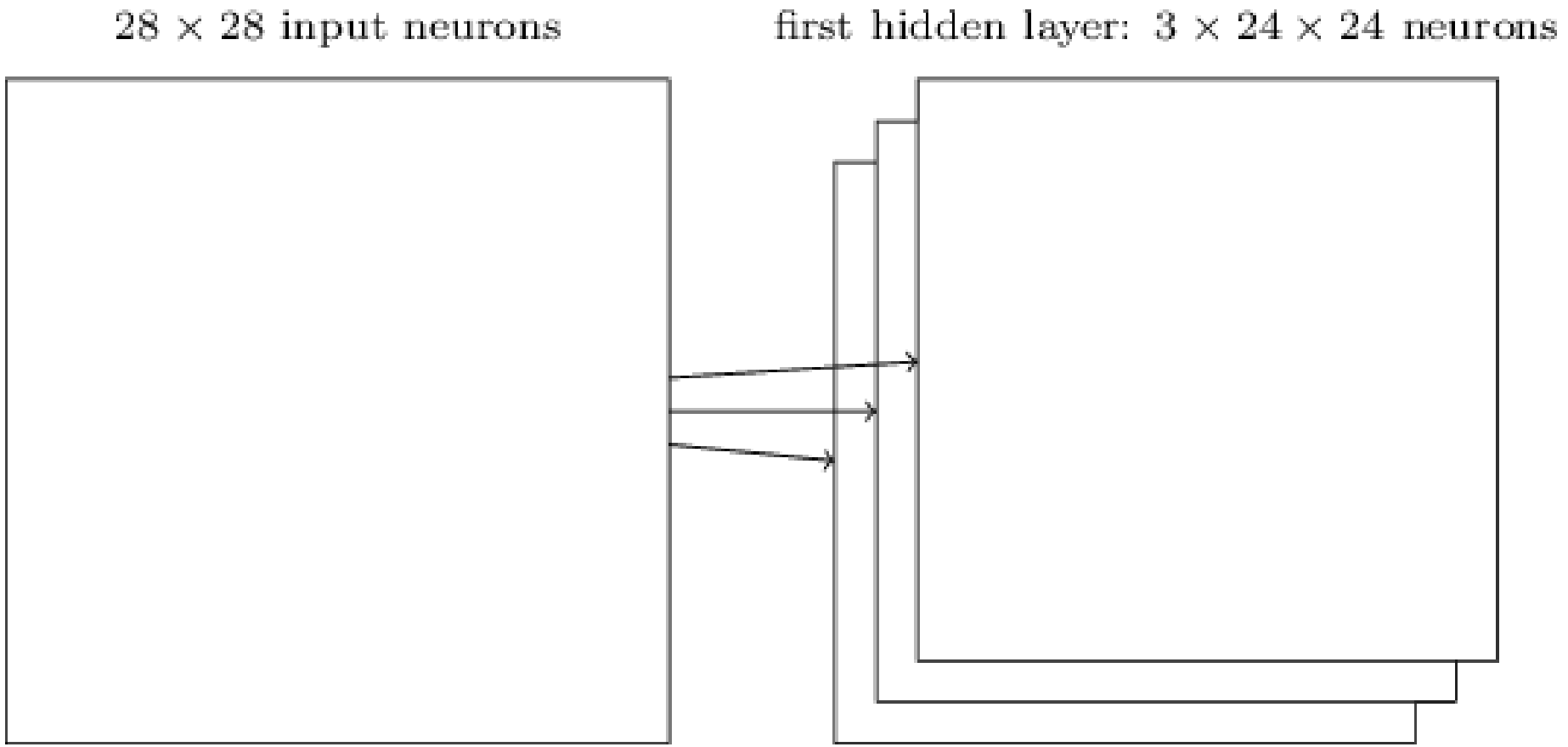


Source: http://neuralnetworksanddeeplearning.com

Convolutional Layer – Hyperparameters

- Parameter sharing
- Depth
- Stride
- Padding
- Dilation

Convolutional Layer – depth



Source: http://neuralnetworksanddeeplearning.com

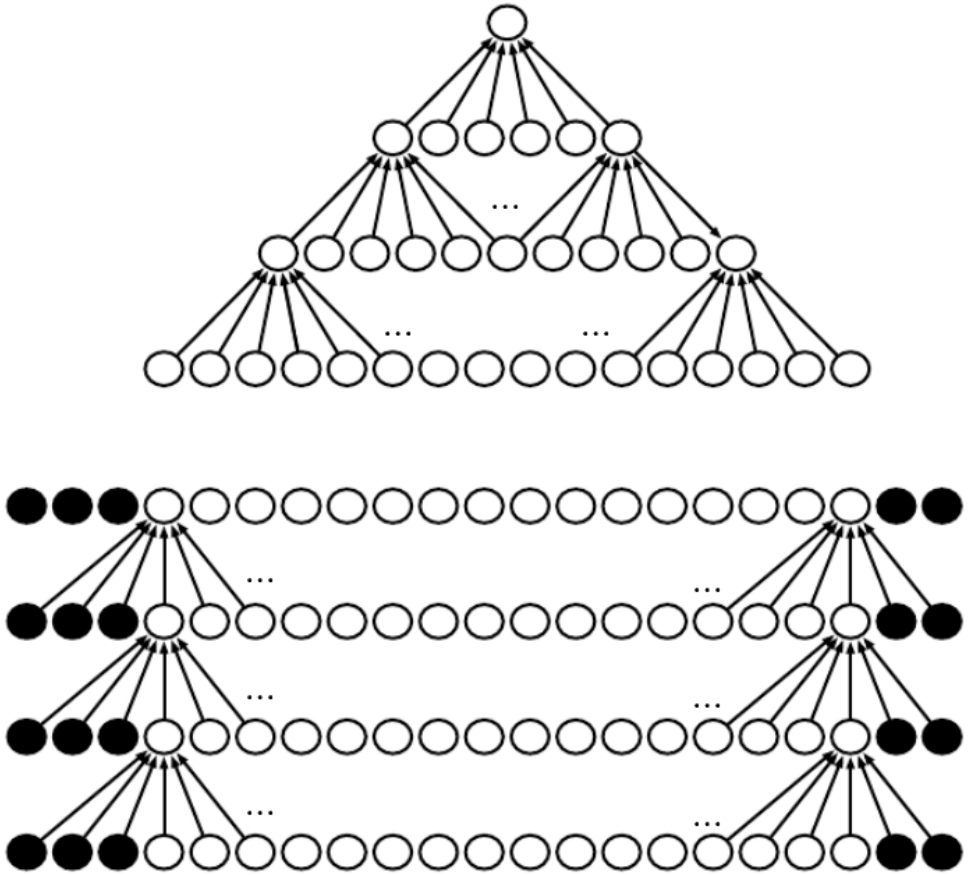
Convolutional Layer – Hyperparameters

- Parameter sharing
- Depth
- **Stride**
- Padding
- Dilation

Convolutional Layer – Hyperparameters

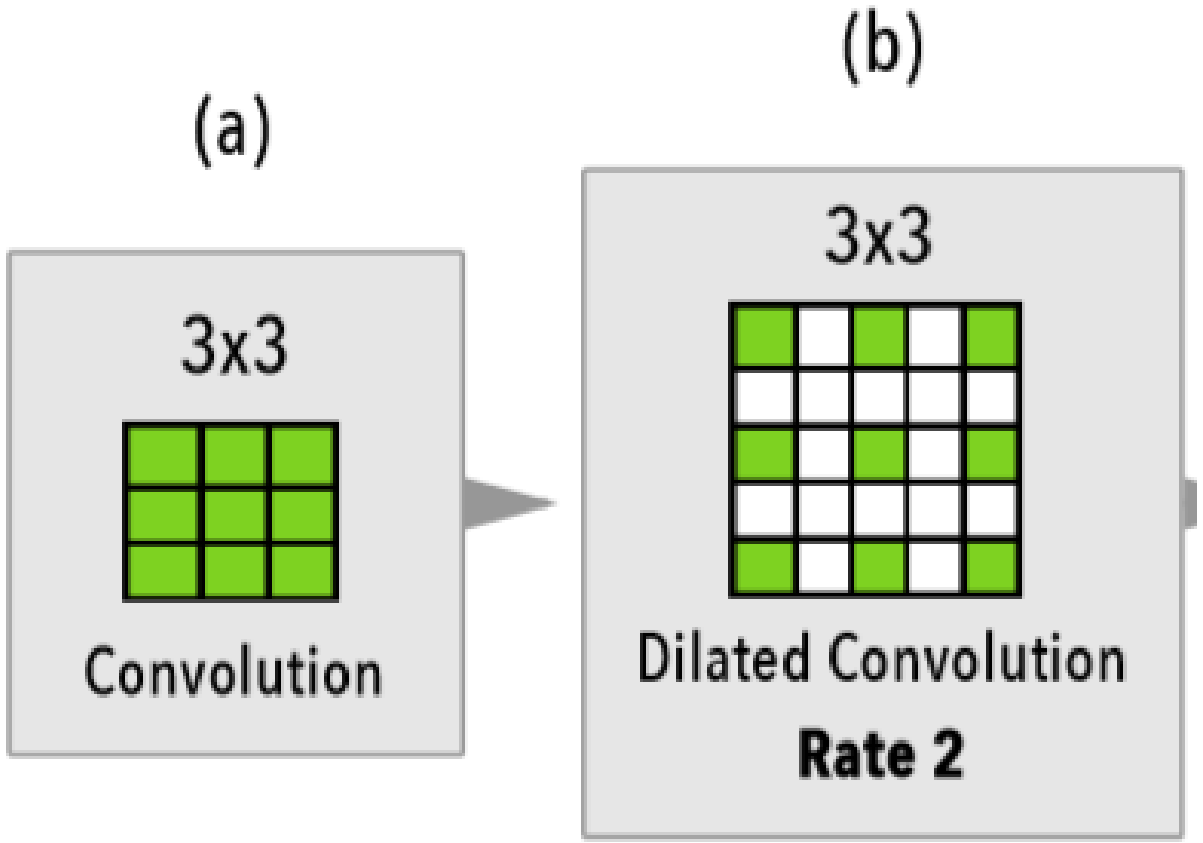
- Parameter sharing
- Depth
- Stride
- **Padding**
- Dilation

Convolutional Layer – padding



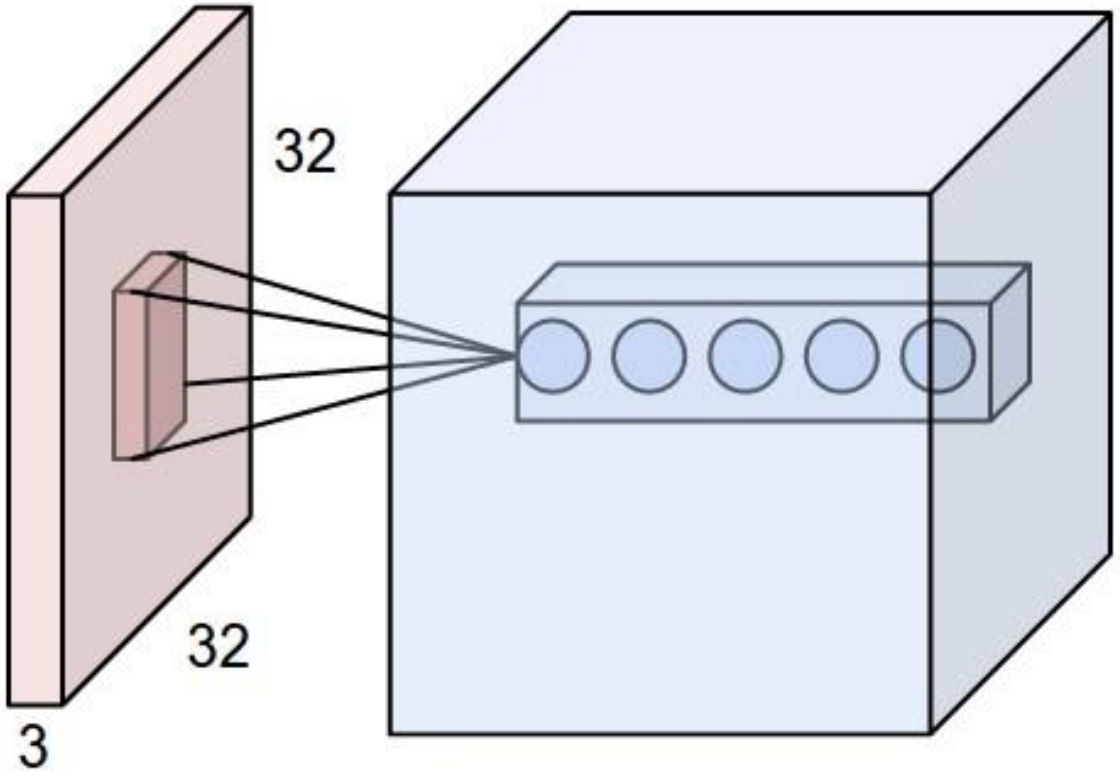
Source: Bengio Y. et al., „Deep Learning”

Convolutional Layer – dilated convolution



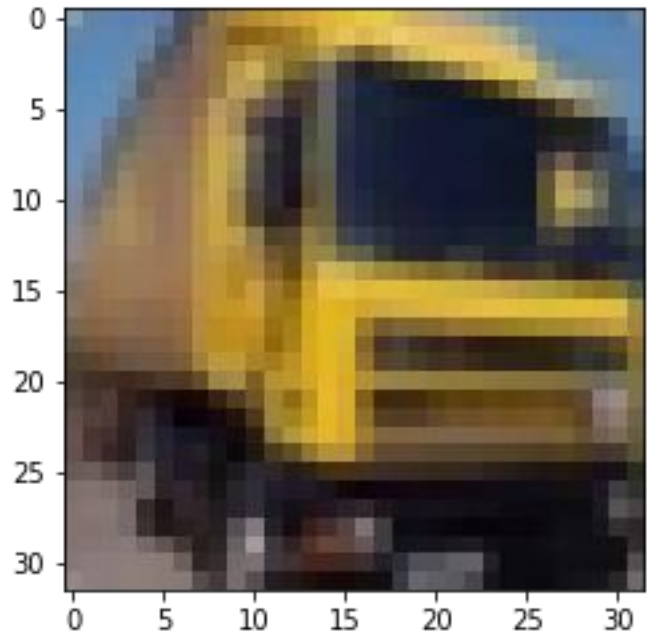
Source: https://www.researchgate.net/publication/320195101_Spinal_cord_gray_matter_segmentation_using_deep_dilated_convolutions

Convolutional Layer – final view

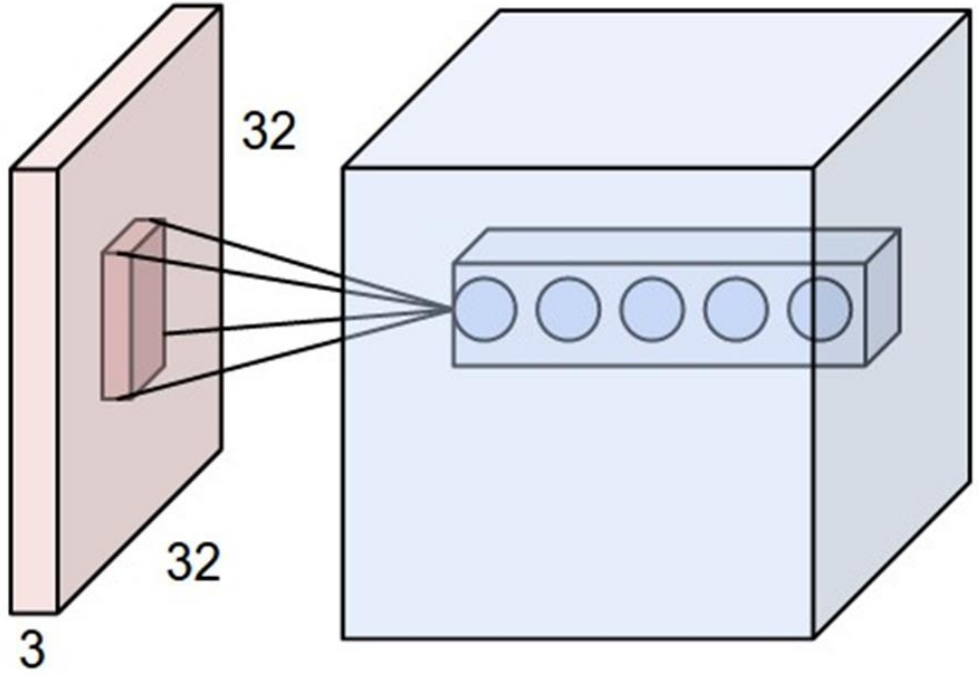


Source: https://cs231n.github.io/convolutional-networks

Convolutional Layer - example

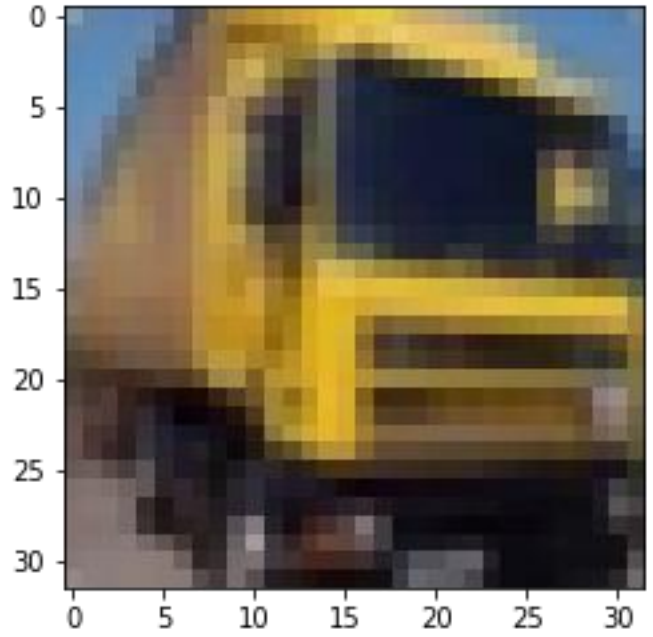


Source: CIFAR-10



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

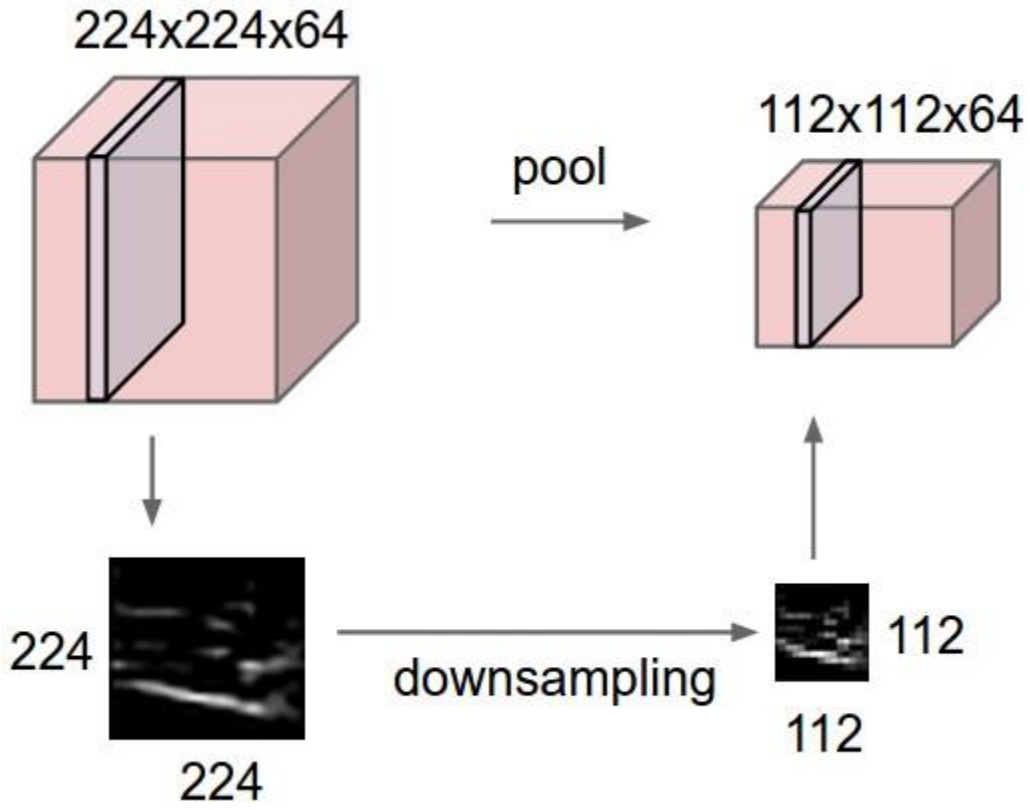
Convolutional Layer - example



Source: CIFAR-10

375 parameters

Pooling Layer

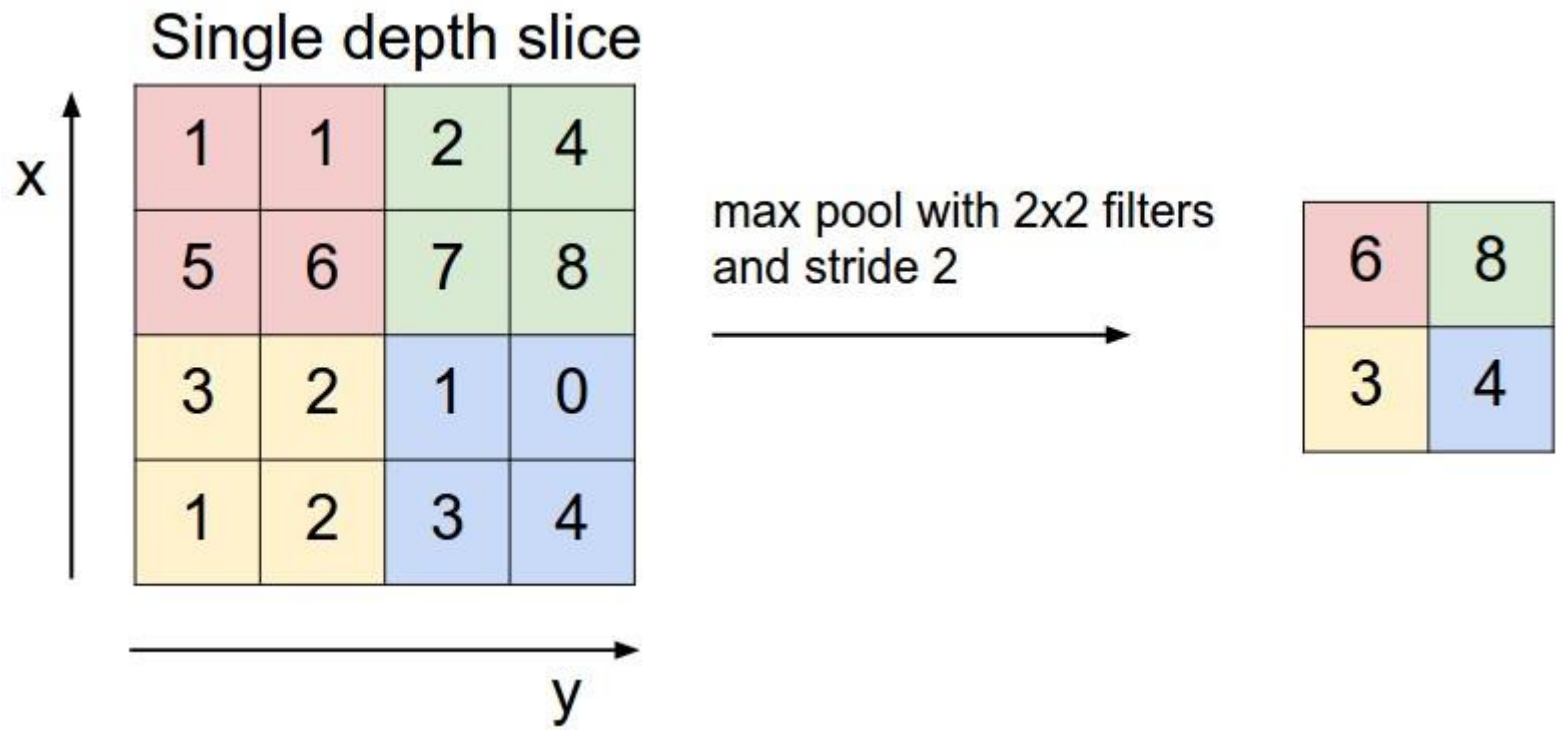


Source: https://cs231n.github.io/convolutional-networks

Pooling Layer - hyperparameters

- Spatial Extent
- Stride
- Type

Pooling Layer - example



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

Layer sizing

- Input many times divisible by 2 (e.g. 32, 64, 96, 224, 384, 512…)

Layer sizing

- Input many times divisible by 2 (e.g. 32, 64, 96, 224, 384, 512…)
- Kernel should have small size (3x3, 5x5), stride = 1 and zero-padding

Layer sizing

- Input many times divisible by 2 (e.g. 32, 64, 96, 224, 384, 512…)
- Kernel should have small size (3x3, 5x5), stride = 1 and zero-padding
- Conv Layers should preserve spatial dimensions of input

Layer sizing

- Input many times divisible by 2 (e.g. 32, 64, 96, 224, 384, 512…)
- Kernel should have small size (3x3, 5x5), stride = 1 and zero-padding
- Conv Layers should preserve spatial dimensions of input
- Max-pooling layer should be 2x2 with stride = 2

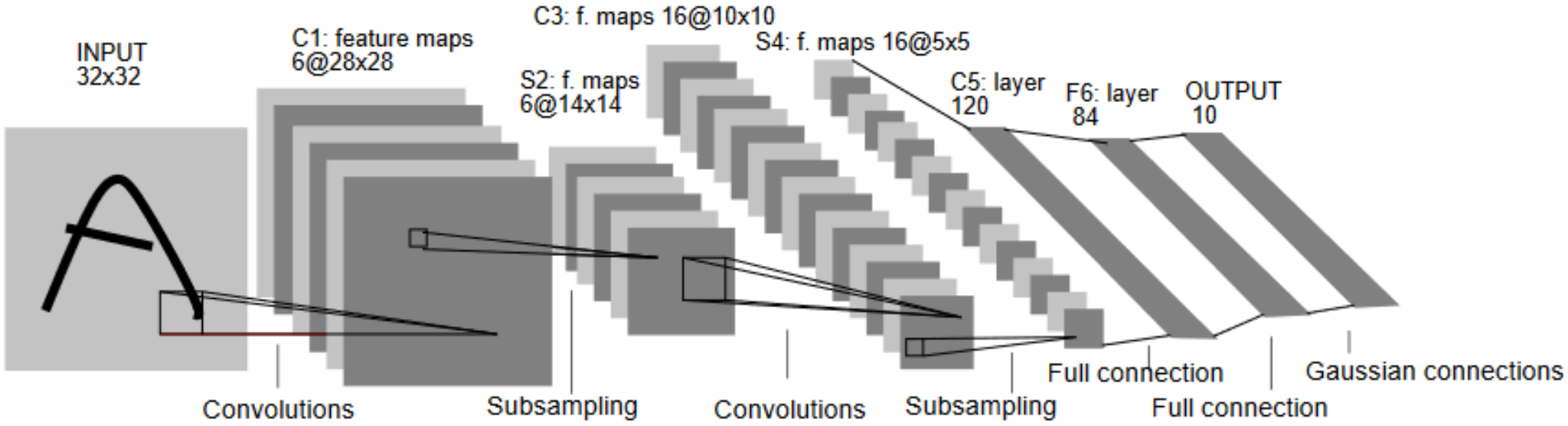
Layer sizing

- Input many times divisible by 2 (e.g. 32, 64, 96, 224, 384, 512…)
- Kernel should have small size (3x3, 5x5), stride = 1 and zero-padding
- Conv Layers should preserve spatial dimensions of input
- Max-pooling layer should be 2x2 with stride = 2
- Be cautious when resizing in Conv Layer

Important CNN architectures

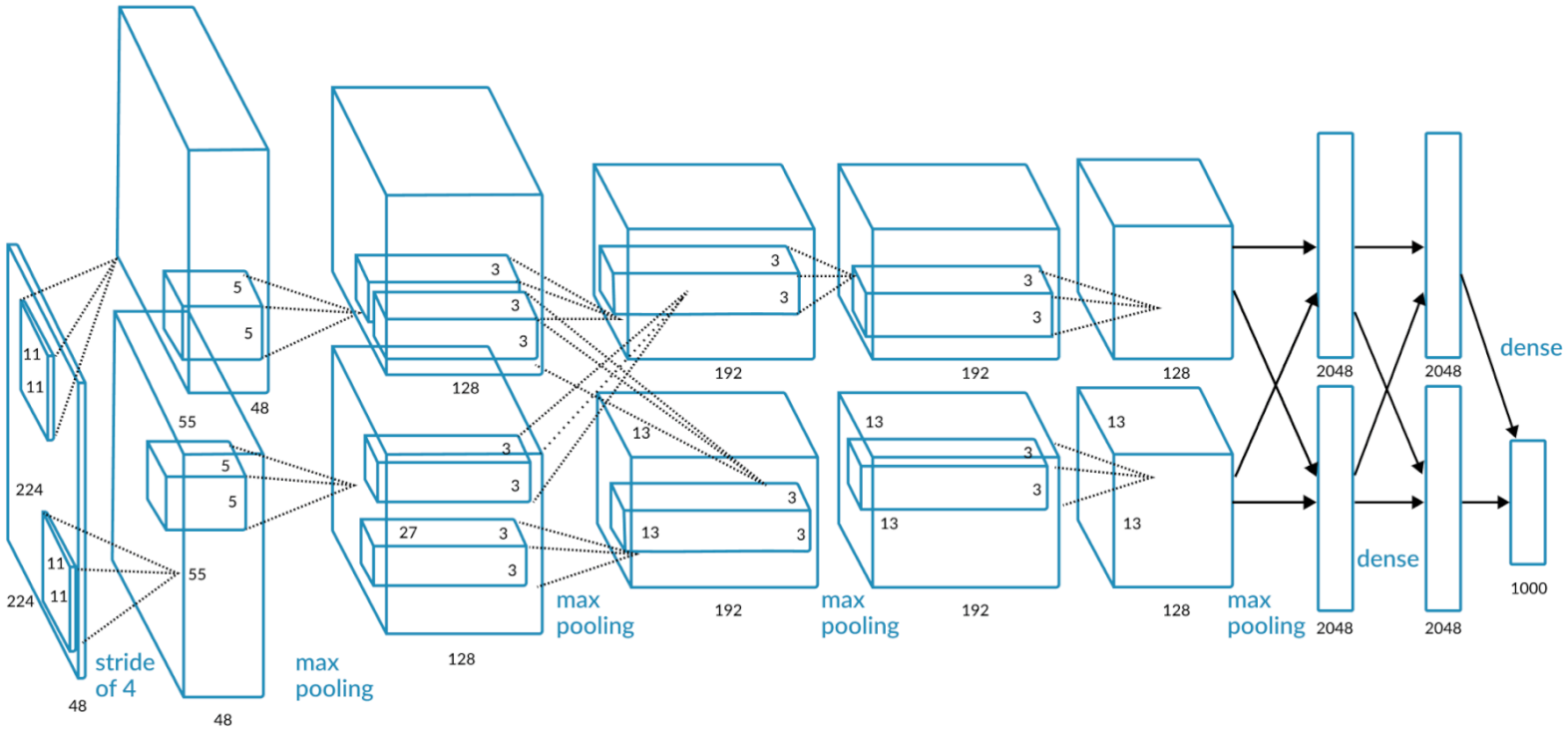
- Basic – LeNet (1990)
- AlexNet (2012)
- GoogleNet (2014)
- ResNet (2015)
- MobileNet (2017)

LeNet (2012)



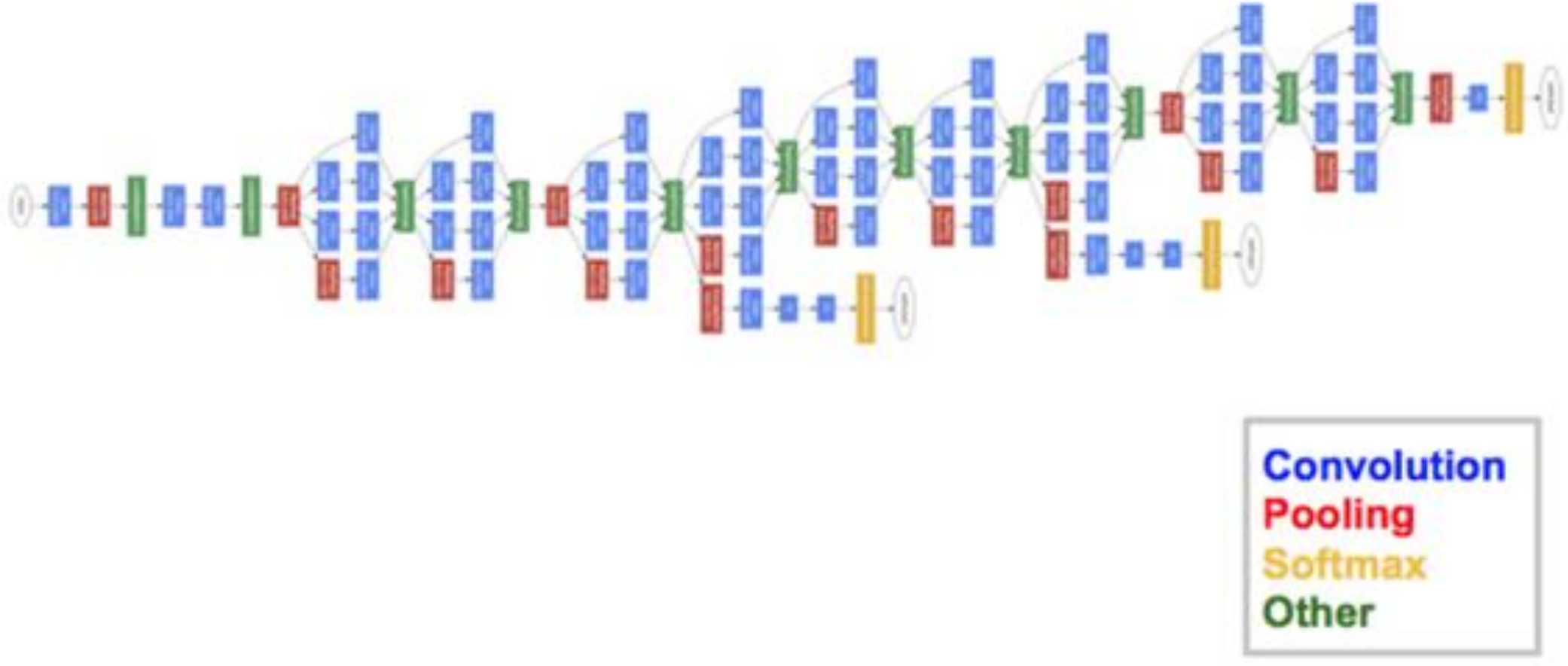
Source: LeCun Y., et al. „Gradient-Based Learning Applied to Document Recognition"

AlexNet (2012)



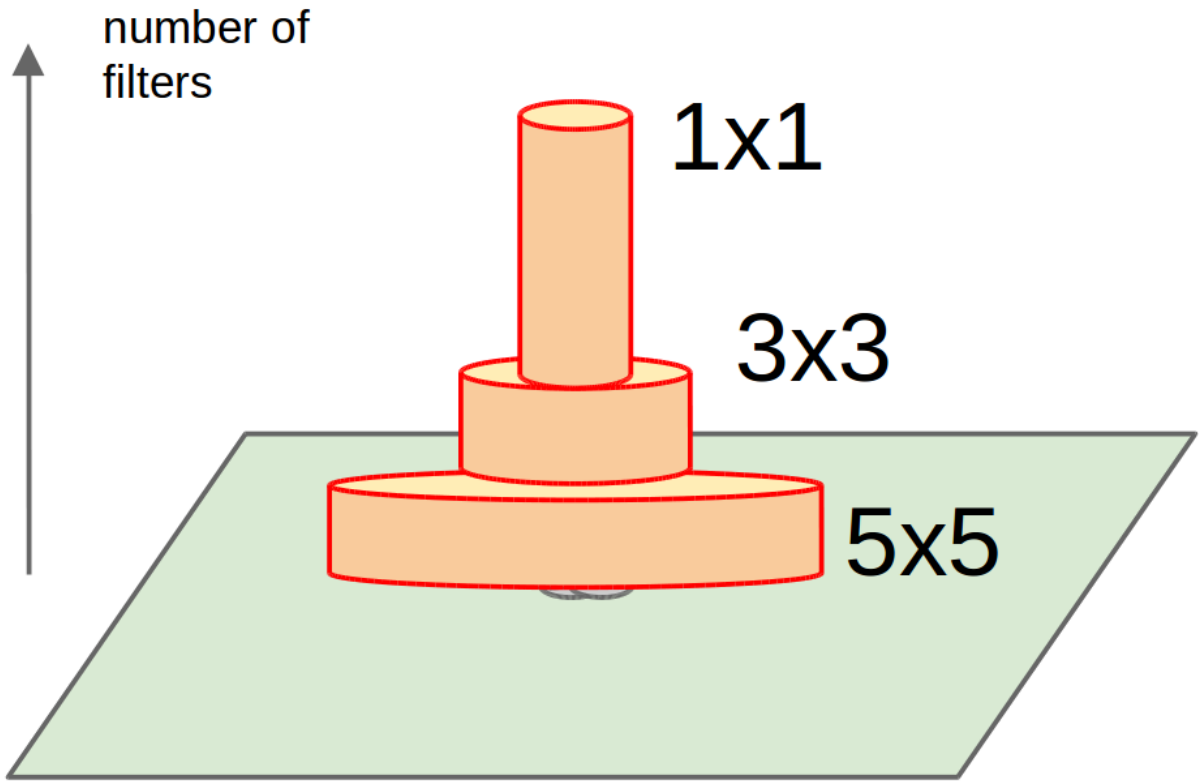
Source: https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future/

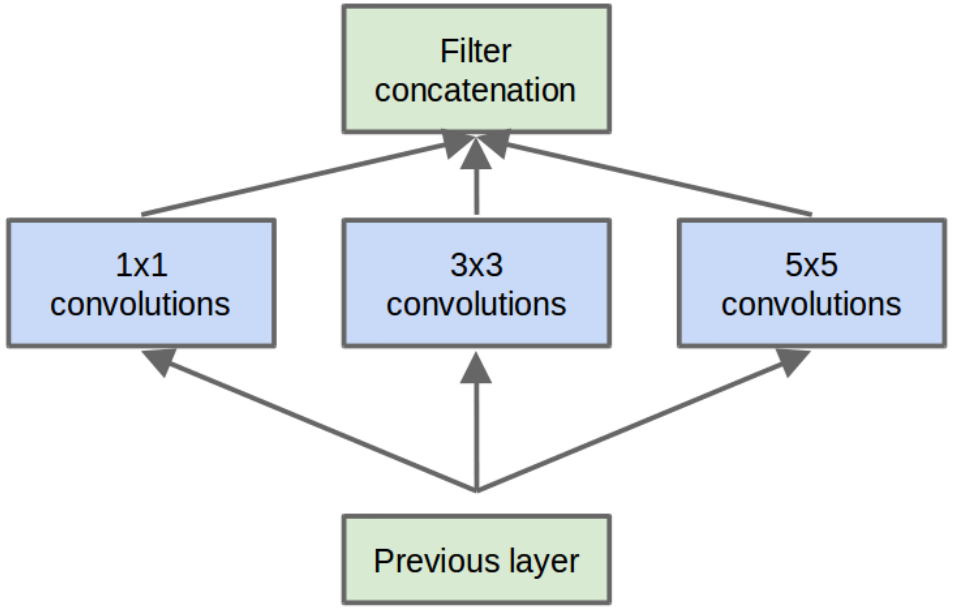
GoogleNet/InceptionNet (2014)



Source: https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future/

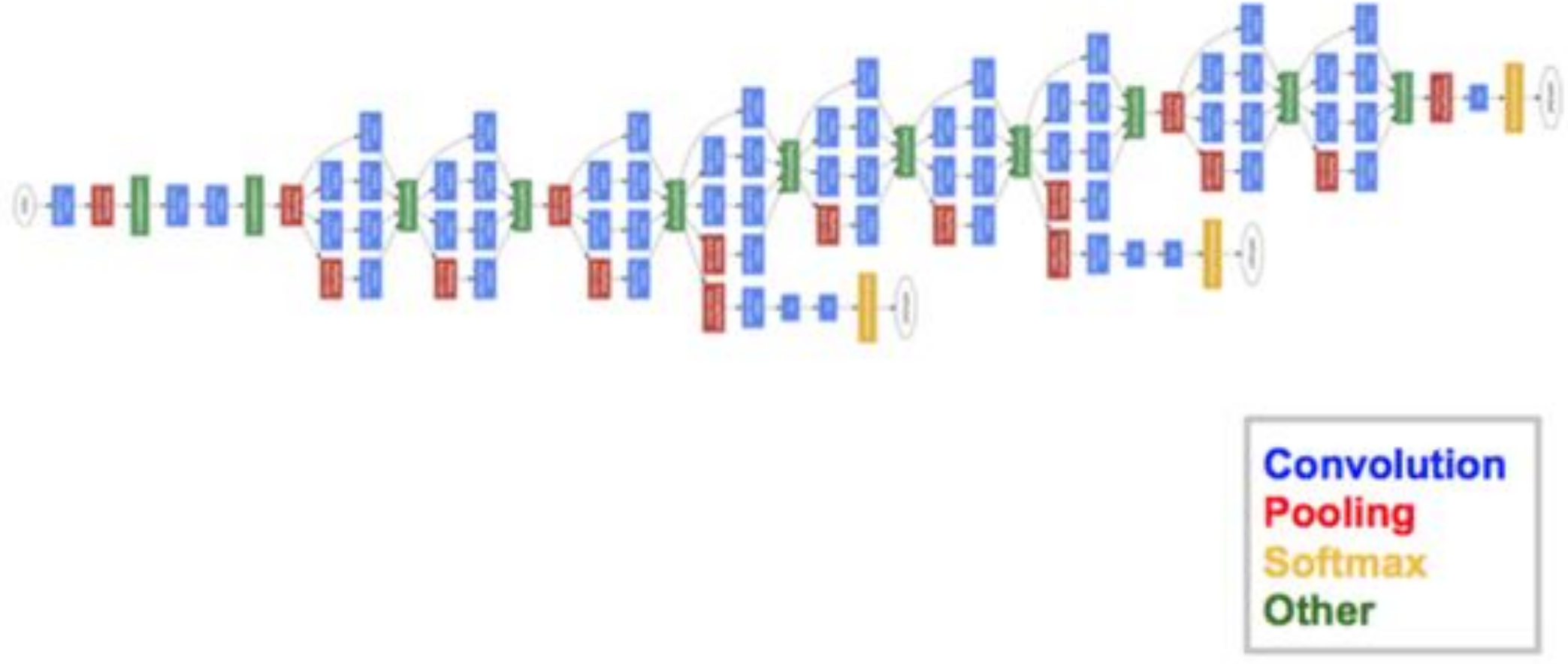
GoogleNet (2014) – inception layer





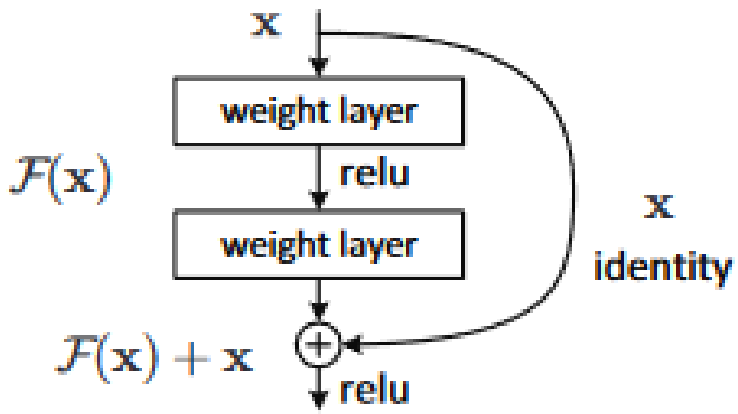
Source: https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/googlenet.html

GoogleNet (2014) – vanishing gradient



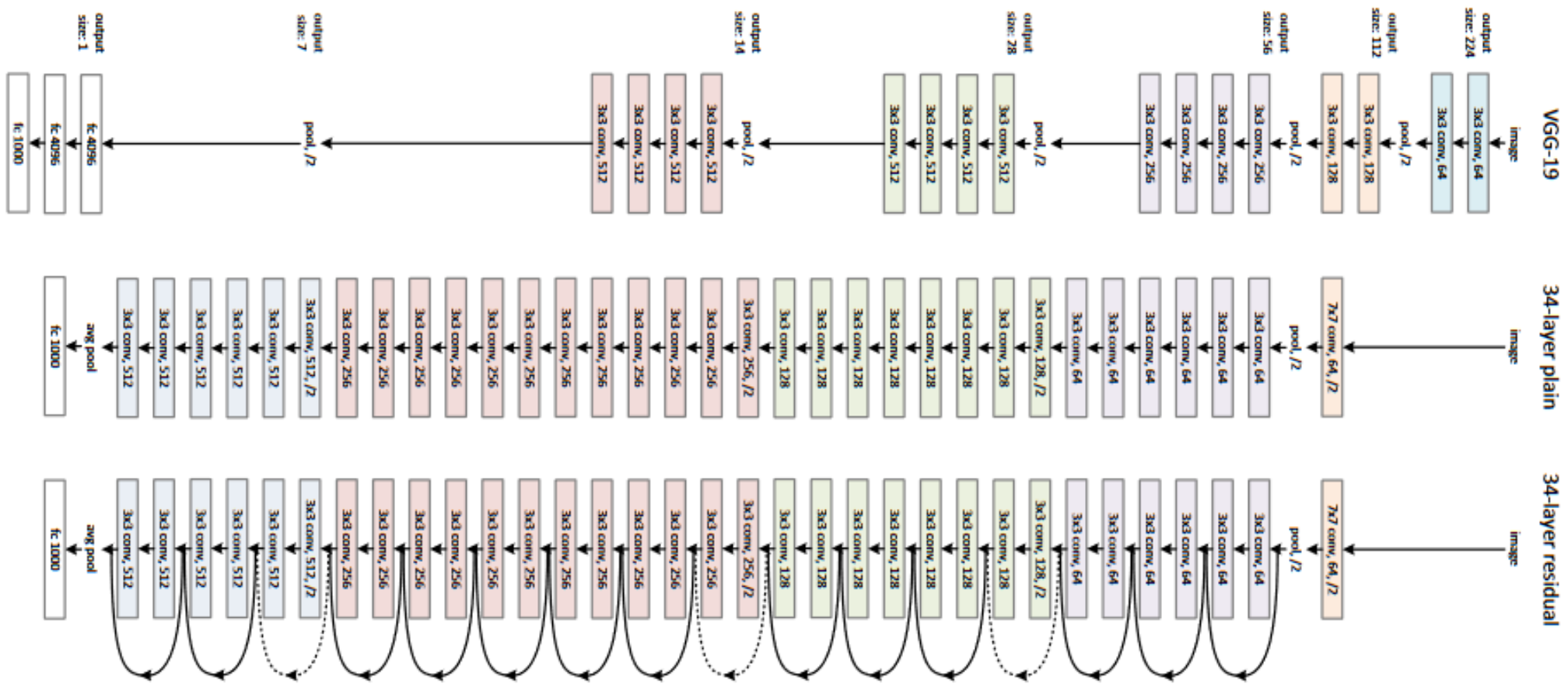
Source: https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future/

ResNet (2015) – shortcut connection



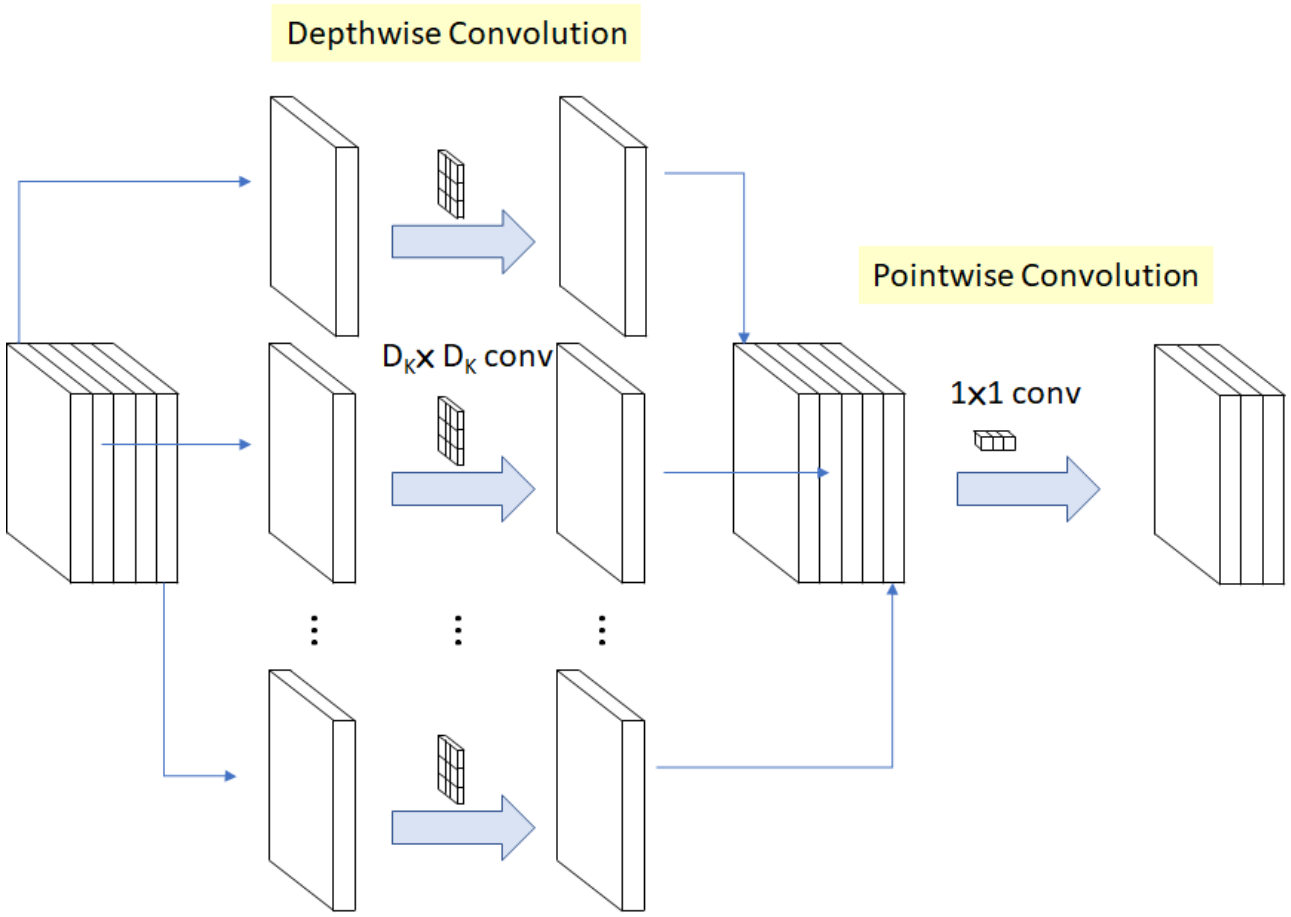
Source: He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

ResNet (2015) – shortcut connection



Source: He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

MobileNet (2017) – Depthwise Separable Convolution



Source: https://towardsdatascience.com/review-mobilenetv1-depthwise-separable-convolution-light-weight-model-a382df364b69

MobileNet (2017)

Model	ImageNet Accuracy	Million Mult-Adds	Million Parameters
Conv MobileNet	71.7%	4866	29.3
MobileNet	70.6%	569	4.2

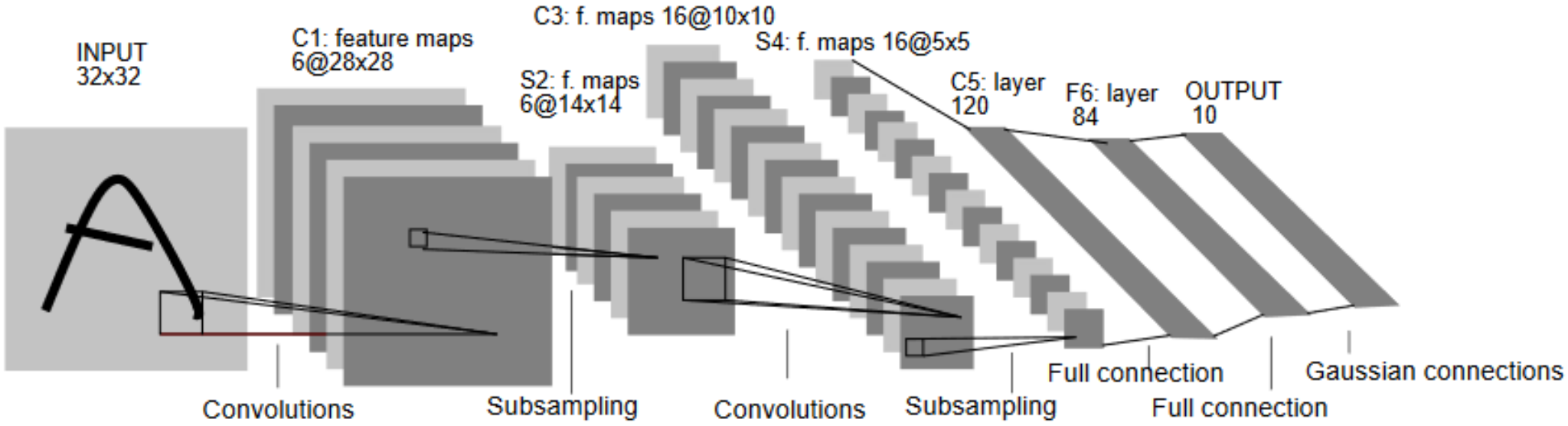
Source: Howard A. G., et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications."

Convolutional Neural Networks Summary

Questions?

Coding

LeNet (2012)



Source: LeCun Y., et al. „Gradient-Based Learning Applied to Document Recognition"

Questions?

Thanks for your attention

Piotr Mazur

piotrmmazur@outlook.com

kontakt@piomazur.pl