MLCon
CONFERENCE & TRAINING

Image Recognition with OpenCV and TensorFlow Workshop

Oliver Zeigermann & Tim Wüllner
OPEN KNOWLEDGE Gmbh

https://bit.ly/mlcon22_image_rec

MLCon
CONFERENCE & TRAINING

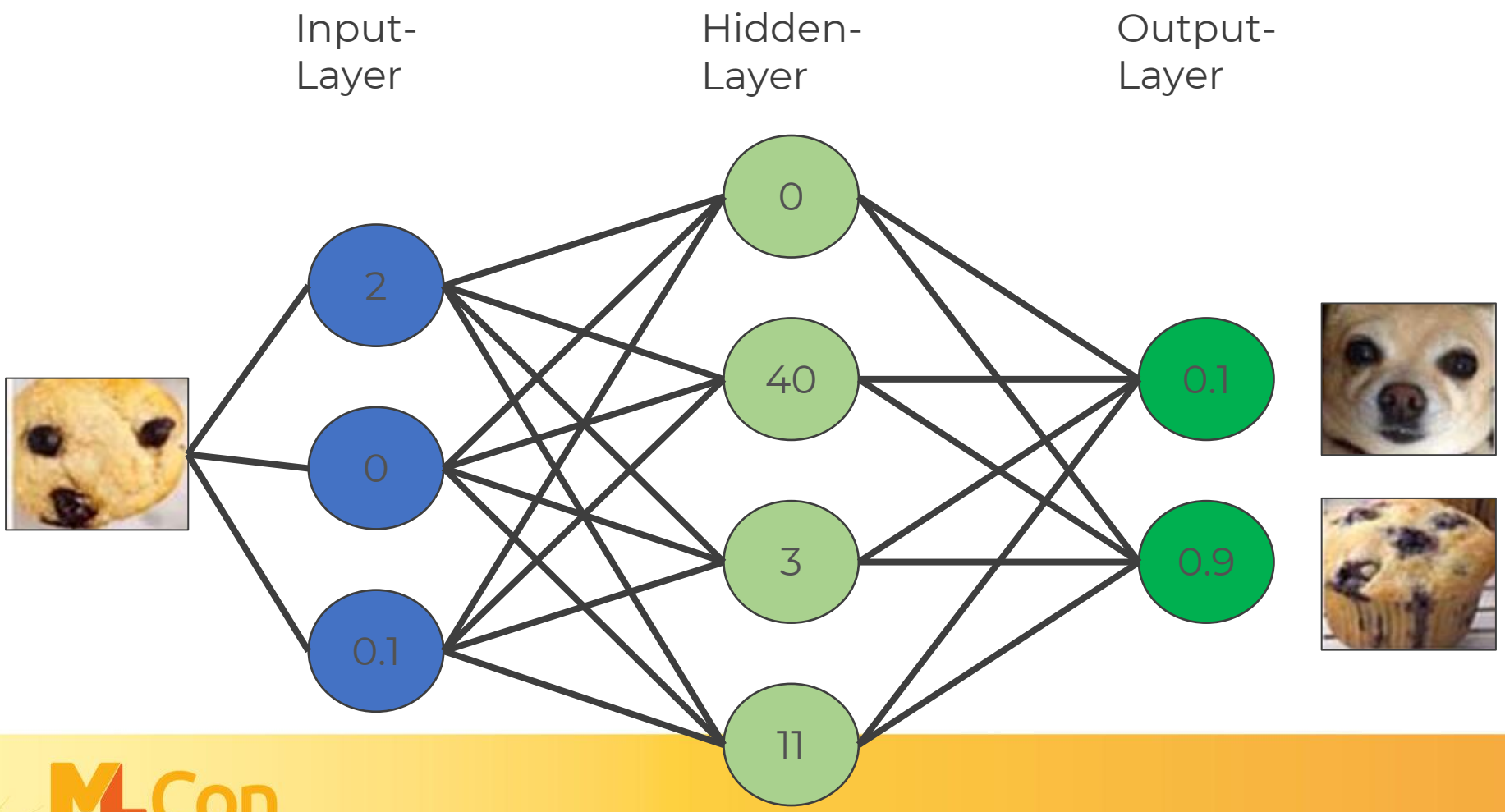
Theoretical Background

Neural Network

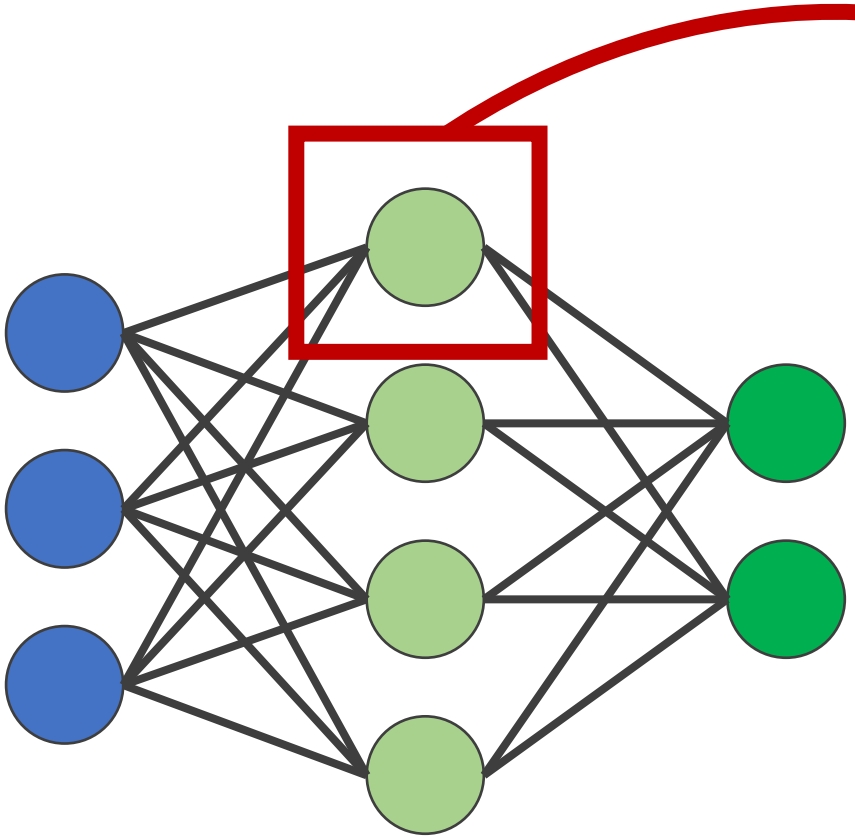
Neurons

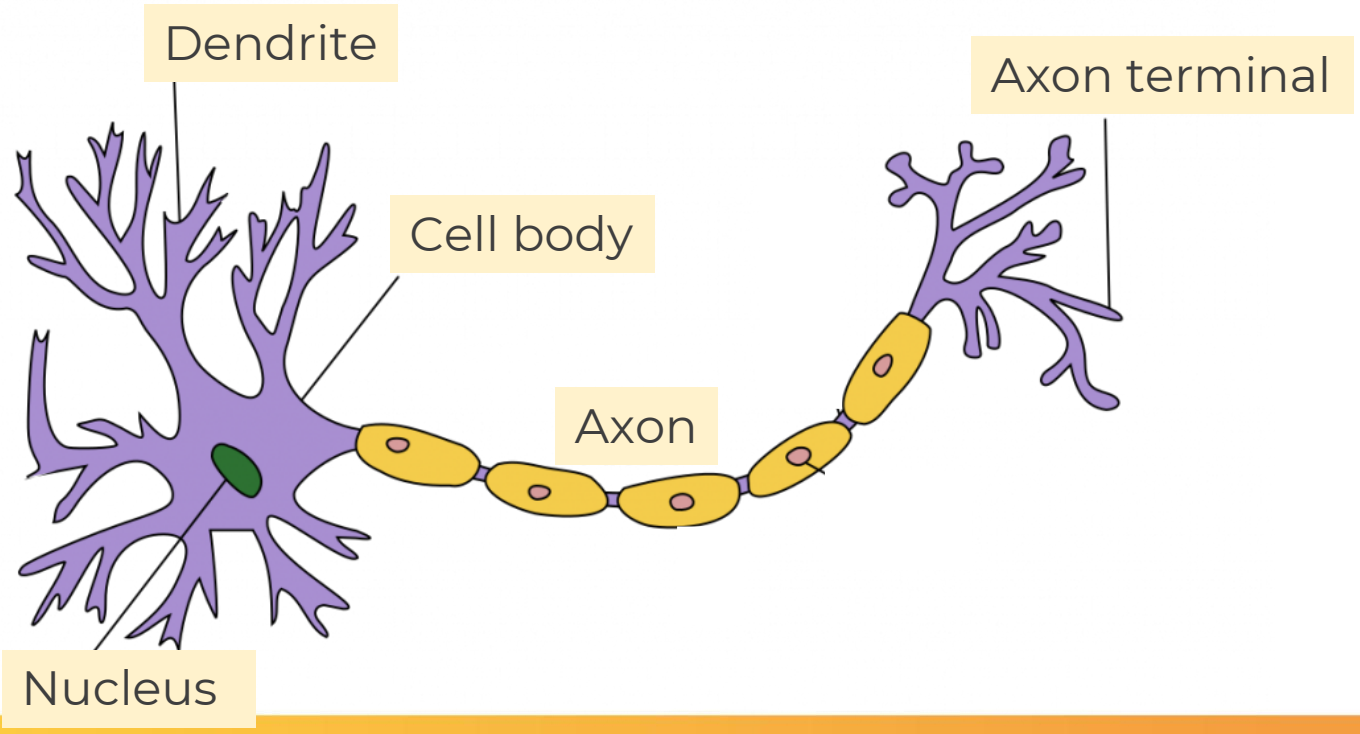
Activation functions

Neural Network

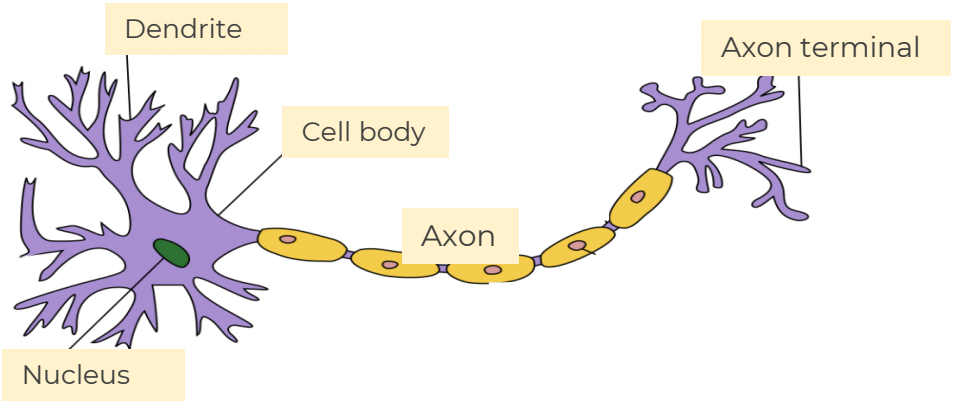


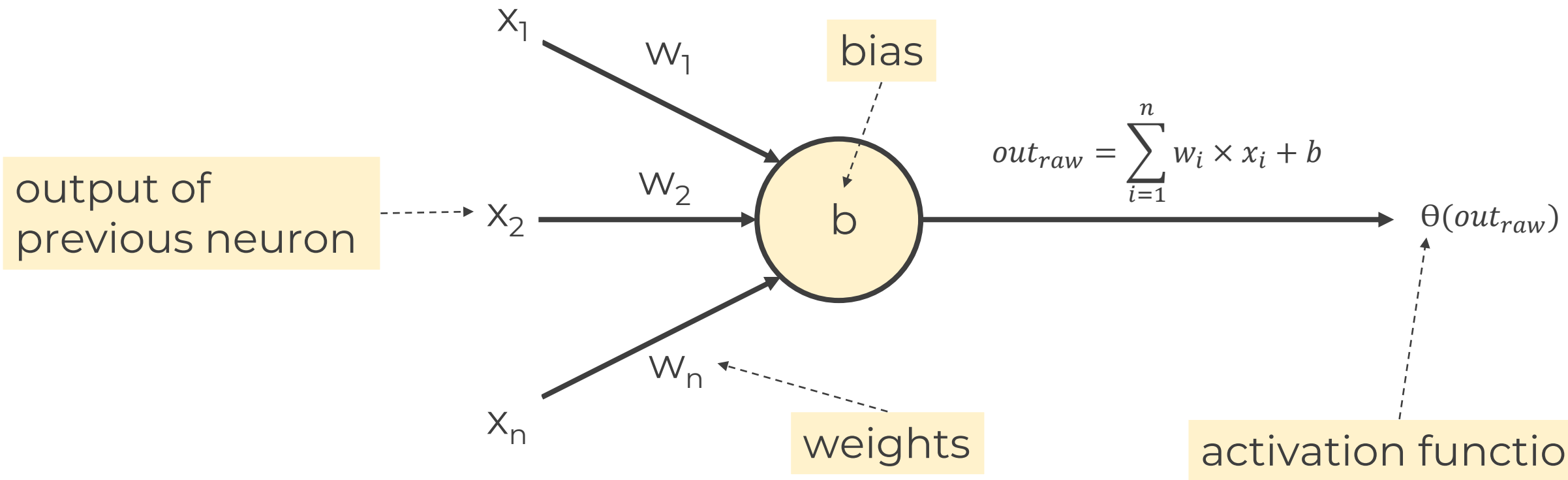
What‘s a neuron?





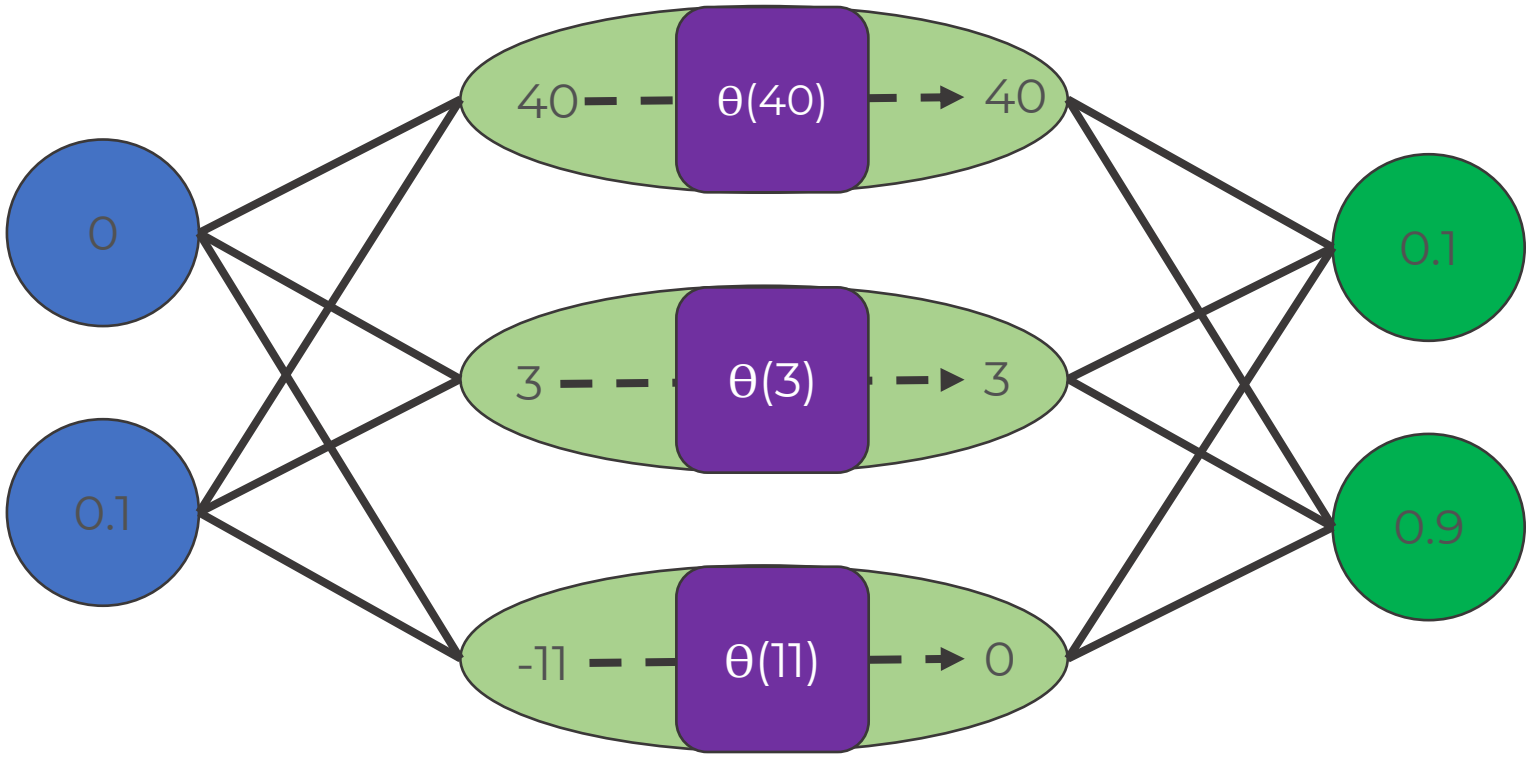
What‘s a neuron?



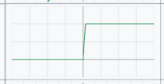
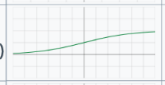
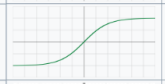
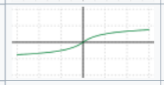


Activation function (θ)

- Subsequent **modification** of the neuron output
- **Layer**-specific (not neuron-specific)
- Adds **non-linear** properties to the network
- Recognition of **complex** patterns

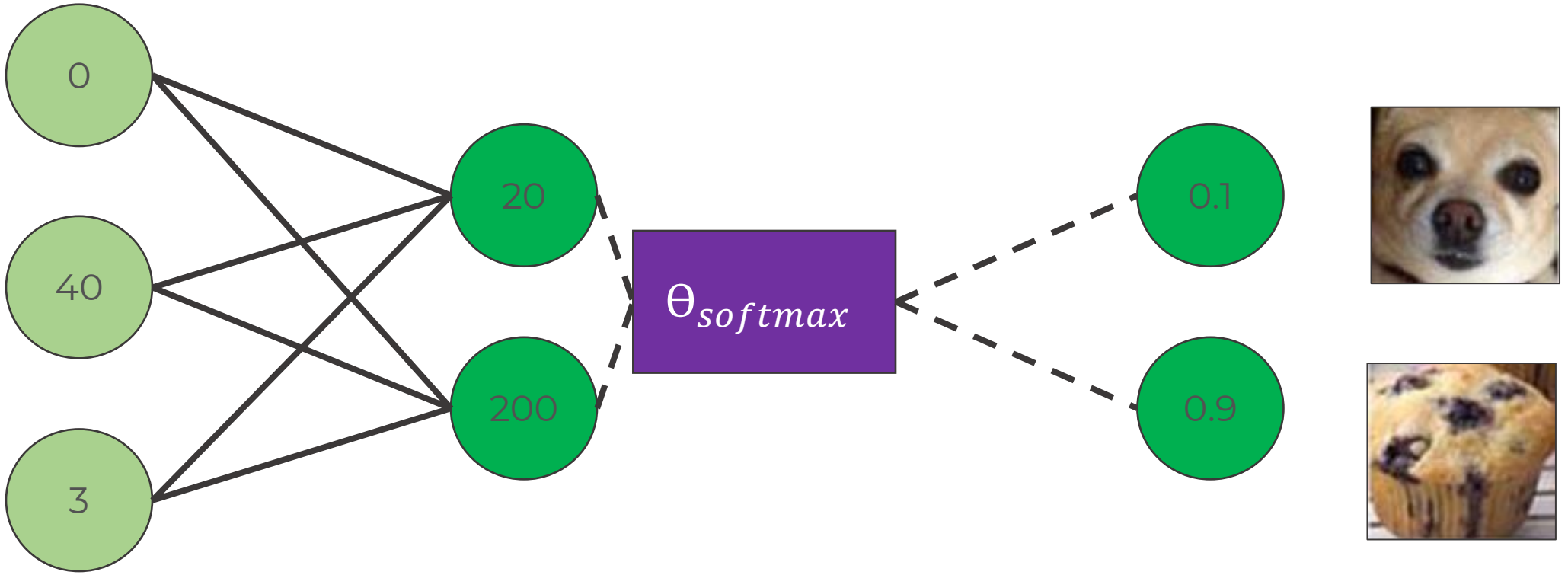


Activation function (θ)

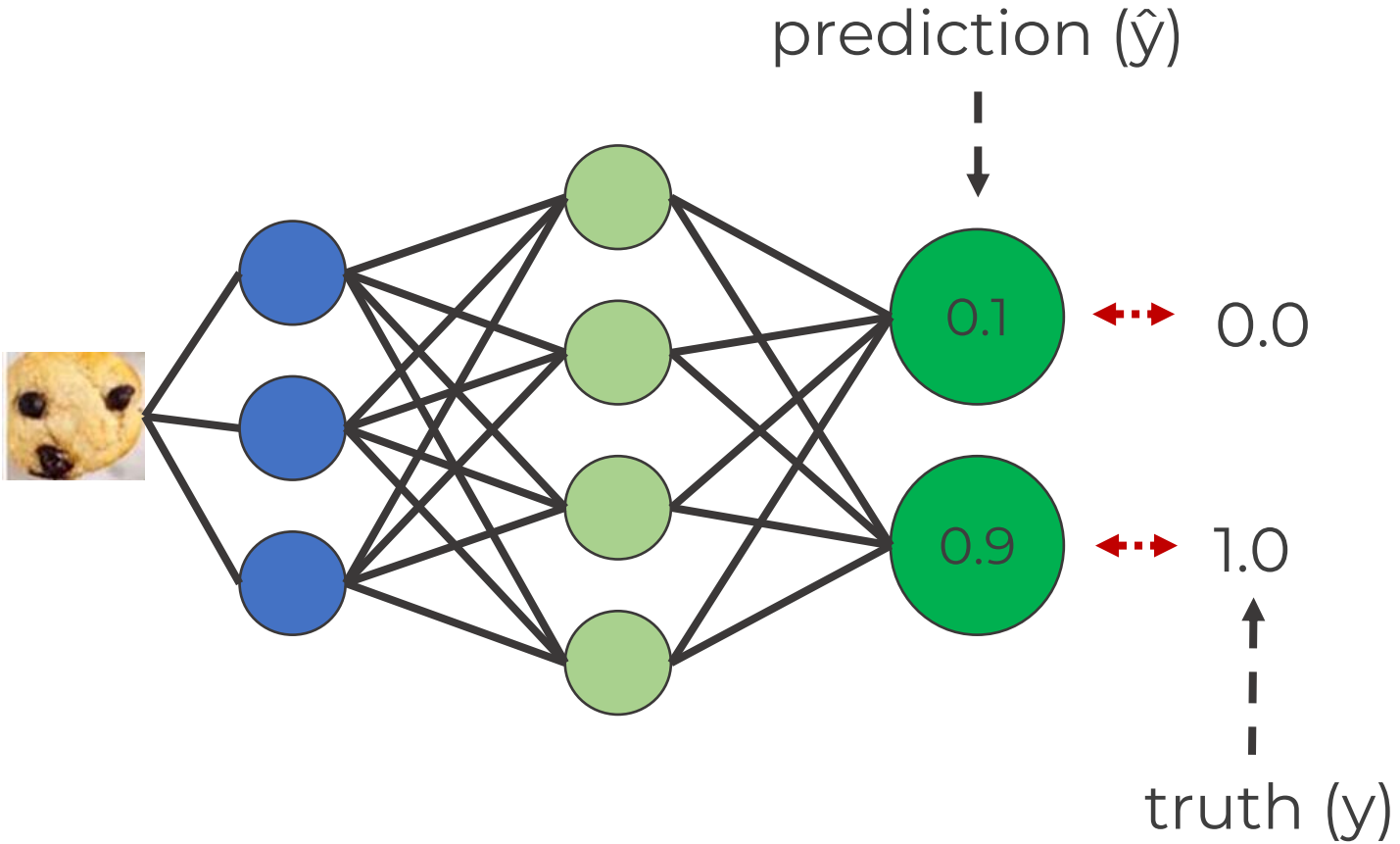
Name	Plot	Equation	Derivative	Range	Order of continuity
Identity		$f(x) = x$	$f'(x) = 1$	$(-\infty, \infty)$	C^∞
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$	$\{0, 1\}$	C^{-1}
Logistic (a.k.a. Sigmoid or Soft step)		$f(x) = \sigma(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1 - f(x))$	$(0, 1)$	C^∞
TanH		$f(x) = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$	$f'(x) = 1 - f(x)^2$	$(-1, 1)$	C^∞
ElliotSig Softsign		$f(x) = \frac{x}{1+\lvert x\rvert}$	$f'(x) = \frac{1}{(1+\lvert x\rvert)^2}$	$(-1, 1)$	C^1
Square Nonlinearity (SQNL)		$f(x) = \begin{cases} 1 & : x > 2.0 \\ x - \frac{x^2}{4} & : 0 \leq x \leq 2.0 \\ x + \frac{x^2}{4} & : -2.0 \leq x < 0 \\ -1 & : x < -2.0 \end{cases}$	$f'(x) = 1 \mp \frac{x}{2}$	$(-1, 1)$	C^2
Rectified linear unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$[0, \infty)$	C^0
Bipolar rectified linear unit (BReLU)		$f(x_i) = \begin{cases} ReLU(x_i) & \text{if } i \bmod 2 = 0 \\ -ReLU(-x_i) & \text{if } i \bmod 2 \neq 0 \end{cases}$	$f'(x_i) = \begin{cases} ReLU'(x_i) & \text{if } i \bmod 2 = 0 \\ -ReLU'(-x_i) & \text{if } i \bmod 2 \neq 0 \end{cases}$	$(-\infty, \infty)$	C^0
Leaky rectified linear unit (Leaky ReLU)		$f(x) = \begin{cases} 0.01x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0.01 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$(-\infty, \infty)$	C^0
Parameteric rectified linear unit (PReLU)		$f(\alpha, x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(\alpha, x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$(-\infty, \infty)$	C^0

[...]

Softmax activation function

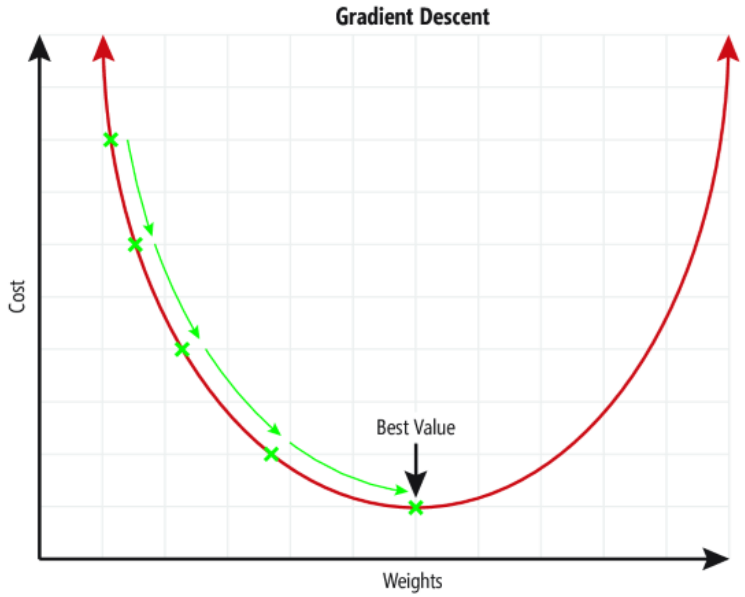


How does a neural network learn?

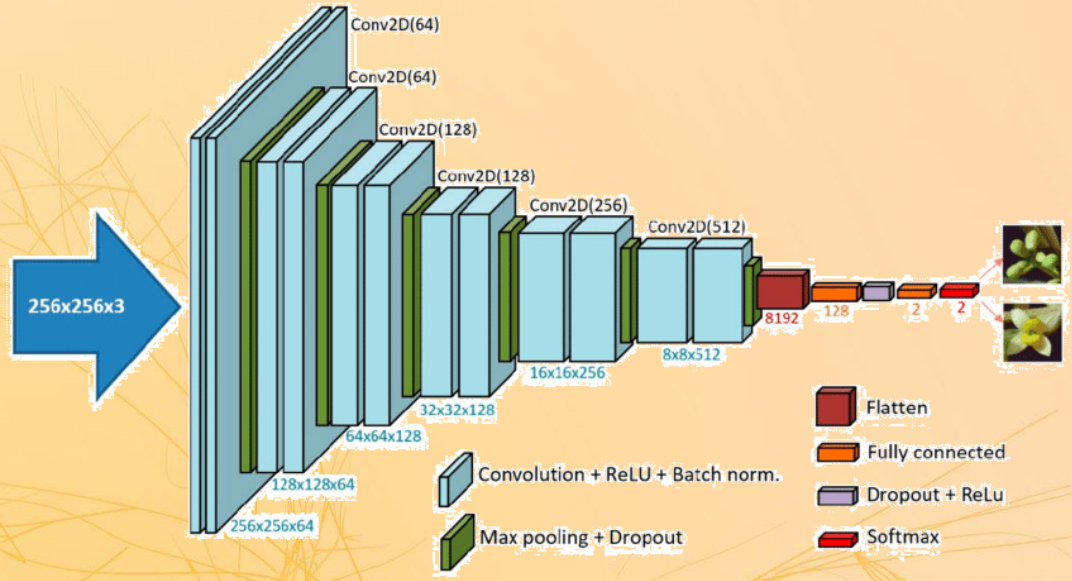


cost function:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$



MLCon
CONFERENCE & TRAINING



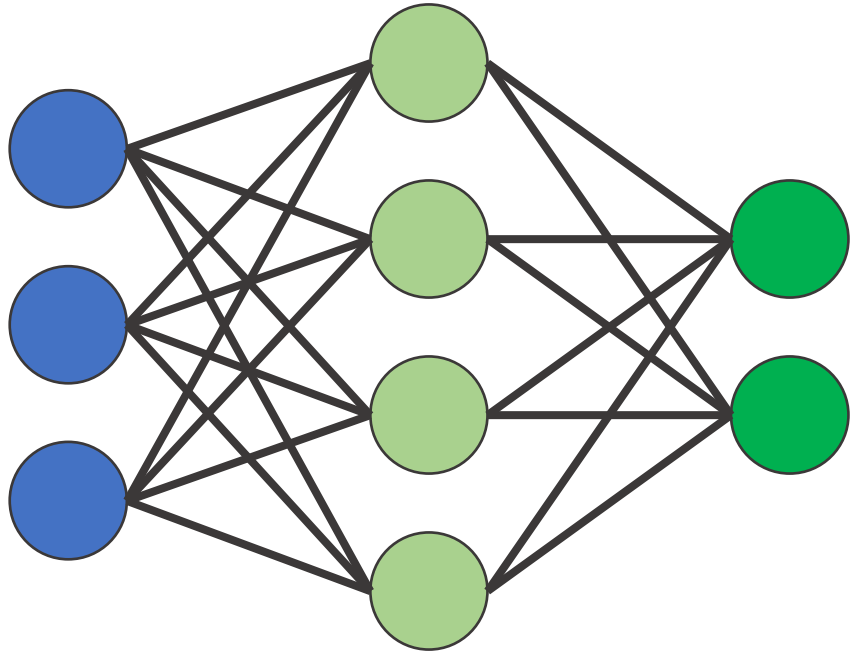
Types of Layers

Dense

Convolutional

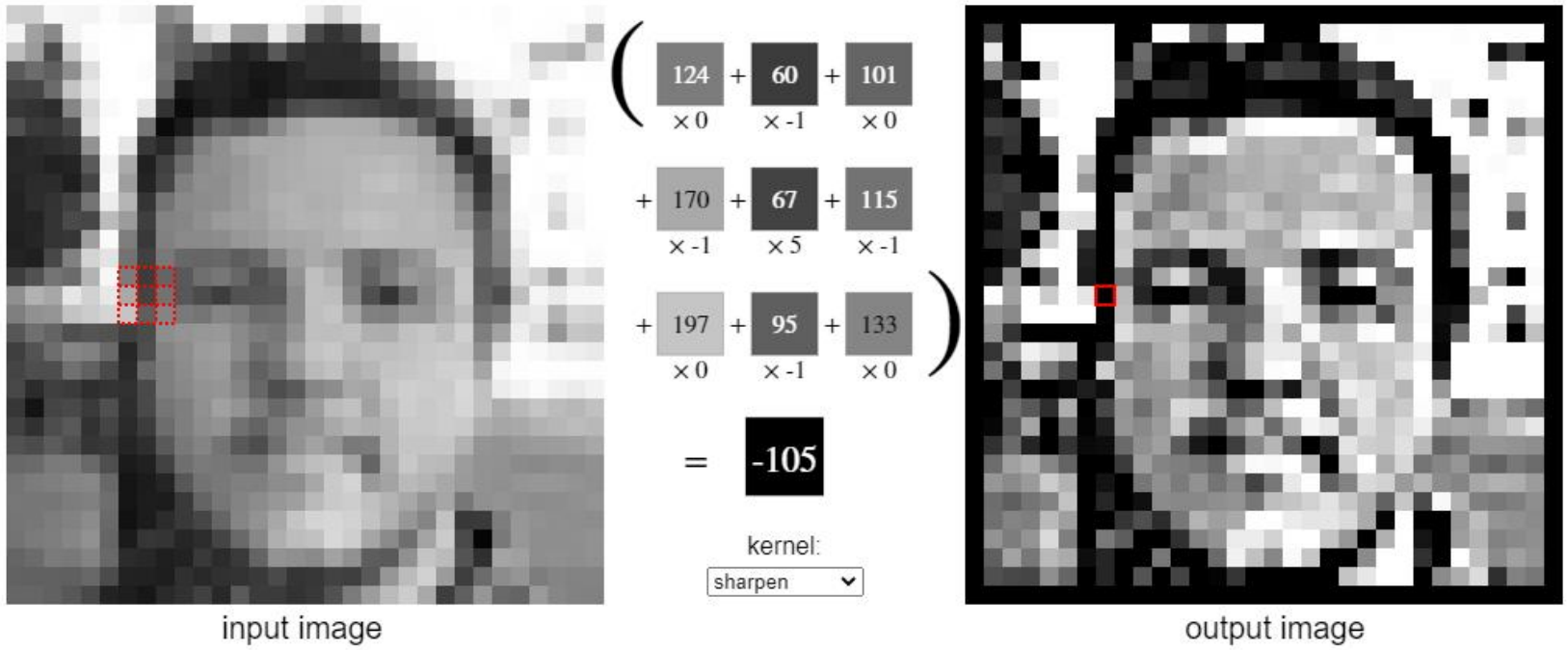
Dense Layer

- Each neuron is connected to each neuron from the previous layer
- High computational effort
- CNN: Output layer



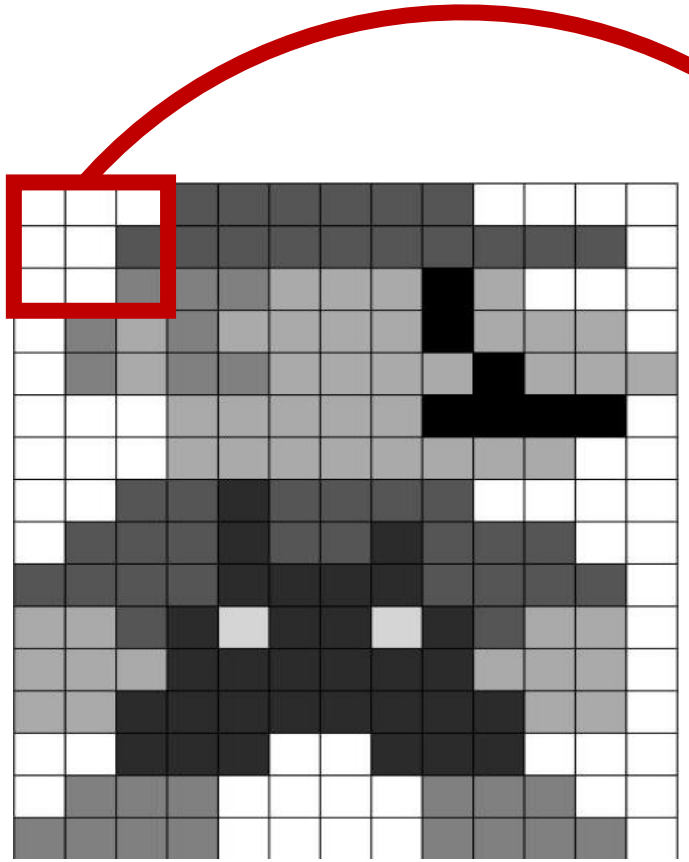
Convolutional Layer

- Defines CNNs
- Extraction of features
- Reduction of computational effort
- Mathematical convolution



by Victor Powell
(https://setosa.io/ev/image-kernels/

Convolutional Layer



Kernel
(trainable weights)

1	1	1
1	1	0.3
1	1	0.5

X

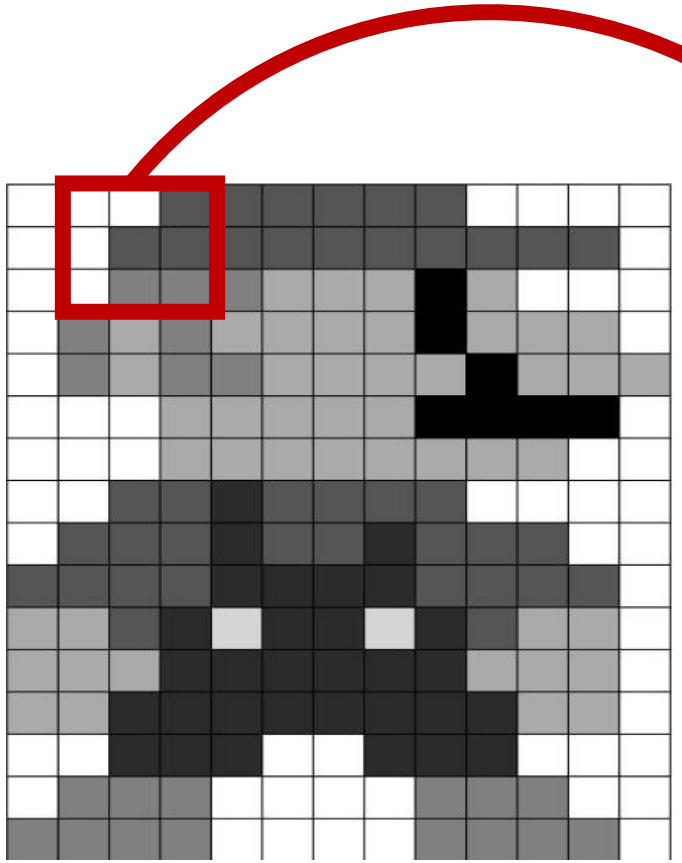
w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Resulting image
(in progress)

42

$= 1 \times w_1 + 1 \times w_2 + 1 \times w_3 + 1 \times w_4 + 1 \times w_5 + 0.3 \times w_6 + [...]$
$= 42$

Convolutional Layer



1	1	0.3
1	0.3	0.3
1	0.5	0.5

X

Kernel
(trainable weights)

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Resulting image
(in progress)

42	34

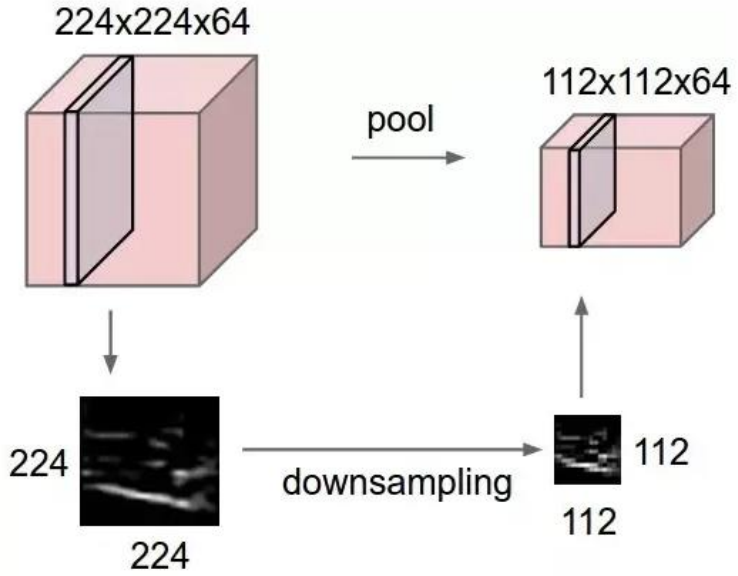
$= 1 \times w_1 + 1 \times w_2 + 0.3 \times w_3 + 1 \times w_4 + 0.3 \times w_5 + 0.3 \times w_6 + [...]$
$= 34$

Pooling Layer

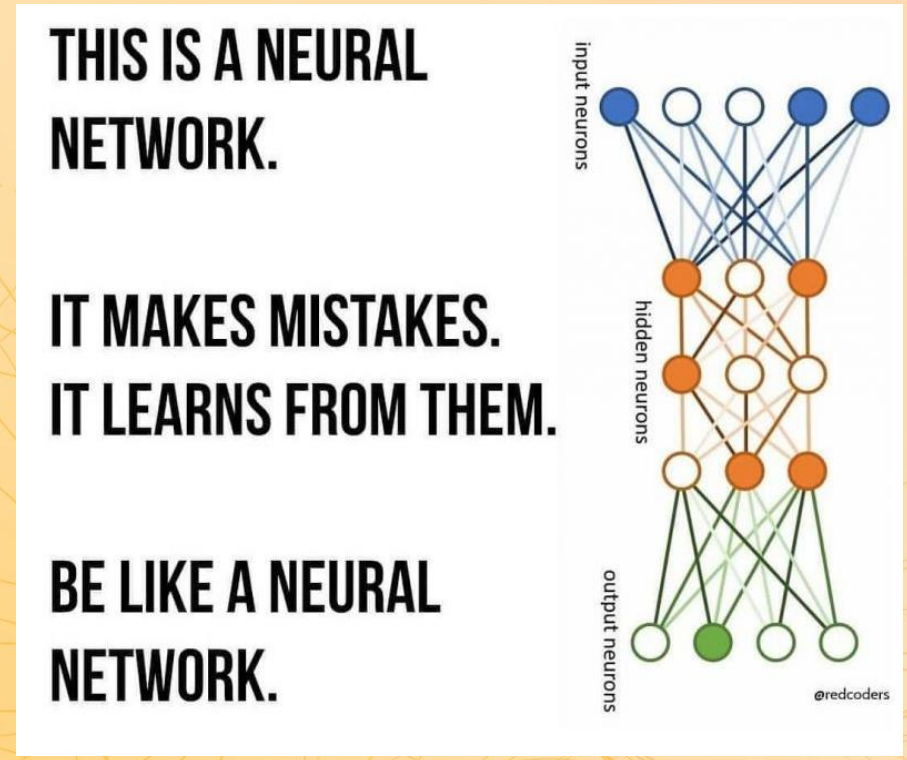
12	20	30	0
8	12	2	0
34	70	37	4
112	100	25	12

2×2 Max-Pool

20	30
112	37



MLCon
CONFERENCE & TRAINING



Let‘s make some mistakes!

Histogram Equalization

4	4	4	4	4
3	4	5	4	3
3	5	5	5	3
3	4	5	4	3
4	4	4	4	4

6	6	6	6	6
2	6	7	6	2
2	7	7	7	2
2	6	7	6	2
6	6	6	6	

Gray Level	No. of pixels (n)	PDF (n/sum)	CDF	CDF * 7 (max gray)	Equalized Histogram
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	6	6/25 = 0.24	0.24	1.68	2
4	14	14/25=0.56	0.8	5.6	6
5	5	5/25=0.2	1.0	7	7
6	0	0	1.0	7	7
7	0	0	1.0	7	7

https://www.youtube.com/watch?v=uqeOrtAzSyU