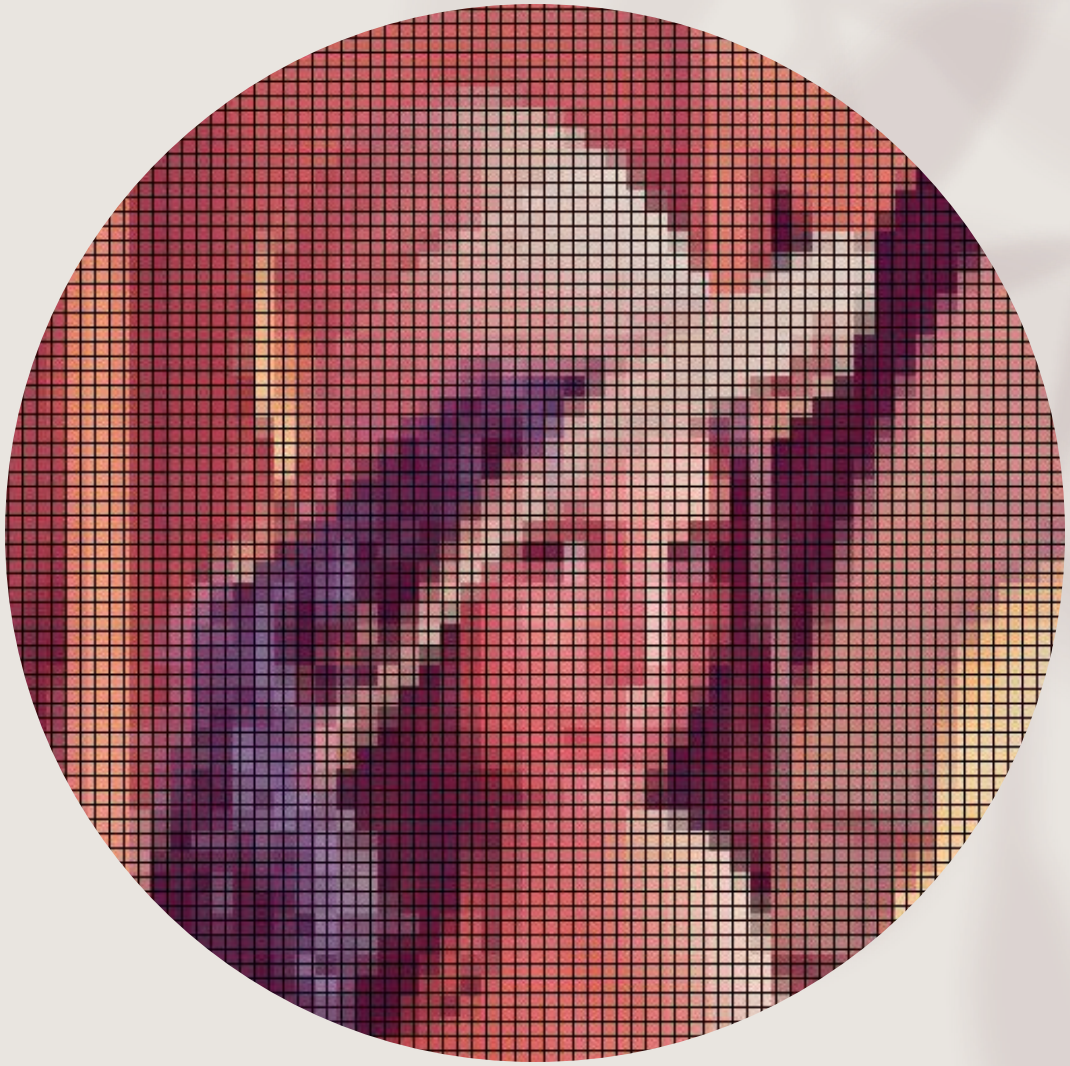
Cyfrowe przetwarzanie obrazów

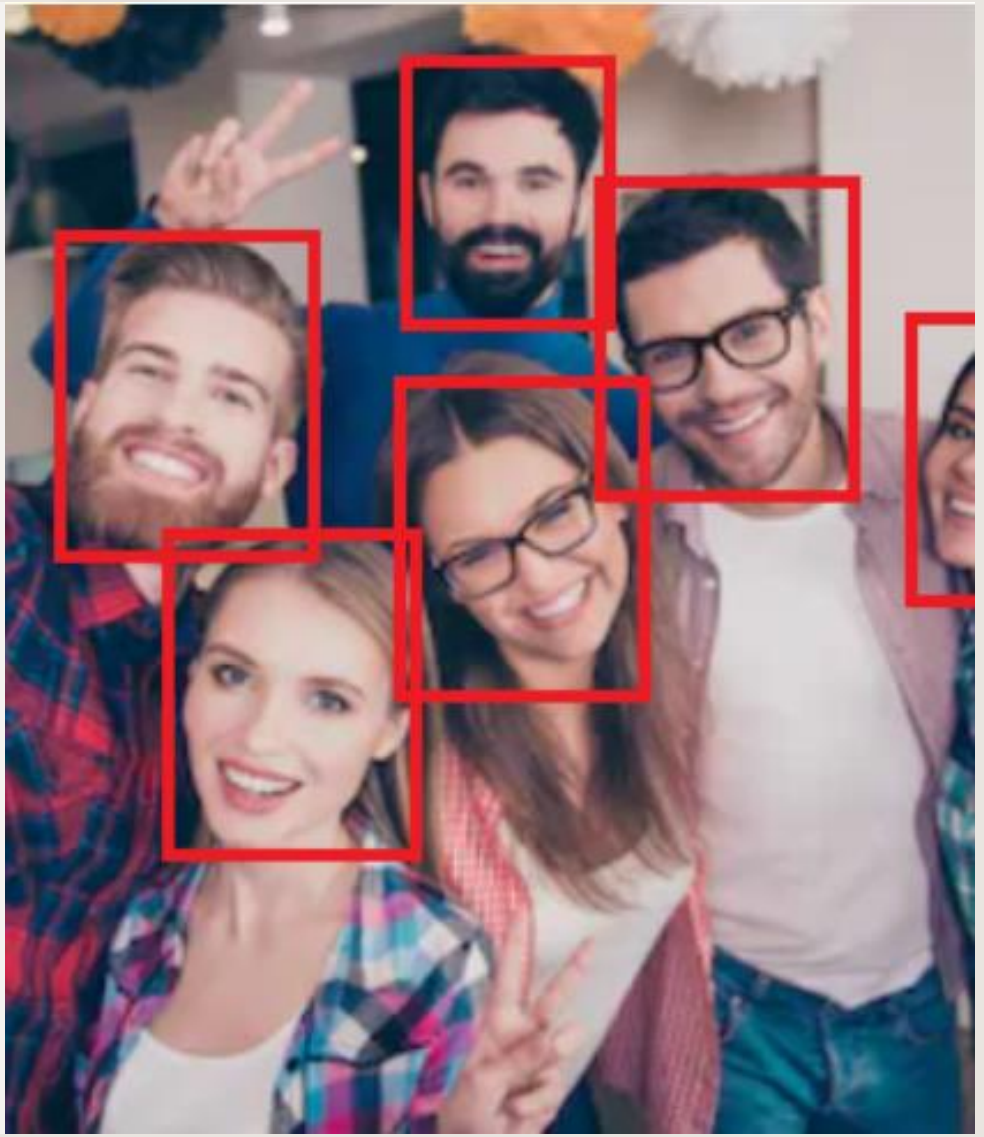
Część 2



Przetwarzanie obrazów

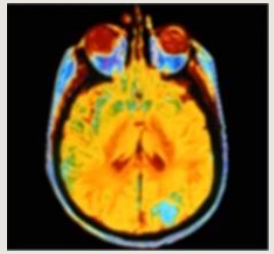
Wykonywanie operacji na zdjęciach w celu ich modyfikacji, wzmocnienia, wydobycia użytecznych informacji, analizy lub podjęcia decyzji.

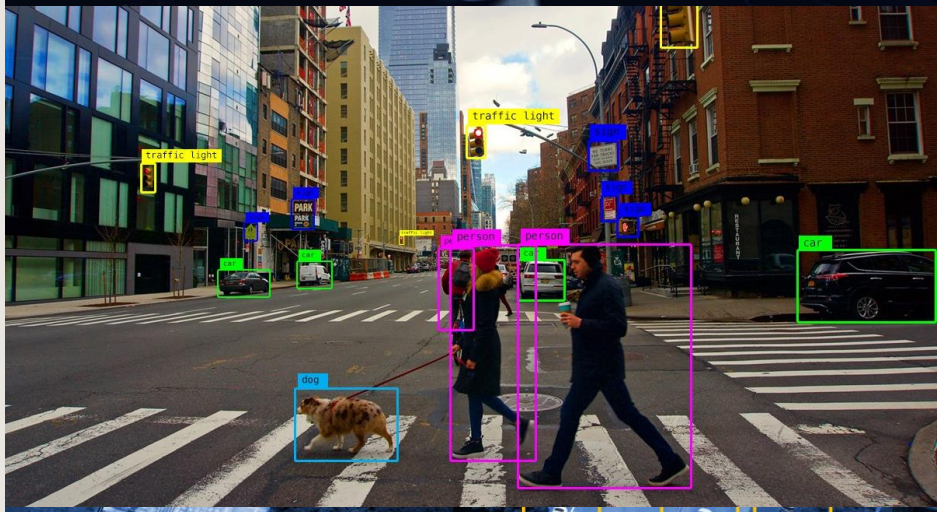
Obecnie dyscyplina zaliczana do dziedziny widzenia komputerowego (computer vision)

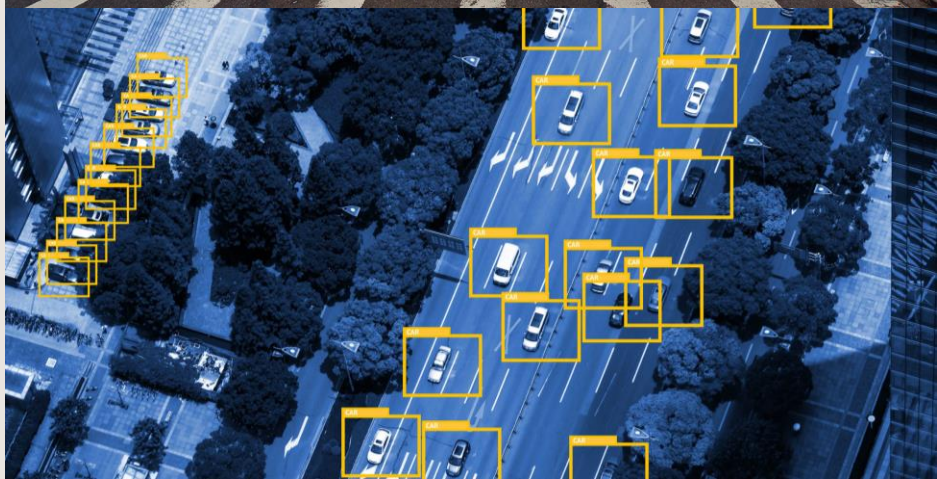


Przykłady zastosowań

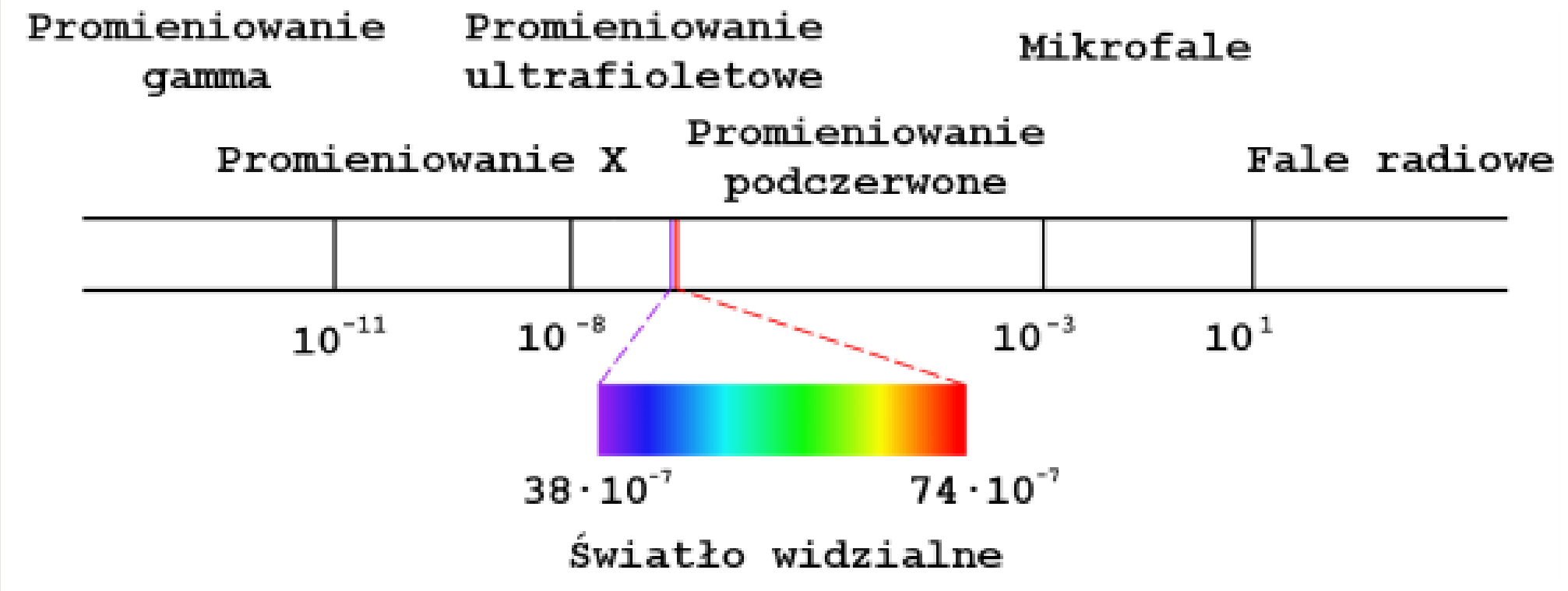
- Sztuczna inteligencja
- Rozpoznawanie obrazów
- Systemy wizyjne
- Analiza przestrzenna
- Obrazowanie medyczne
- inne ...







Akwizycja sygnału



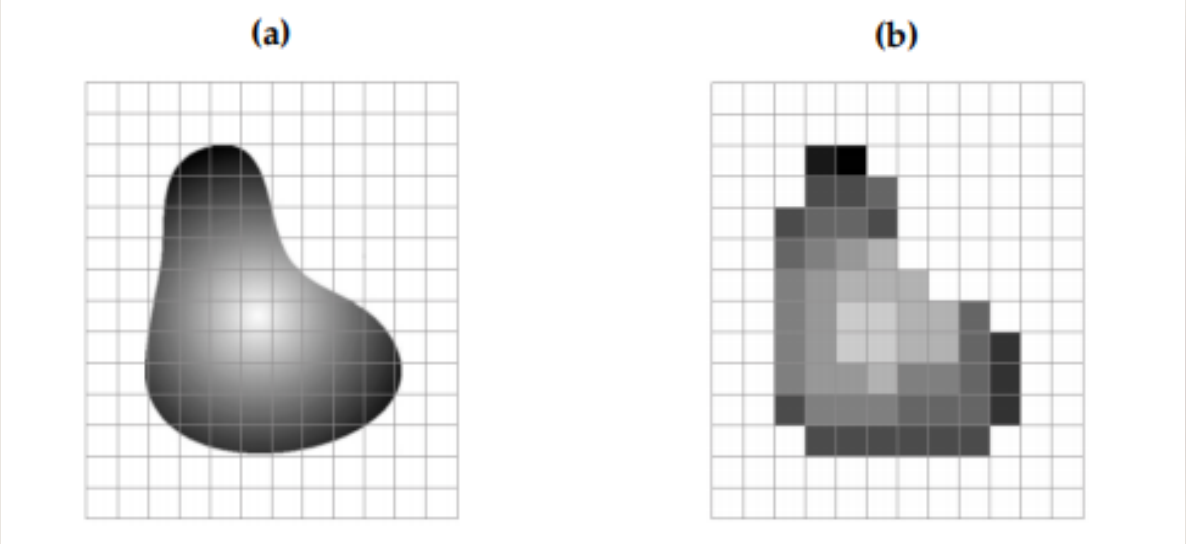
Spektrum promieniowania elektromagnetycznego [m]

Preprocessing

W ramach digitalizacji (cyfryzacji) sygnał podlega:

próbkowaniu

kwantyzacji

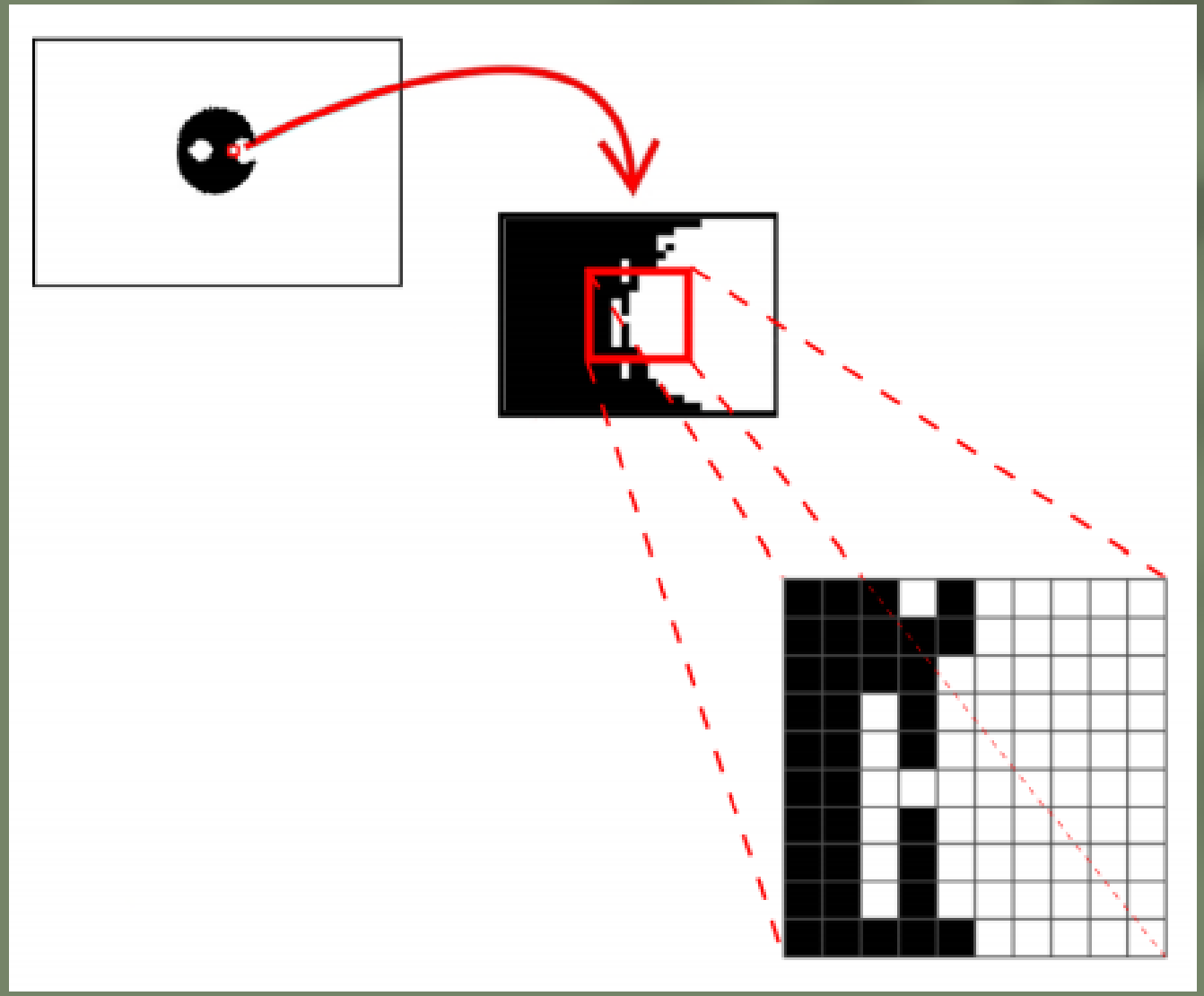


a. Ciągła projekcja na światłoczułą matrycę b. Efekt próbkowania i kwantyzacji.

Obraz reprezentatywny

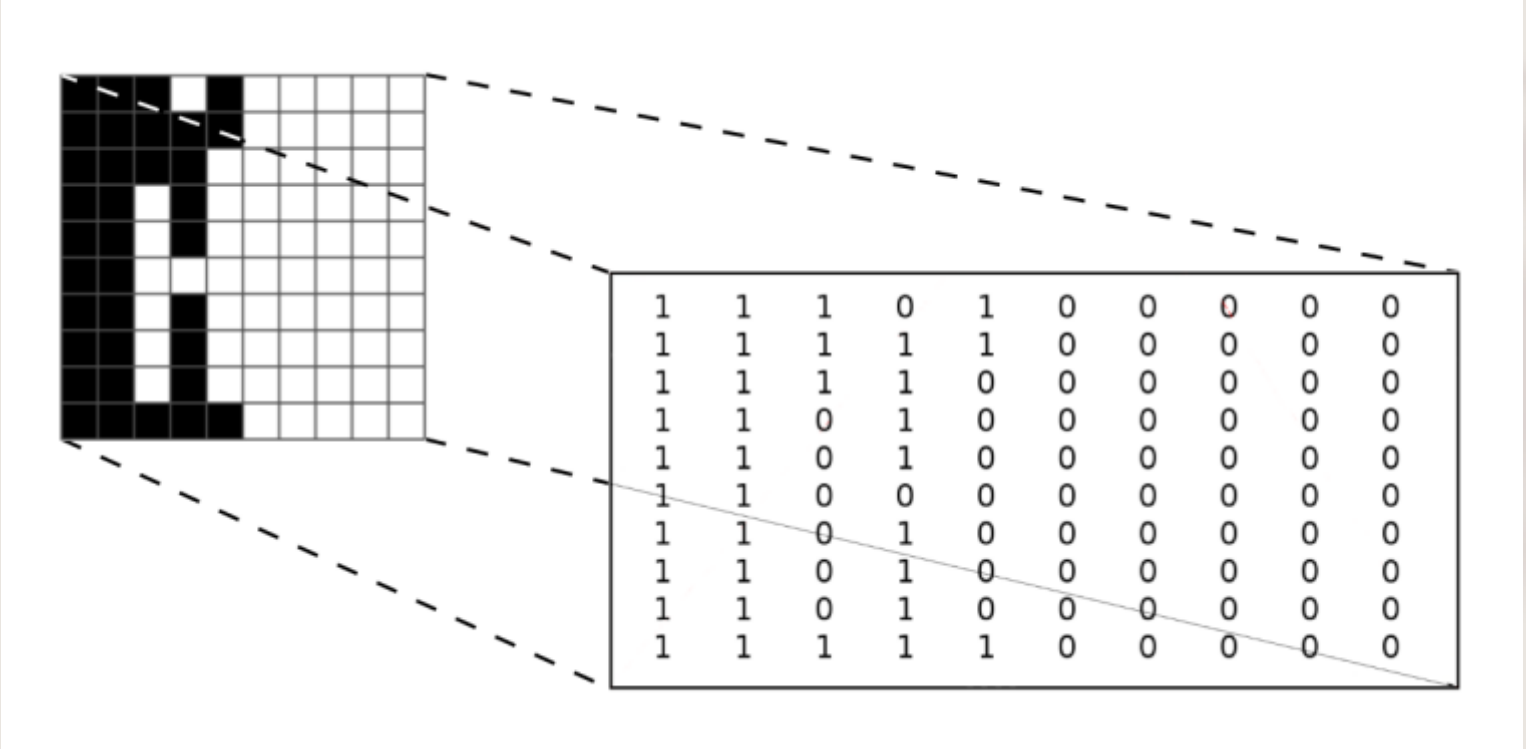
Obraz źrenicy oka w bliskiej podczerwieni

Struktura czarno-białego (binarnego) obrazu



Metody reprezentacji obrazu

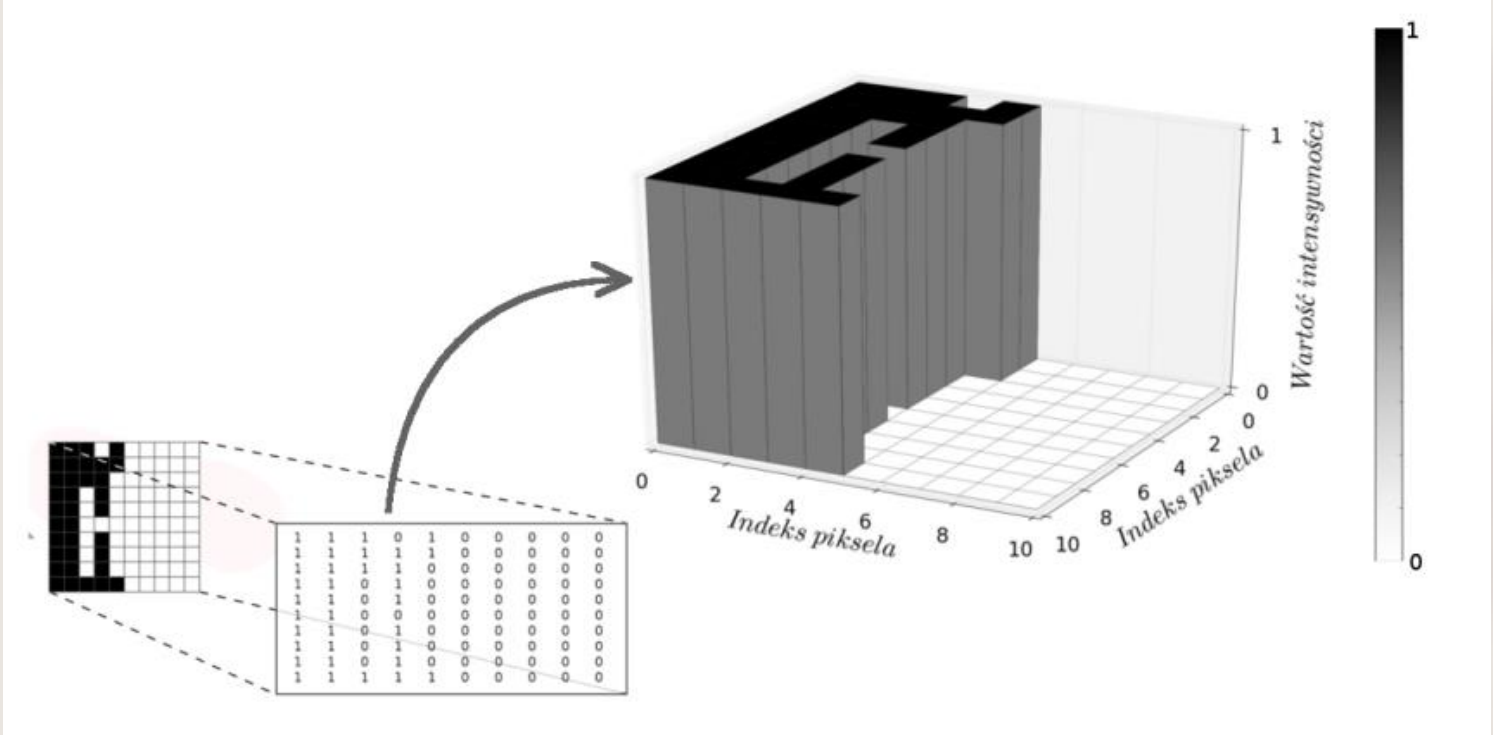
Macierzowa reprezentacja obrazu



Metody reprezentacji obrazu

Funkcyjna reprezentacja obrazu

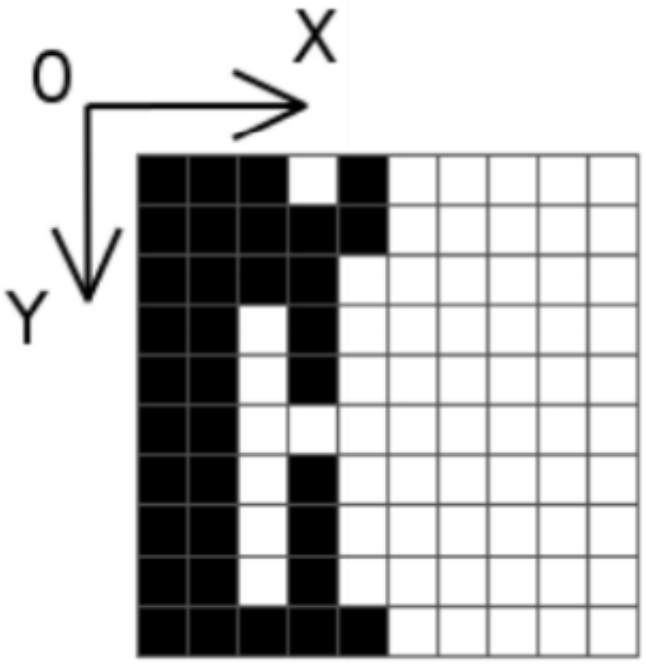
W ujęciu algebraicznym macierz może reprezentować odwzorowanie (funkcję). Tym samym o obrazie możemy myśleć jako o funkcji (odwzorowaniu).



Metody reprezentacji obrazu

Teoriomnogościowa reprezentacja obrazu

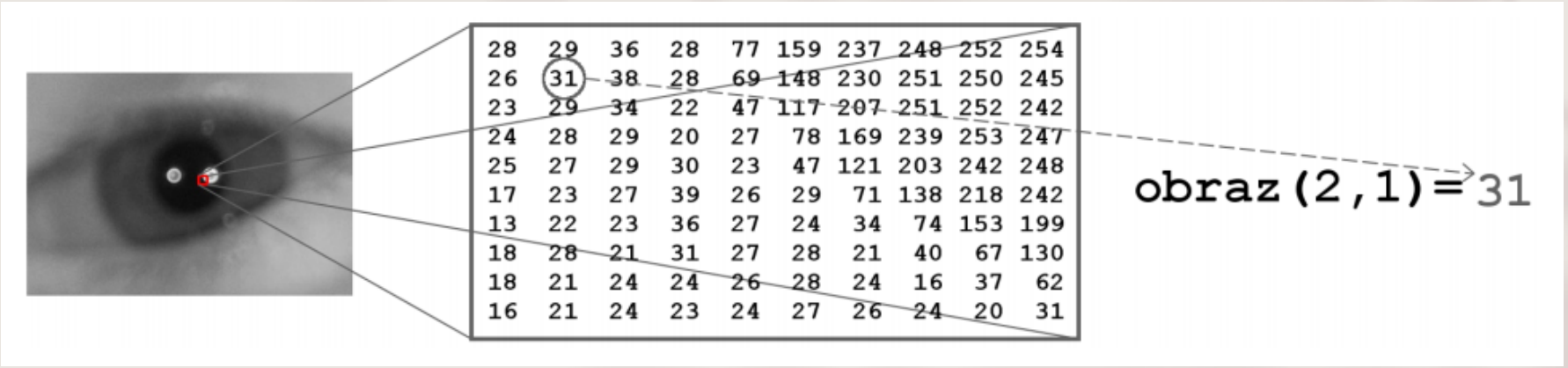
Obraz binarny jako zbiór tych punktów (x, y), których wartość intensywności wynosi 1.



$$
\begin{aligned}
A = \{ & \\
& (0,0), (1,0), (2,0), (4,0), \\
& (0,1), (1,1), (2,1), (3,1), (4,1), \\
& (0,2), (1,2), (2,2), (3,2), \\
& (0,3), (1,3), (3,3), \\
& (0,4), (1,4), (3,4), \\
& (0,5), (1,5), \\
& (0,6), (1,6), (3,6), \\
& (0,7), (1,7), (3,7), \\
& (0,8), (1,8), (3,8), \\
& (0,9), (1,9), (2,9), (3,9), (4,9) \\
\} &
\end{aligned}
$$

Uogólnienie na obrazy w skali szarości

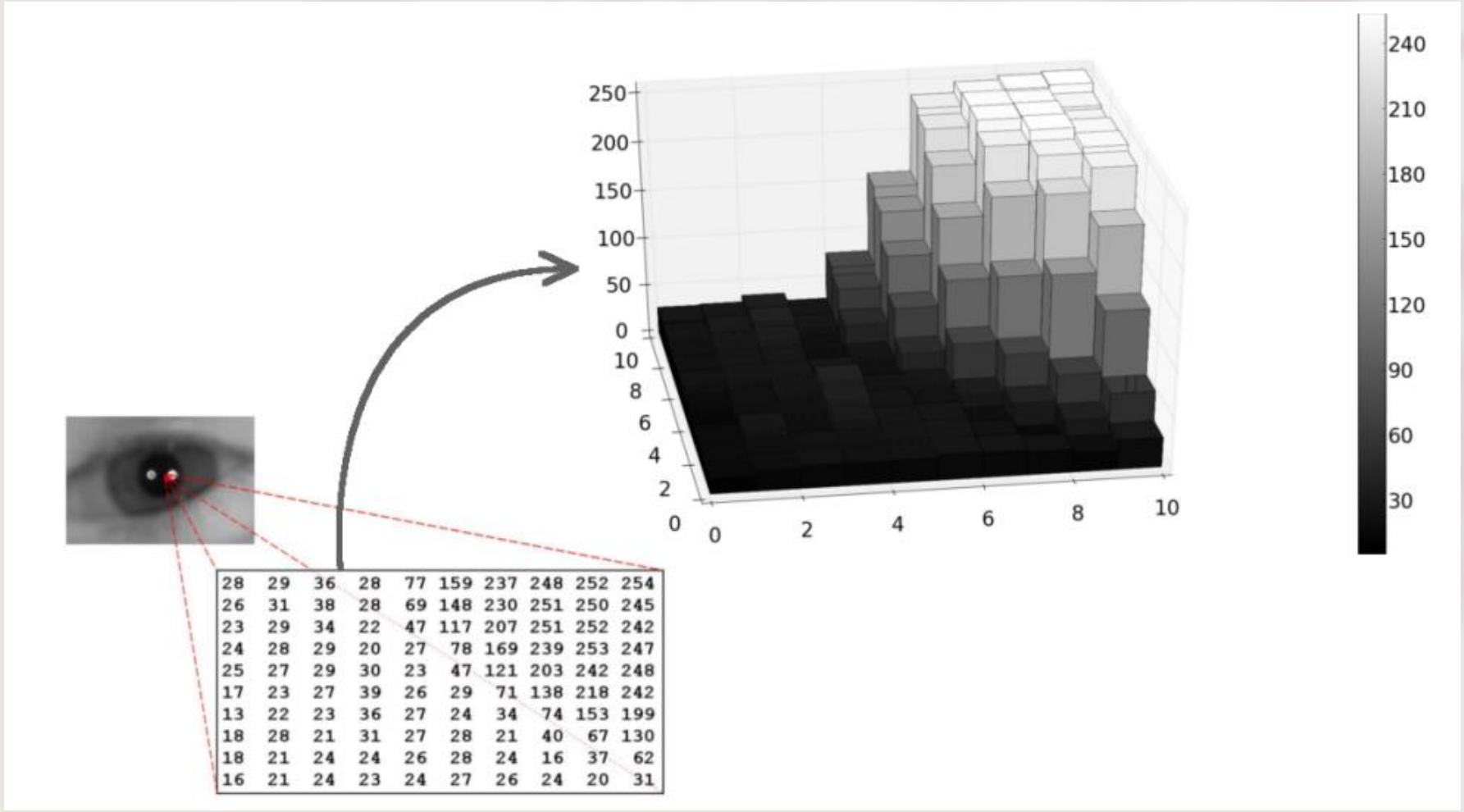
Reprezentacja macierzowa



Macierz jako reprezentacja cyforwego obrazu w skali szarości oraz skalar jako reprezentacja piksela w skali szarości.

Uogólnienie na obrazy w skali szarości

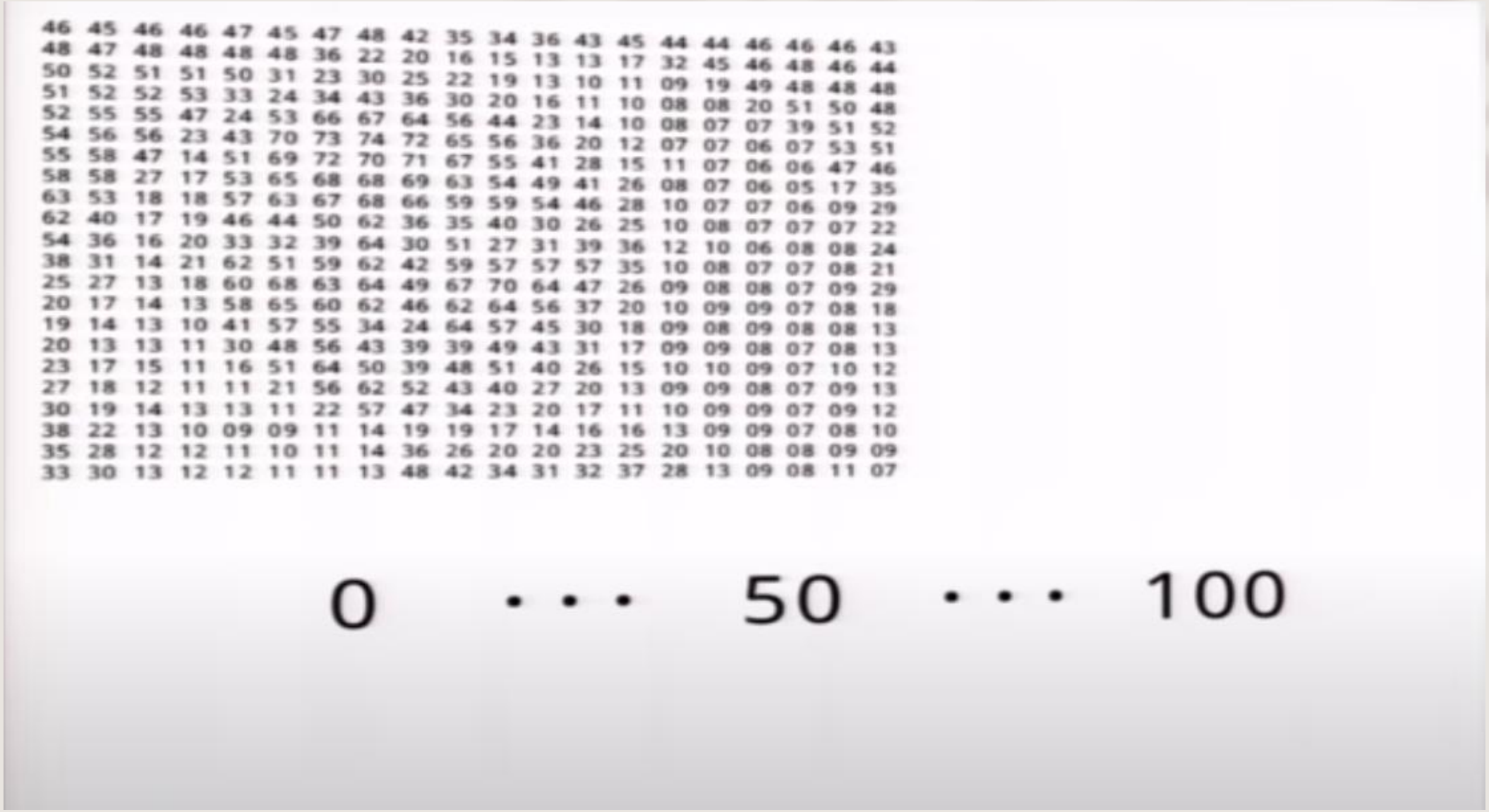
Reprezentacja funkcyjna



Funkcja jako reprezentacja cyfrowego obrazu w skali szarości

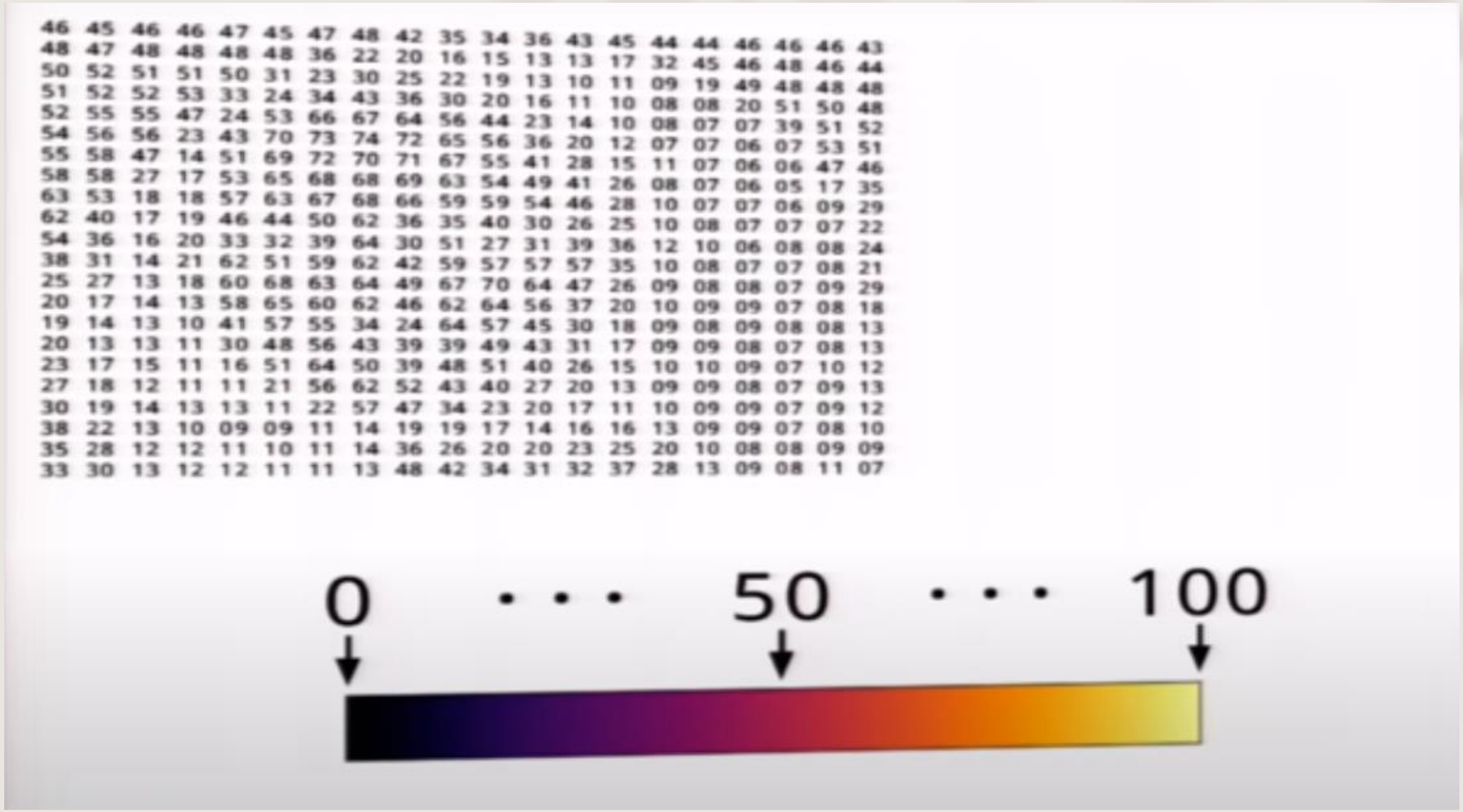
Mapa kolorów

Mapa kolorów

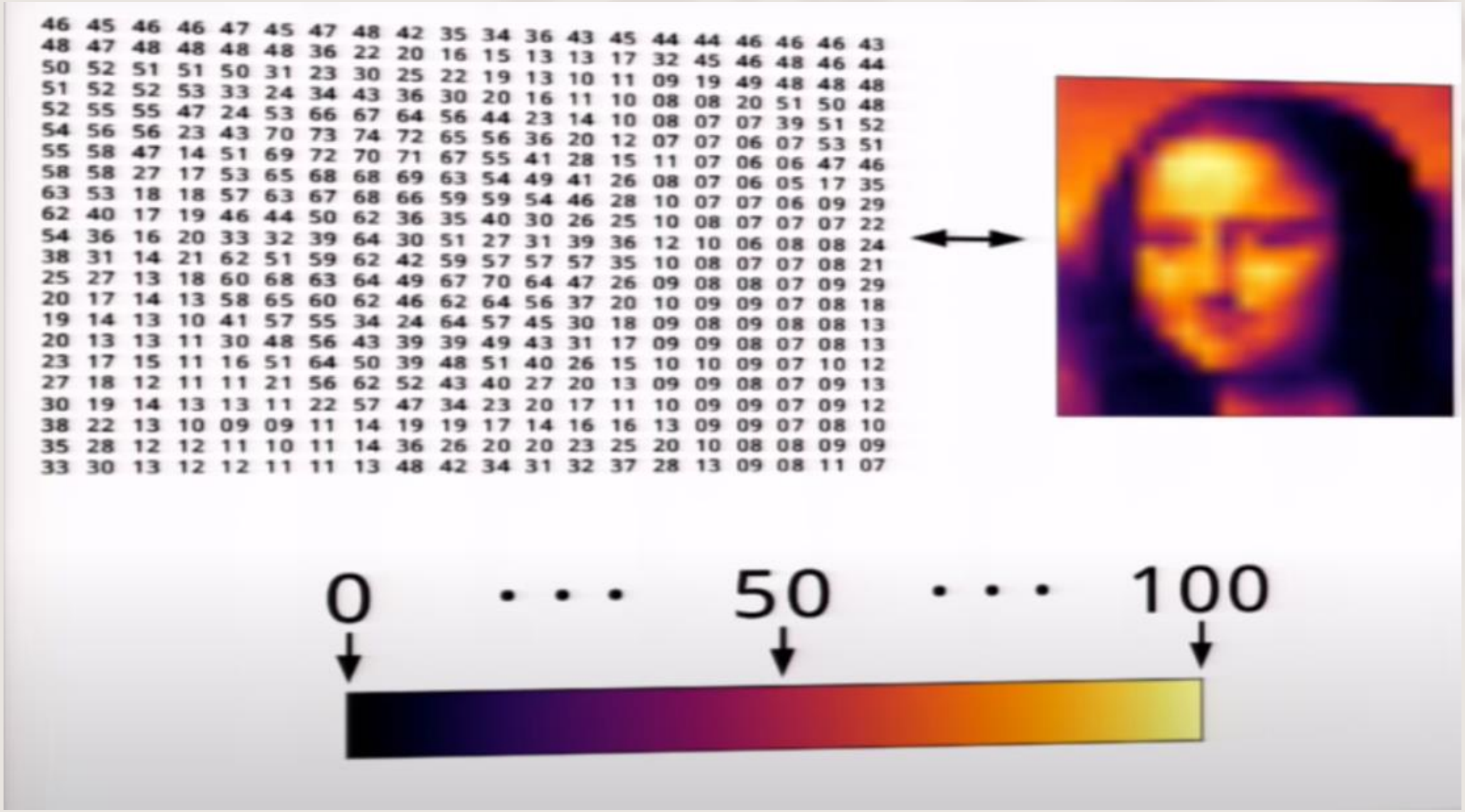


Mapa kolorów

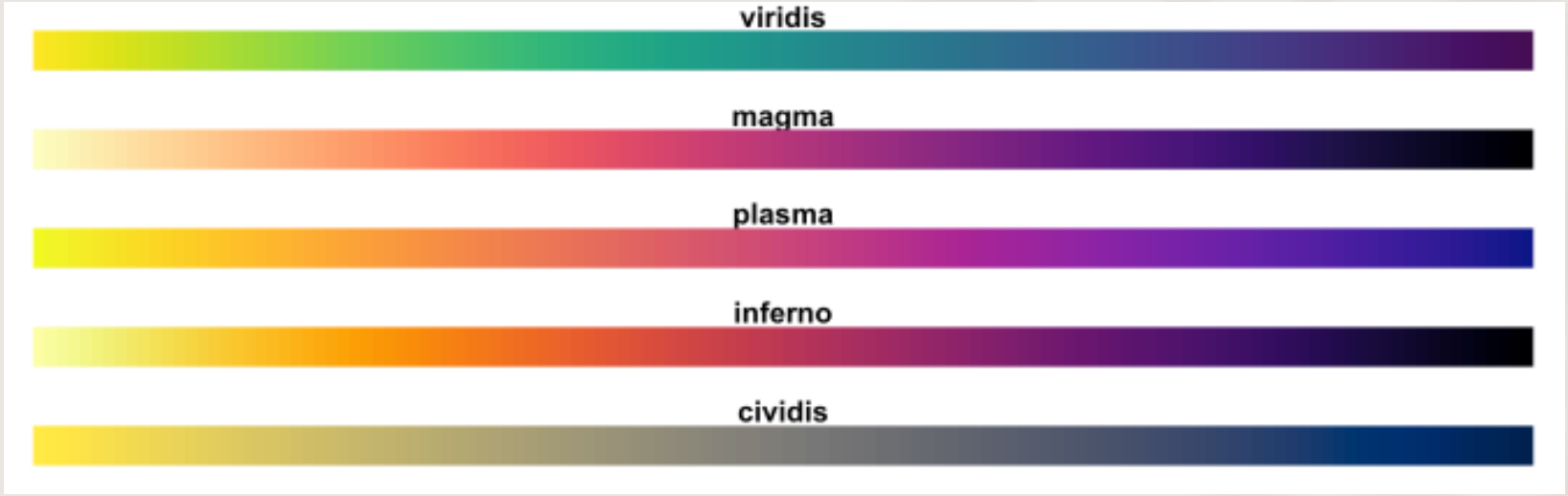
46	45	46	46	47	45	47	48	42	35	34	36	43	45	44	44	46	46	46	43
48	47	48	48	48	48	36	22	20	16	15	13	13	17	32	45	46	48	46	44
50	52	51	51	50	31	23	30	25	22	19	13	10	11	09	19	49	48	48	48
51	52	52	53	33	24	34	43	36	30	20	16	11	10	08	08	20	51	50	48
52	55	55	47	24	53	66	67	64	56	44	23	14	10	08	07	07	39	51	52
54	56	56	23	43	70	73	74	72	65	56	36	20	12	07	07	06	07	53	51
55	58	47	14	51	69	72	70	71	67	55	41	28	15	11	07	06	06	47	46
58	58	27	17	53	65	68	68	69	63	54	49	41	26	08	07	06	05	17	35
63	53	18	18	57	63	67	68	66	59	59	54	46	28	10	07	07	06	09	29
62	40	17	19	46	44	50	62	36	35	40	30	26	25	10	08	07	07	07	22
54	36	16	20	33	32	39	64	30	51	27	31	39	36	12	10	06	08	08	24
38	31	14	21	62	51	59	62	42	59	57	57	57	35	10	08	07	07	08	21
25	27	13	18	60	68	63	64	49	67	70	64	47	26	09	08	08	07	09	29
20	17	14	13	58	65	60	62	46	62	64	56	37	20	10	09	09	07	08	18
19	14	13	10	41	57	55	34	24	64	57	45	30	18	09	08	09	08	08	13
20	13	13	11	30	48	56	43	39	39	49	43	31	17	09	09	08	07	08	13
23	17	15	11	16	51	64	50	39	48	51	40	26	15	10	10	09	07	10	12
27	18	12	11	11	21	56	62	52	43	40	27	20	13	09	09	08	07	09	13
30	19	14	13	13	11	22	57	47	34	23	20	17	11	10	09	09	07	09	12
38	22	13	10	09	09	11	14	19	19	17	14	16	16	13	09	09	07	08	10
35	28	12	12	11	10	11	14	36	26	20	20	23	25	20	10	08	08	09	09
33	30	13	12	12	11	11	13	48	42	34	31	32	37	28	13	09	08	11	07

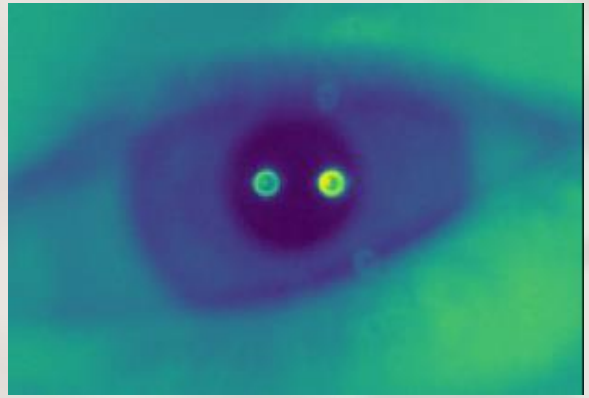


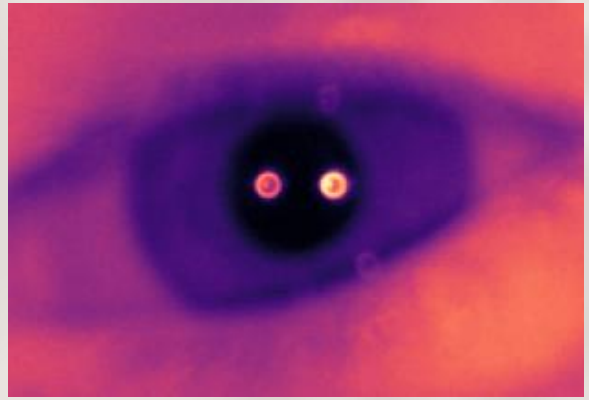
Mapa kolorów

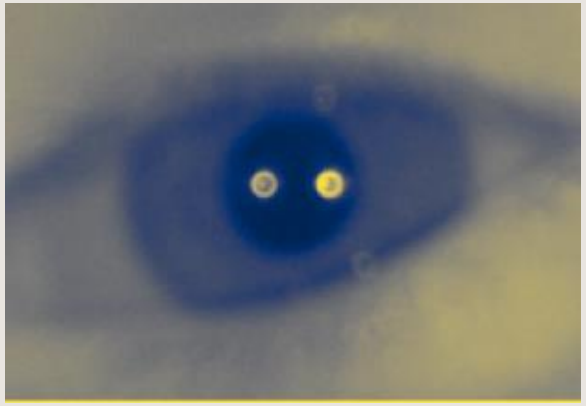


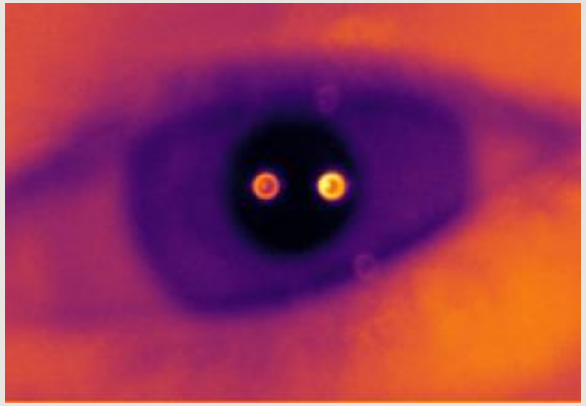
Mapy kolorów

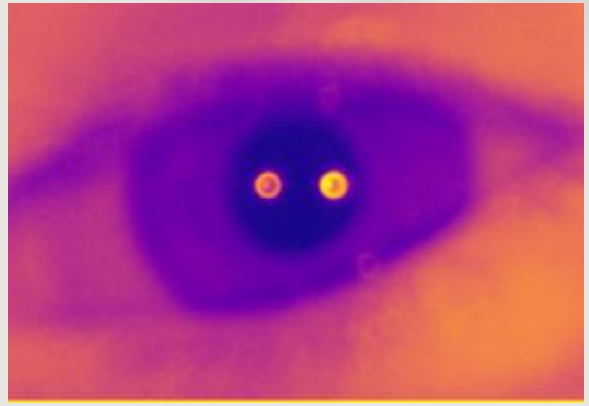












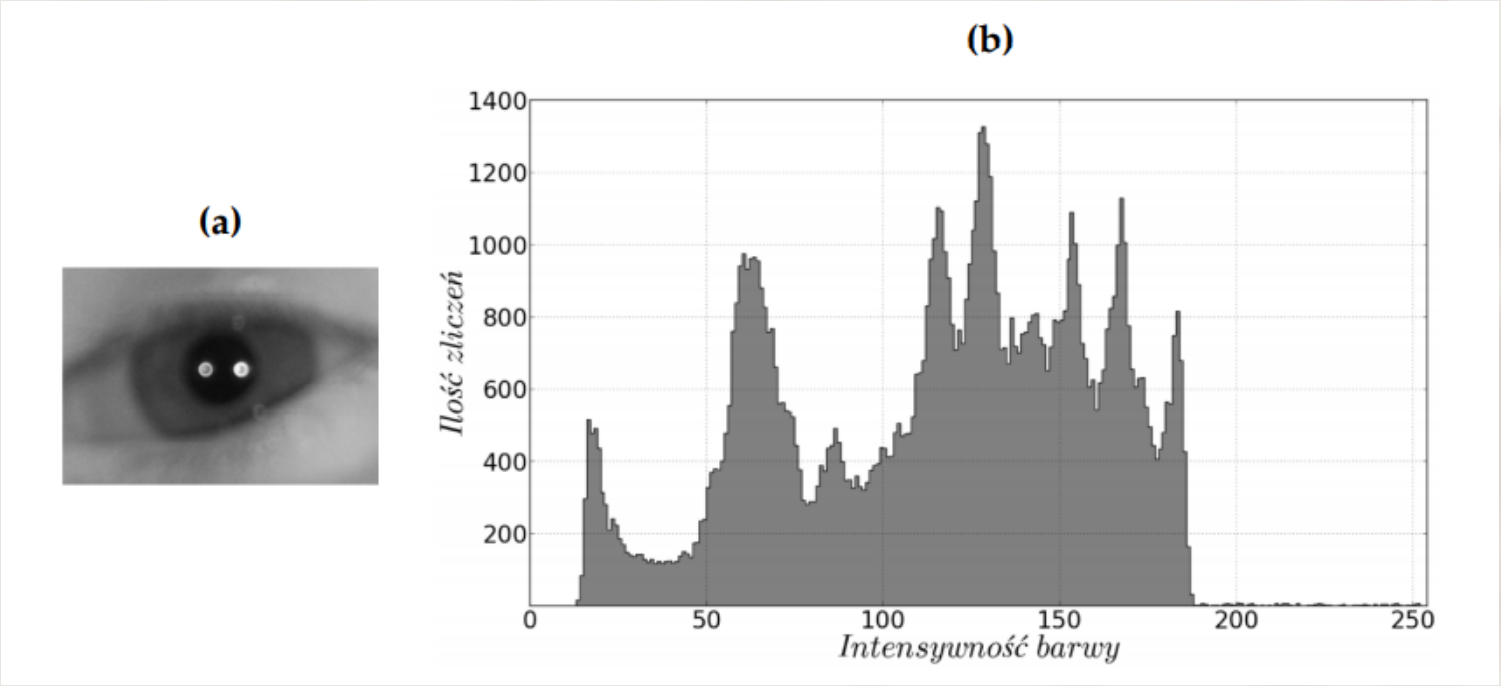
Uogólnienie na obrazy w skali szarości

Reprezentacja teoriomnogościowa może zostać uogólniona poprzez zastosowanie:

- trójwymiarowych zbiorów
- zbioru trójwymiarowych wektorów
- rodziny zbiorów dwuwymiarowych

Uogólnienie na obrazy w skali szarości

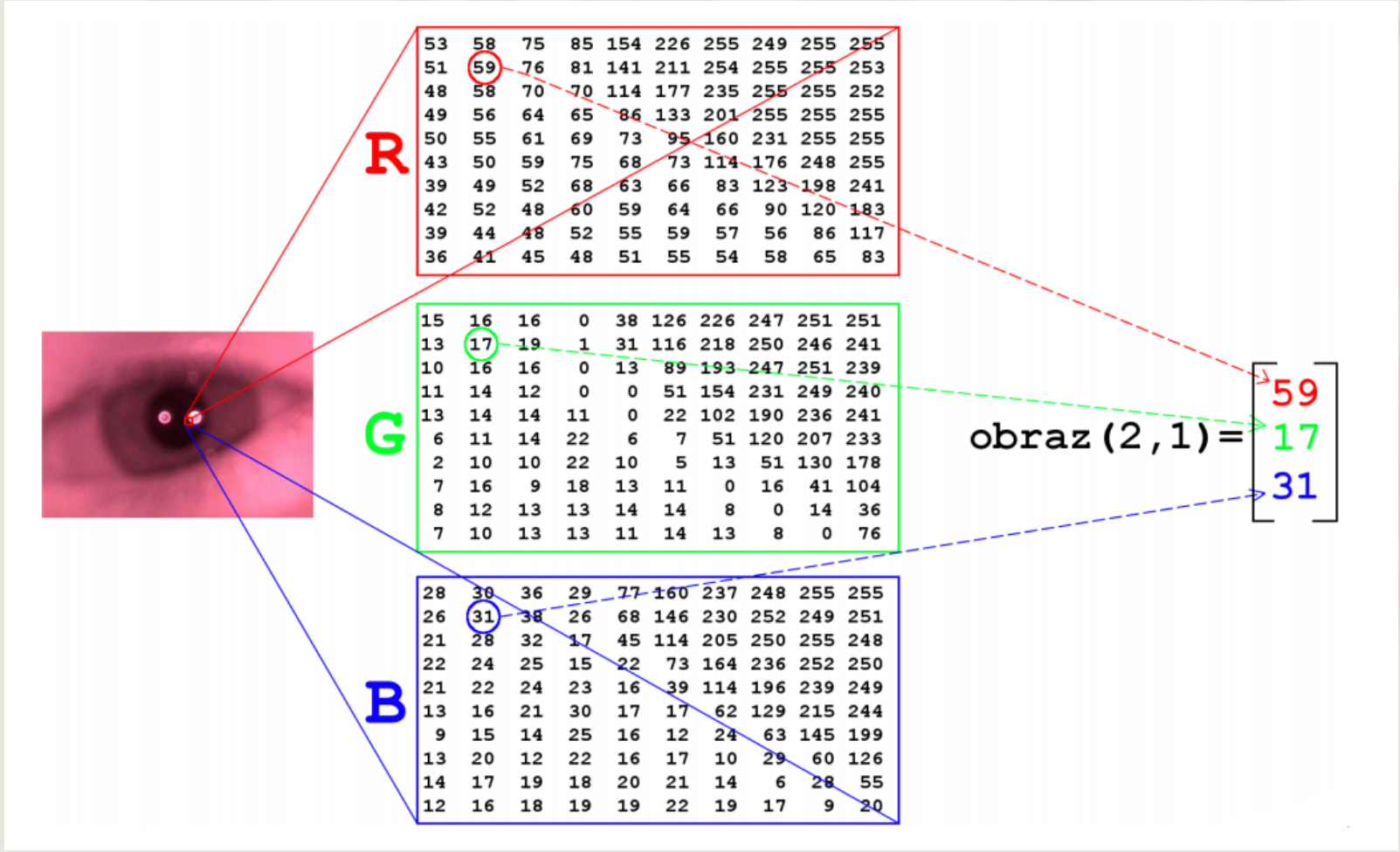
Probabilistyczna reprezentacja obrazu



a. Obraz w skali szarości b. Histogram

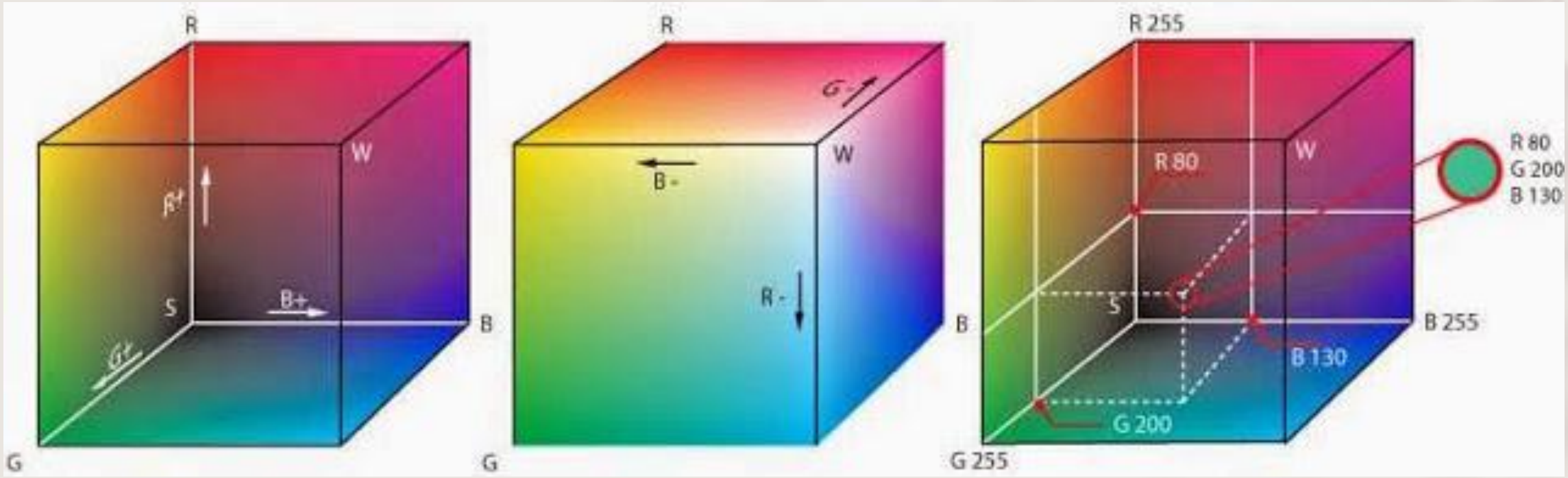
Uogólnienie na obrazy kolorowe

Macierzowa reprezentacja obrazu

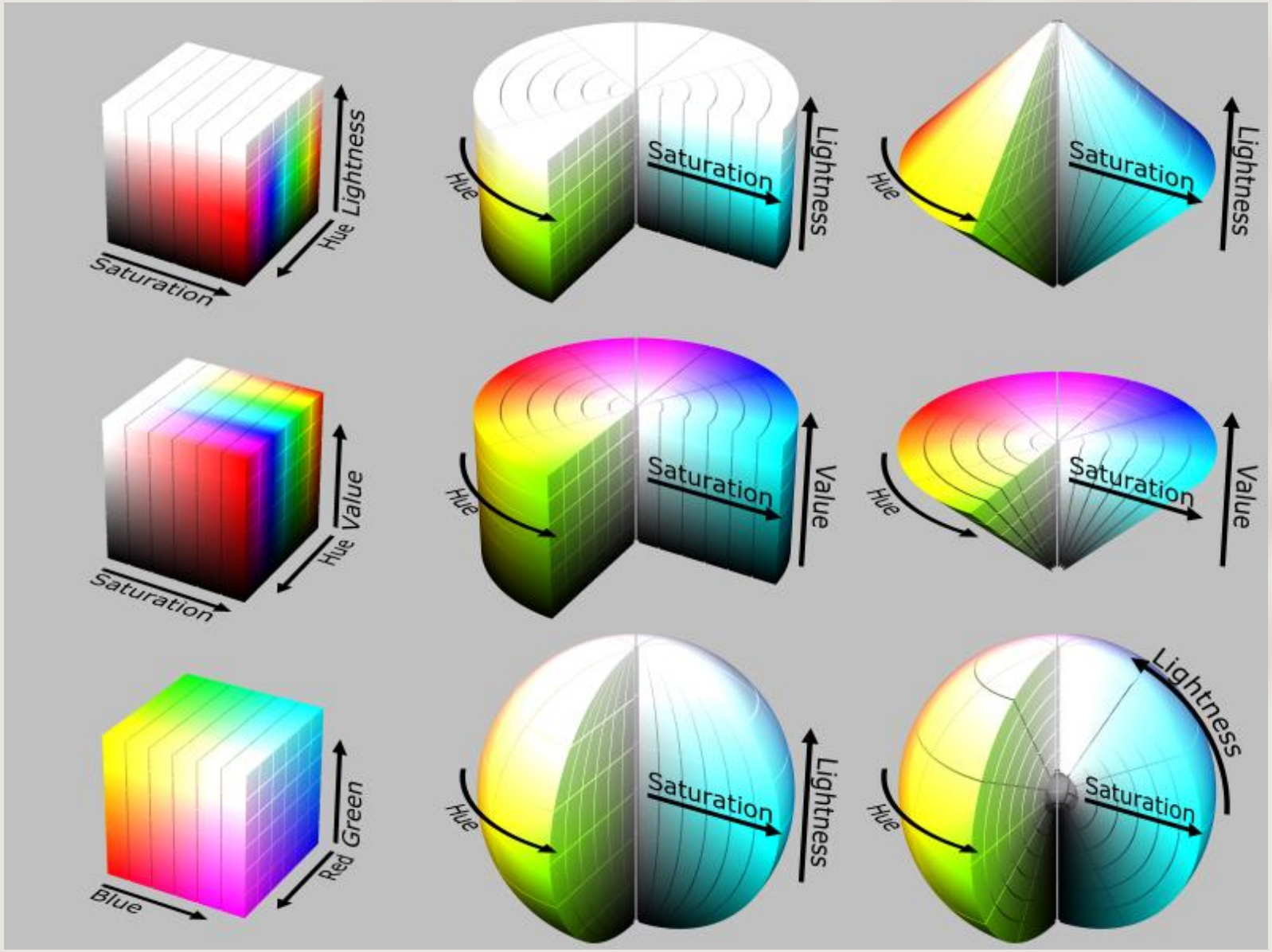


Trzy macierze jako reprezentacja RGB cyfrowego obrazu kolorowego oraz trójwymiarowy wektor jako reprezentacja kolorowego piksela.

Przestrzenie kolorów



Przestrzenie kolorów



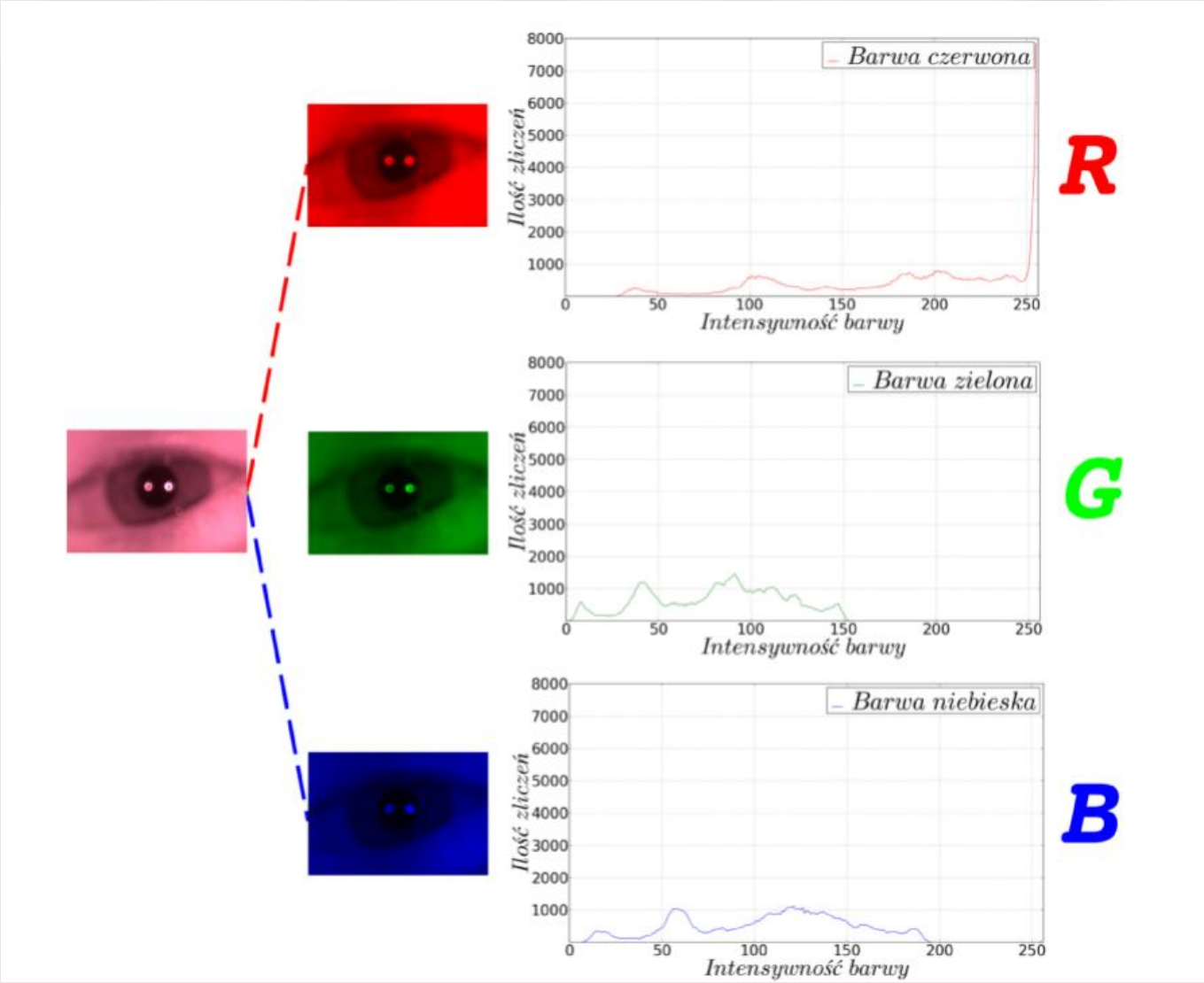
Uogólnienie na obrazy kolorowe

Funkcyjna reprezentacja obrazu zostaje uogólniona poprzez rozpatrywanie funkcji wektorowych.

Teoriomnogościowa reprezentacja obrazu zostaje uogólniona poprzez zwiększenie wymiarowości zagadnienia.

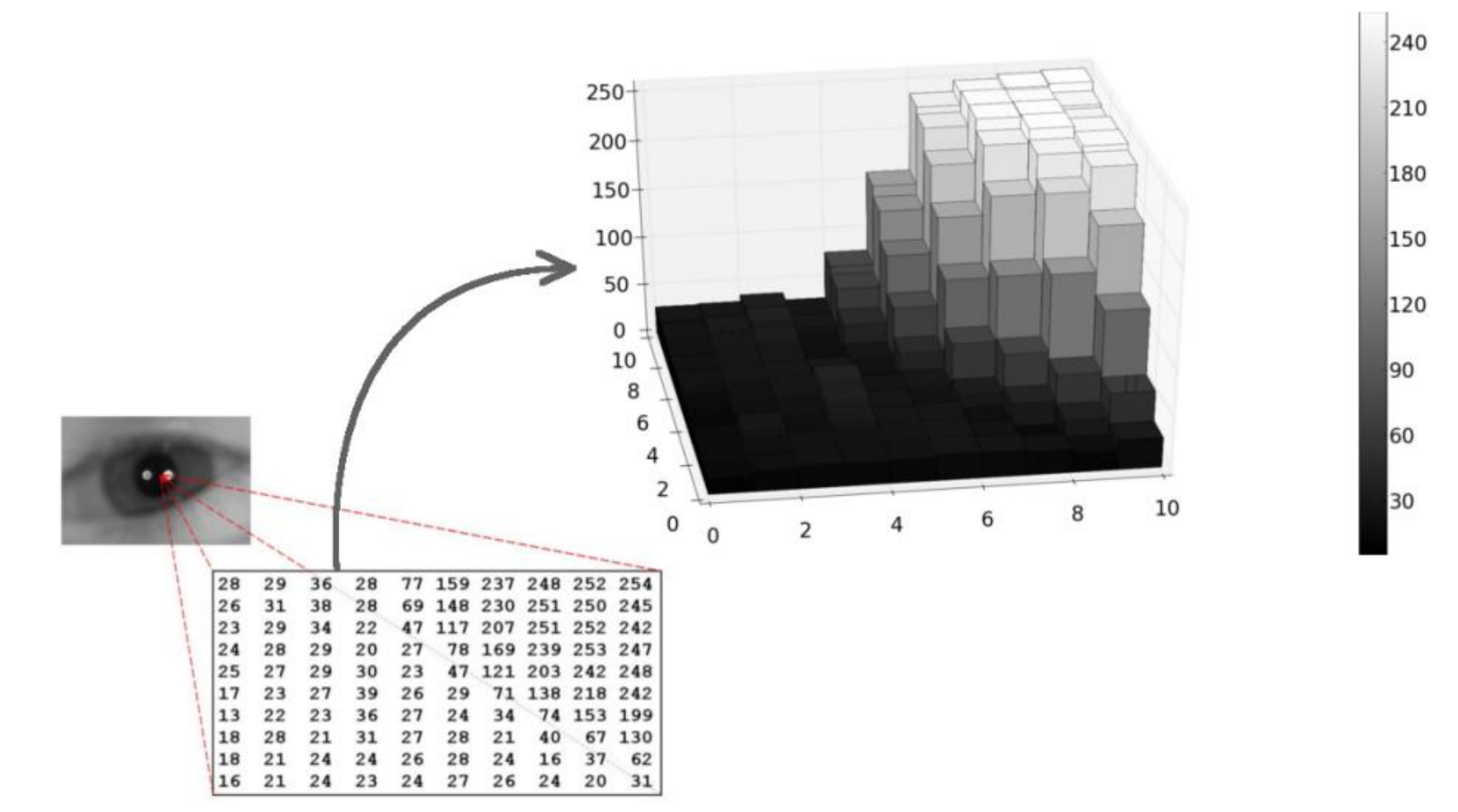
Uogólnienie na obrazy kolorowe

Probabilistyczna reprezentacja obrazu

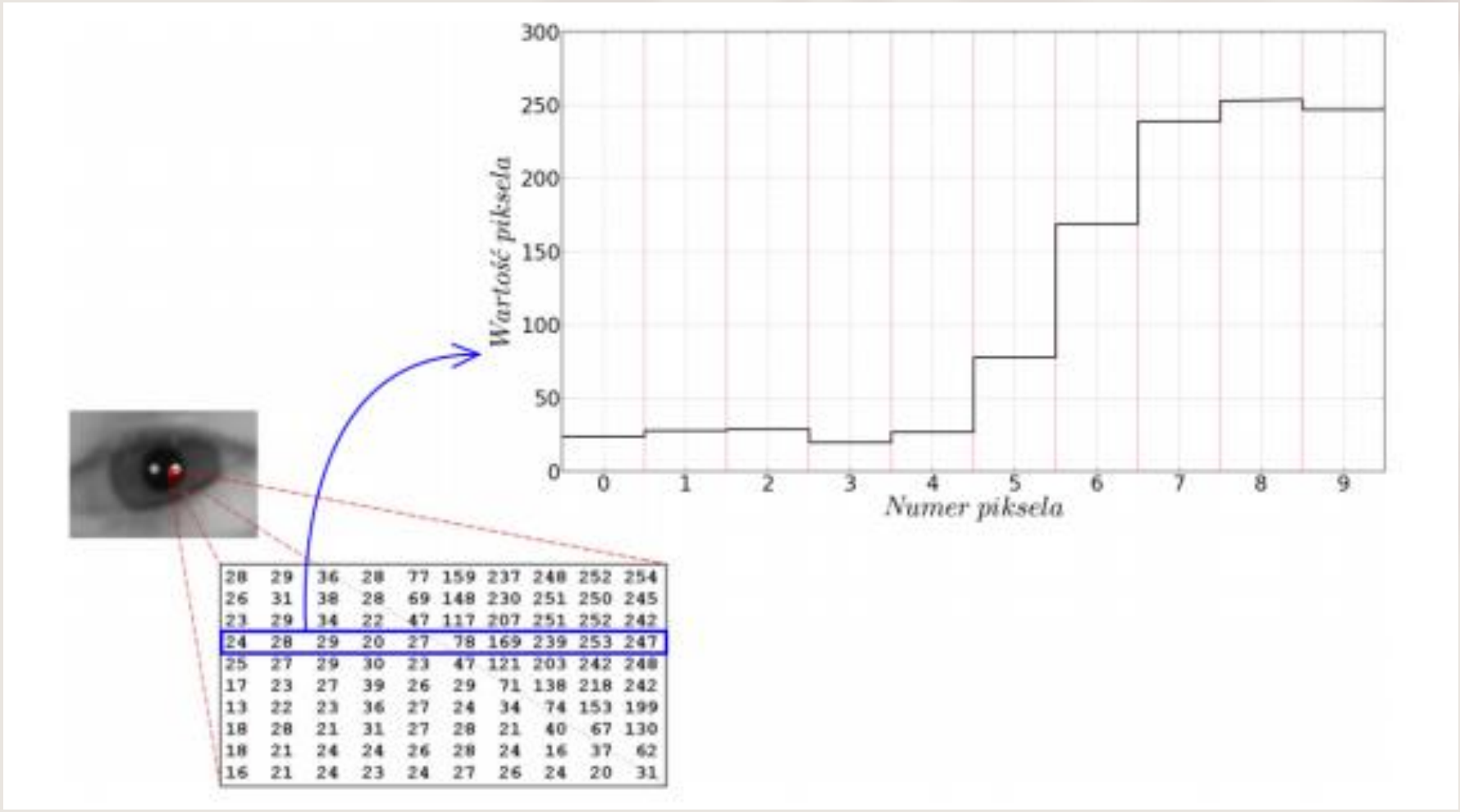


Histogramy trzech podstawowych składowych obrazu (RGB)

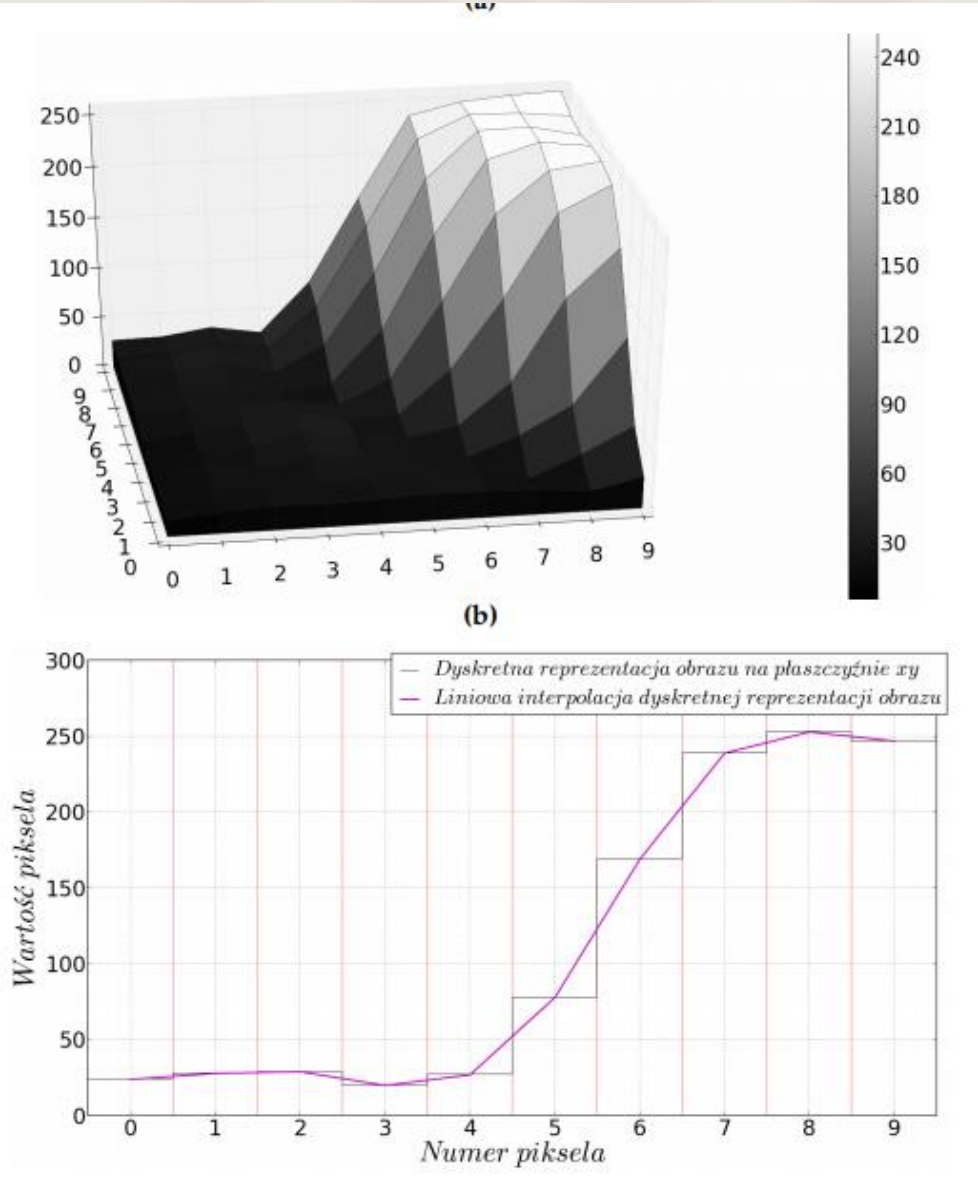
Wykres 3D jako reprezentacja obrazu



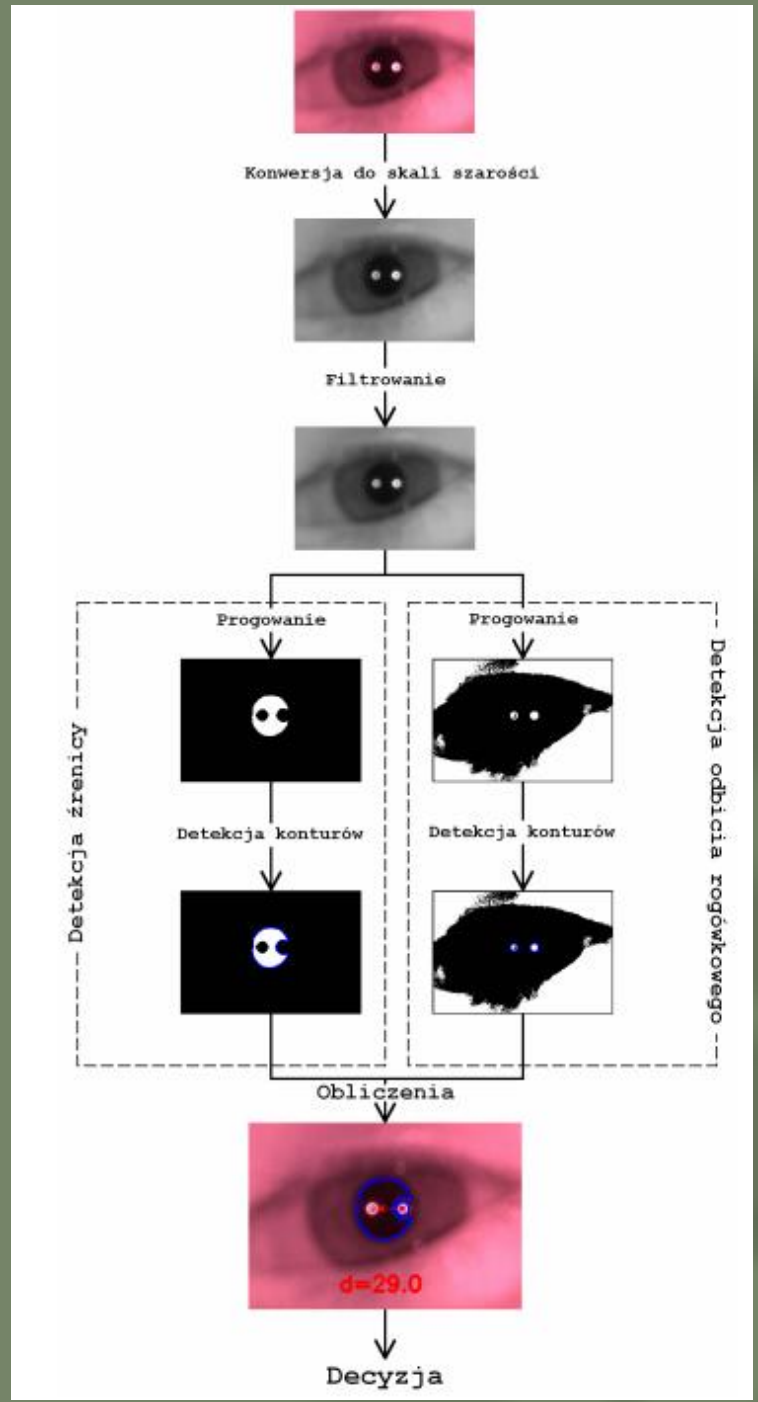
Wykres 2D jako reprezentacja obrazu



Liniowa interpolacja obrazu

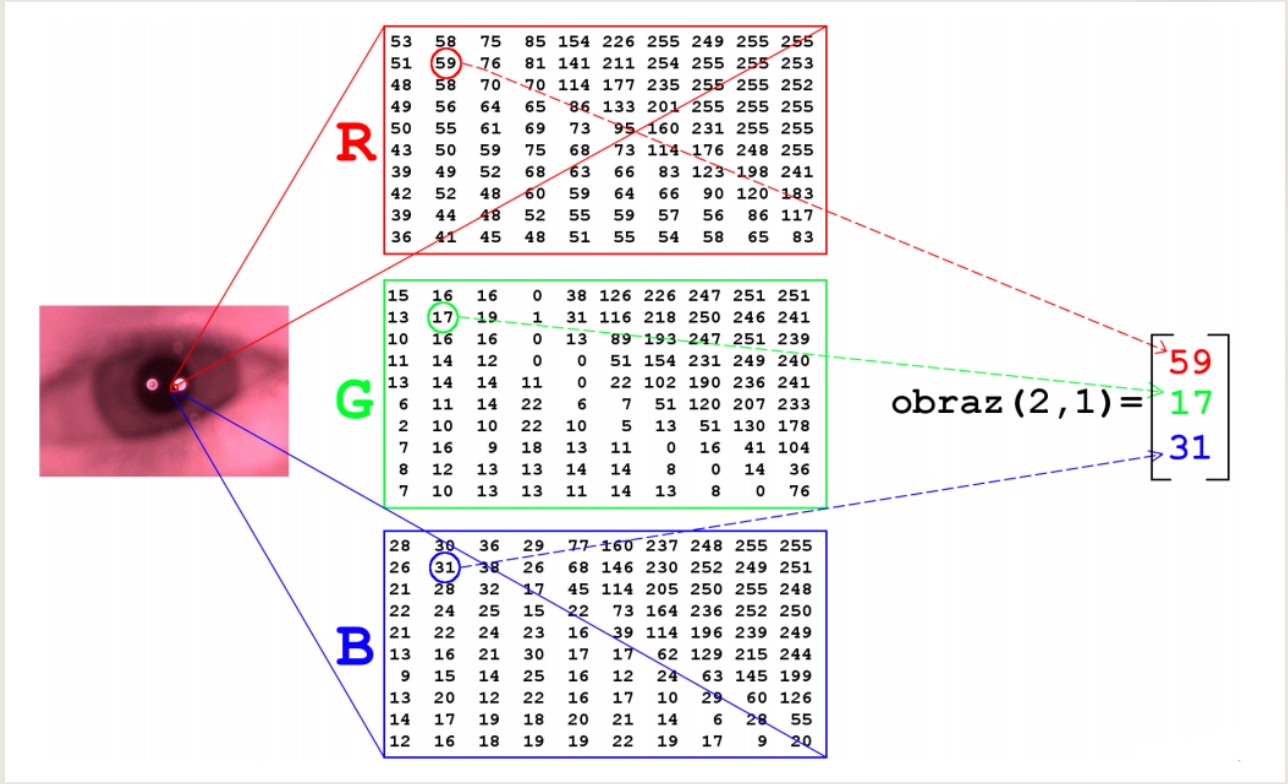


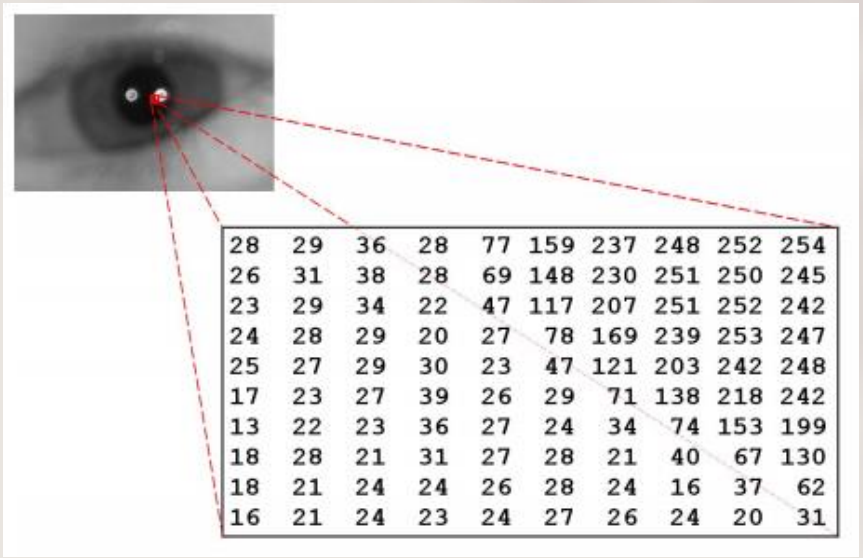
Przykładowy pipeline



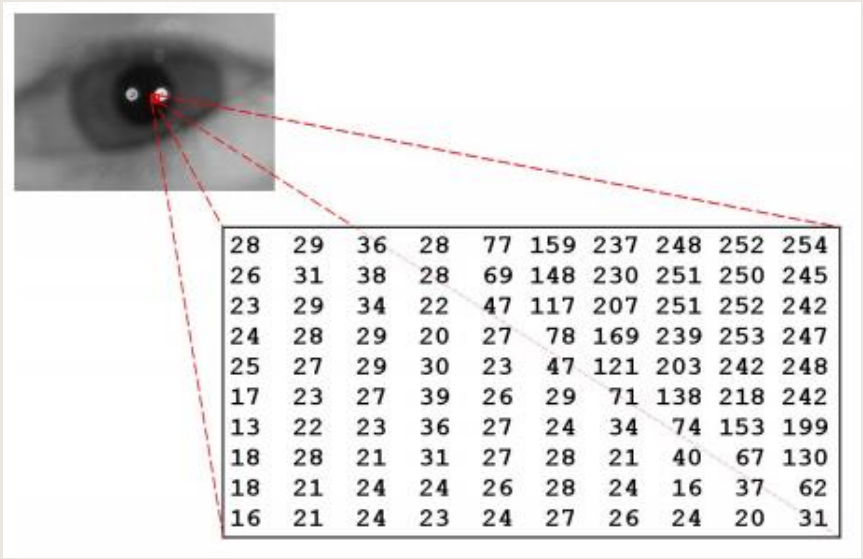
Konwersja do skali szarości

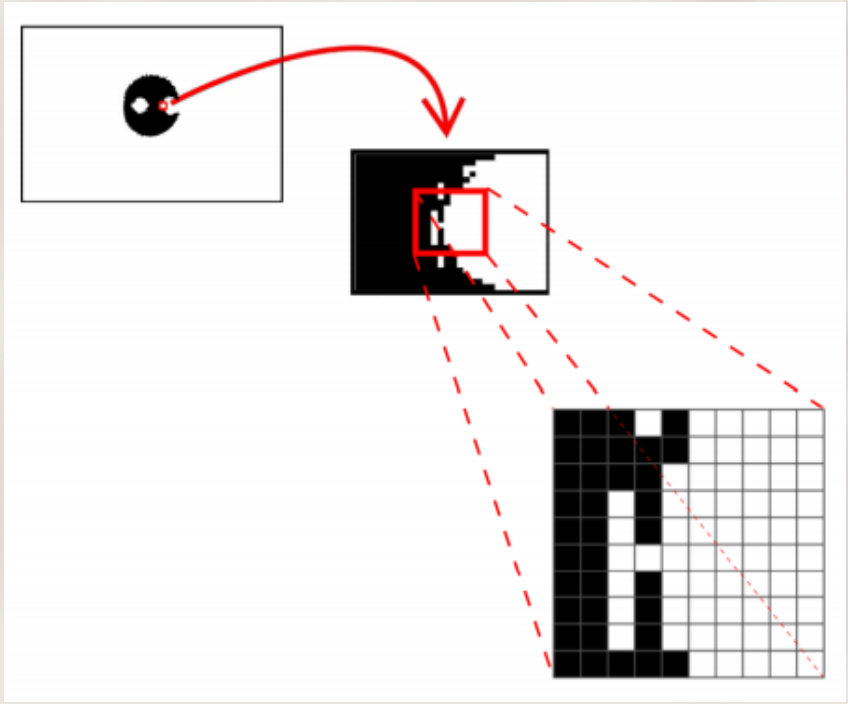
$$GY = 0.299 \cdot R + 0.587 \cdot G + 0.144 \cdot B$$





Progowanie

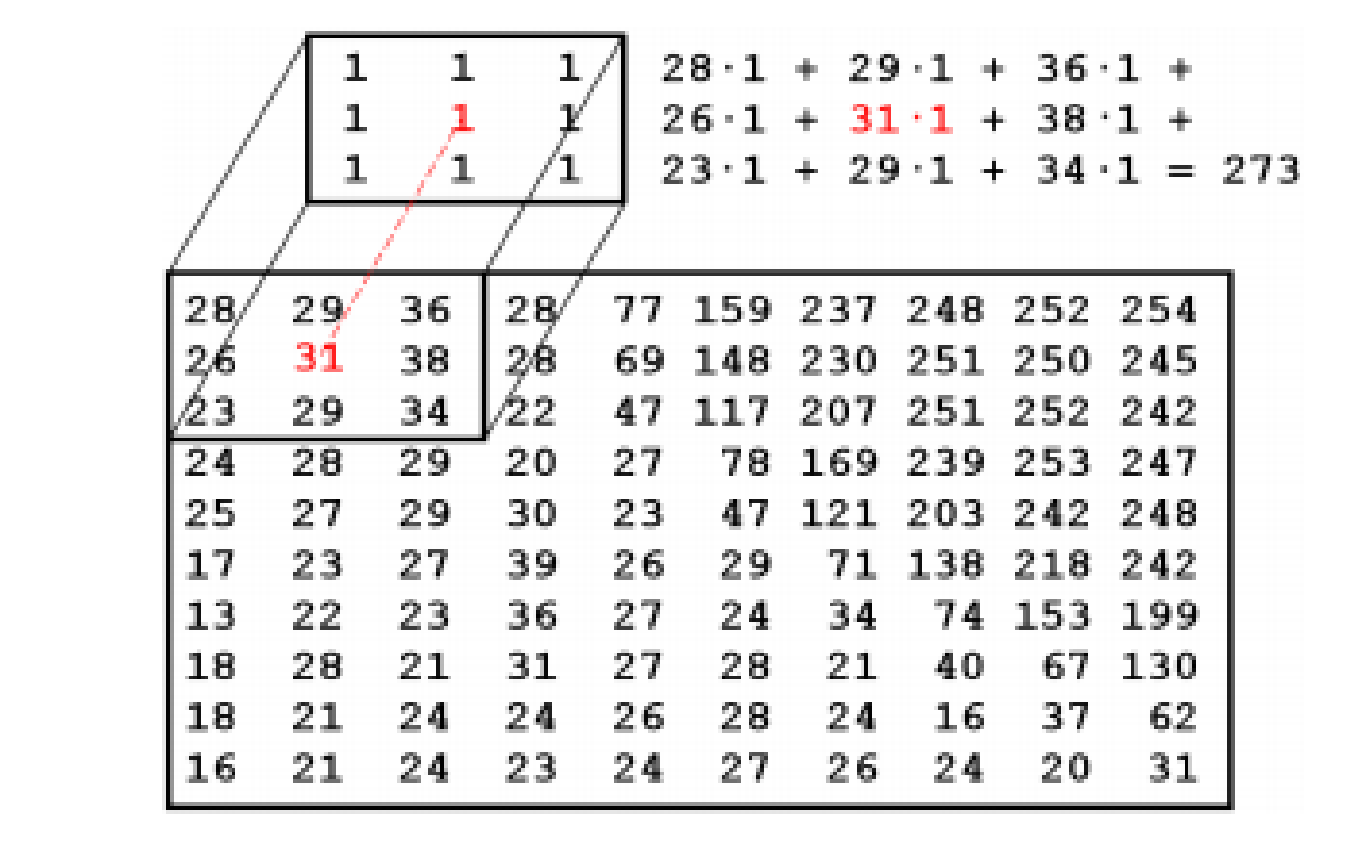

28 29 36 28 77 159 237 248 252 254
26 31 38 28 69 148 230 251 250 245
23 29 34 22 47 117 207 251 252 242
24 28 29 20 27 78 169 239 253 247
25 27 29 30 23 47 121 203 242 248
17 23 27 39 26 29 71 138 218 242
13 22 23 36 27 24 34 74 153 199
18 28 21 31 27 28 21 40 67 130
18 21 24 24 26 28 24 16 37 62
16 21 24 23 24 27 26 24 20 31



Splot (konwolucja)

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Macierz splotu (jądro)
(maska)

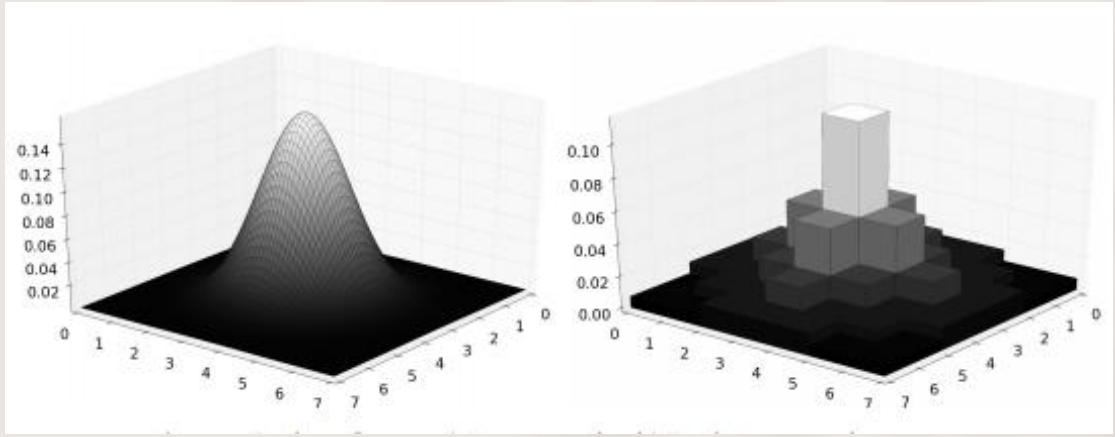


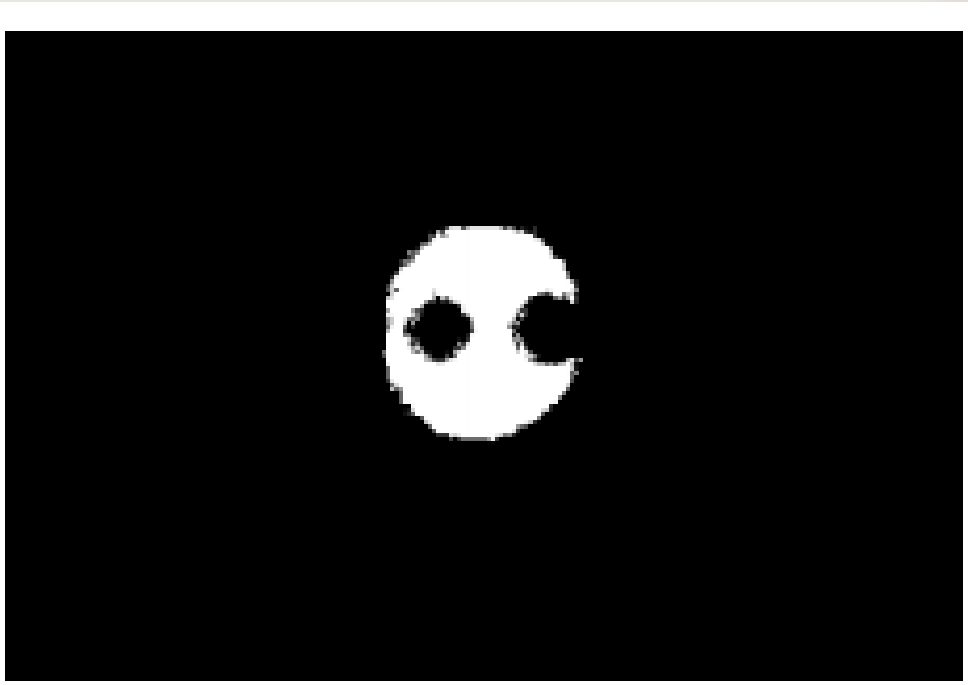
Uśrednienie (filtr wygładzający)

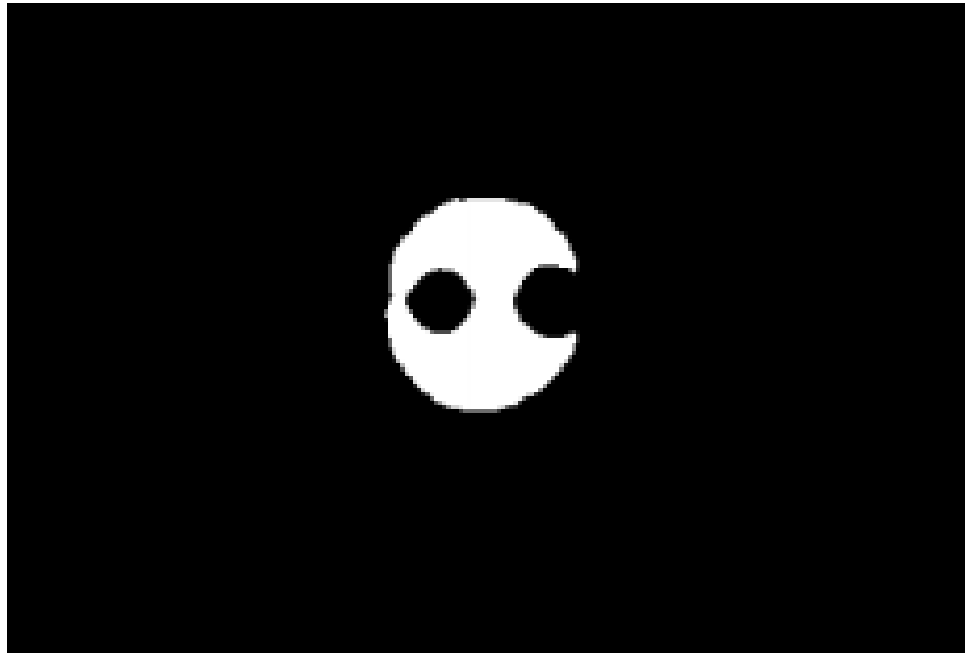
$$\begin{bmatrix} 28 & 29 & 36 & 28 & 77 & 159 & 237 & 248 & 252 & 254 \\ 26 & 31 & 38 & 28 & 69 & 148 & 230 & 251 & 250 & 245 \\ 23 & 29 & 34 & 22 & 47 & 117 & 207 & 251 & 252 & 242 \\ 24 & 28 & 29 & 20 & 27 & 78 & 169 & 239 & 253 & 247 \\ 25 & 27 & 29 & 30 & 23 & 47 & 121 & 203 & 242 & 248 \\ 17 & 23 & 27 & 39 & 26 & 29 & 71 & 138 & 218 & 242 \\ 13 & 22 & 23 & 36 & 27 & 24 & 34 & 74 & 153 & 199 \\ 18 & 28 & 21 & 31 & 27 & 28 & 21 & 40 & 67 & 130 \\ 18 & 21 & 24 & 24 & 26 & 28 & 24 & 16 & 37 & 62 \\ 16 & 21 & 24 & 23 & 24 & 27 & 26 & 24 & 20 & 31 \end{bmatrix} * \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 27 & 32 & 33 & 48 & 86 & 154 & 212 & 245 & 250 & 248 \\ 26 & 30 & 31 & 42 & 77 & 143 & 205 & 242 & 249 & 244 \\ 25 & 29 & 29 & 35 & 62 & 121 & 188 & 234 & 248 & 241 \\ 25 & 28 & 28 & 29 & 46 & 93 & 159 & 215 & 242 & 242 \\ 23 & 25 & 28 & 28 & 35 & 66 & 122 & 184 & 226 & 241 \\ 22 & 23 & 28 & 29 & 30 & 32 & 51 & 91 & 140 & 183 \\ 20 & 21 & 28 & 29 & 30 & 32 & 51 & 91 & 140 & 183 \\ 20 & 21 & 26 & 27 & 28 & 27 & 32 & 52 & 86 & 127 \\ 19 & 21 & 24 & 25 & 26 & 26 & 26 & 31 & 47 & 74 \\ 18 & 20 & 22 & 23 & 24 & 25 & 25 & 25 & 29 & 41 \end{bmatrix}$$

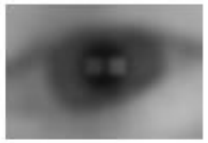
Filtr gaussowski 7x7



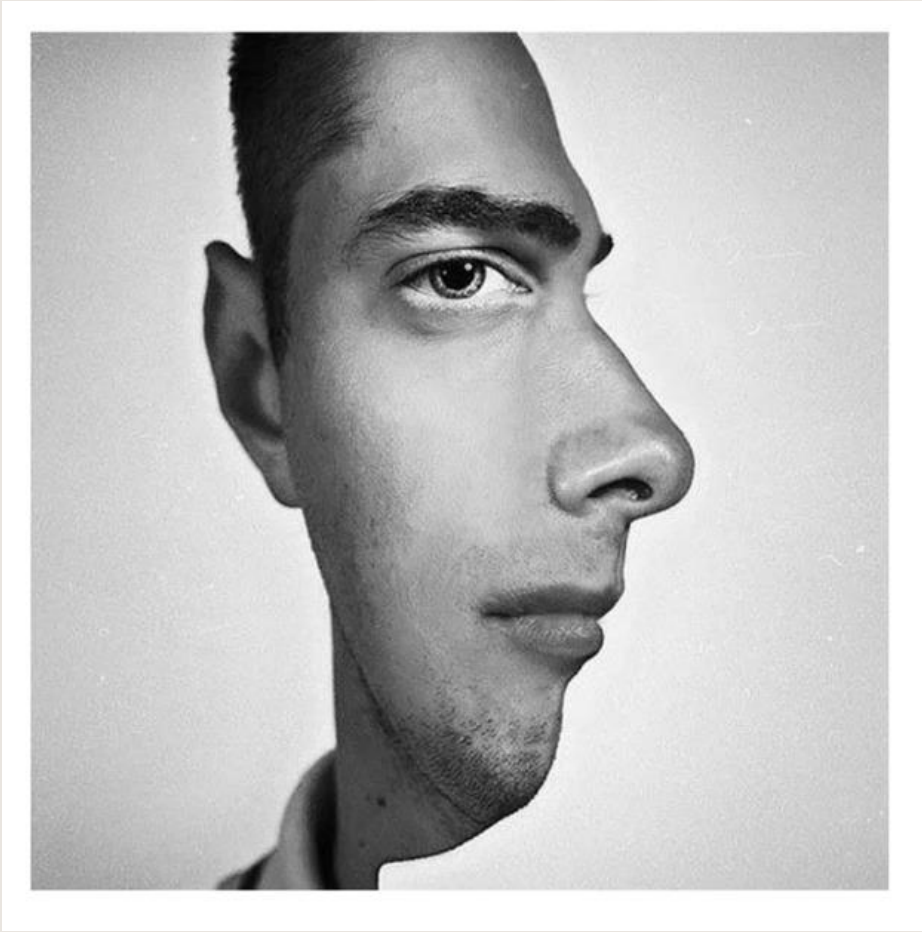




Zestawienie

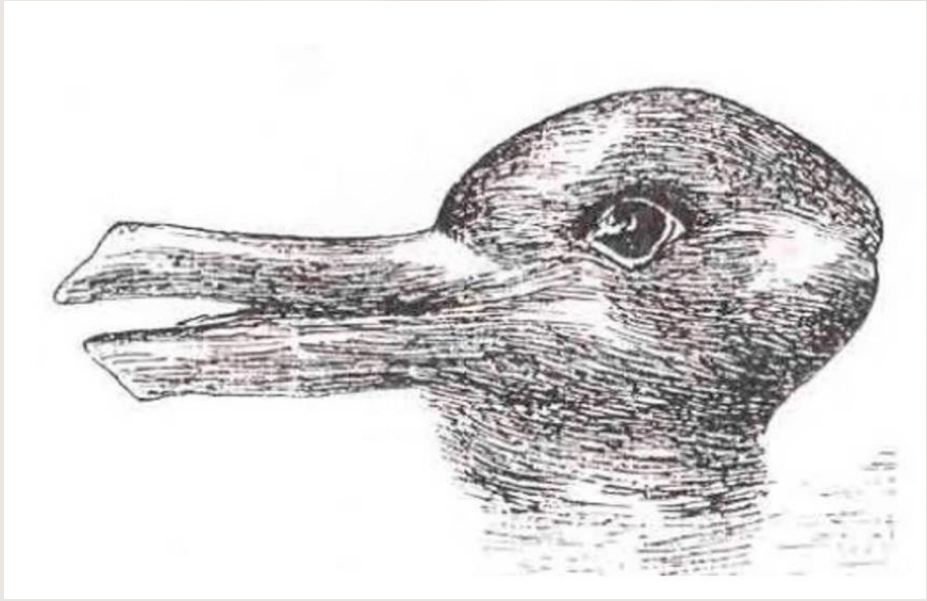
Filtr uśredniający (ang. *mean filter*).		
		
Maska 3x3	Maska 17x17	Maska 33x33
Filtr gaussowski (ang. *Gaussian filter*).		
		
Maska 3x3	Maska 17x17	Maska 33x33
Filtr medianowy (ang. *median filter*).		
		
Maska 3x3	Maska 17x17	Maska 33x33
Filtr bilateralny (ang. bilateral filter).		
		
Maska 3x3	Maska 17x17	Maska 33x33

Sieci splotowe

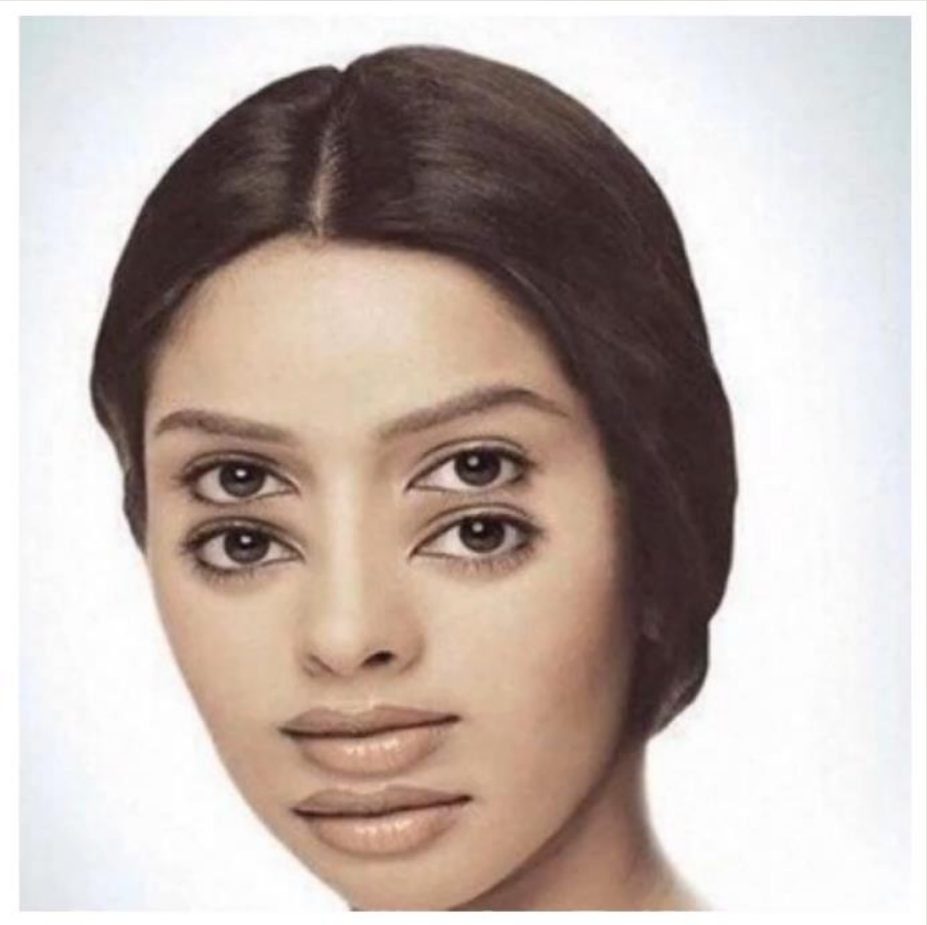


Sieci splotowe

Sieci splotowe



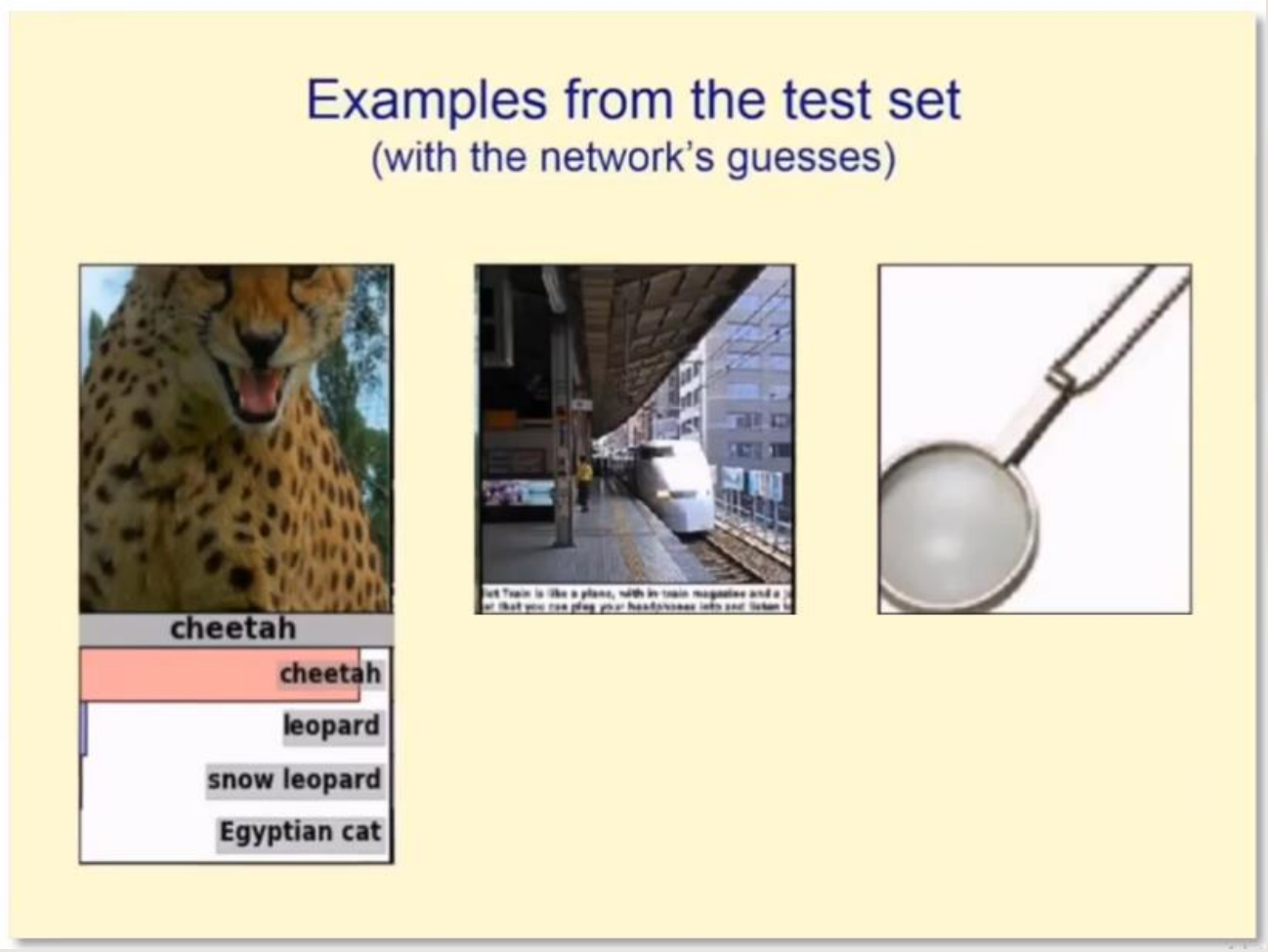
Sieci splotowe



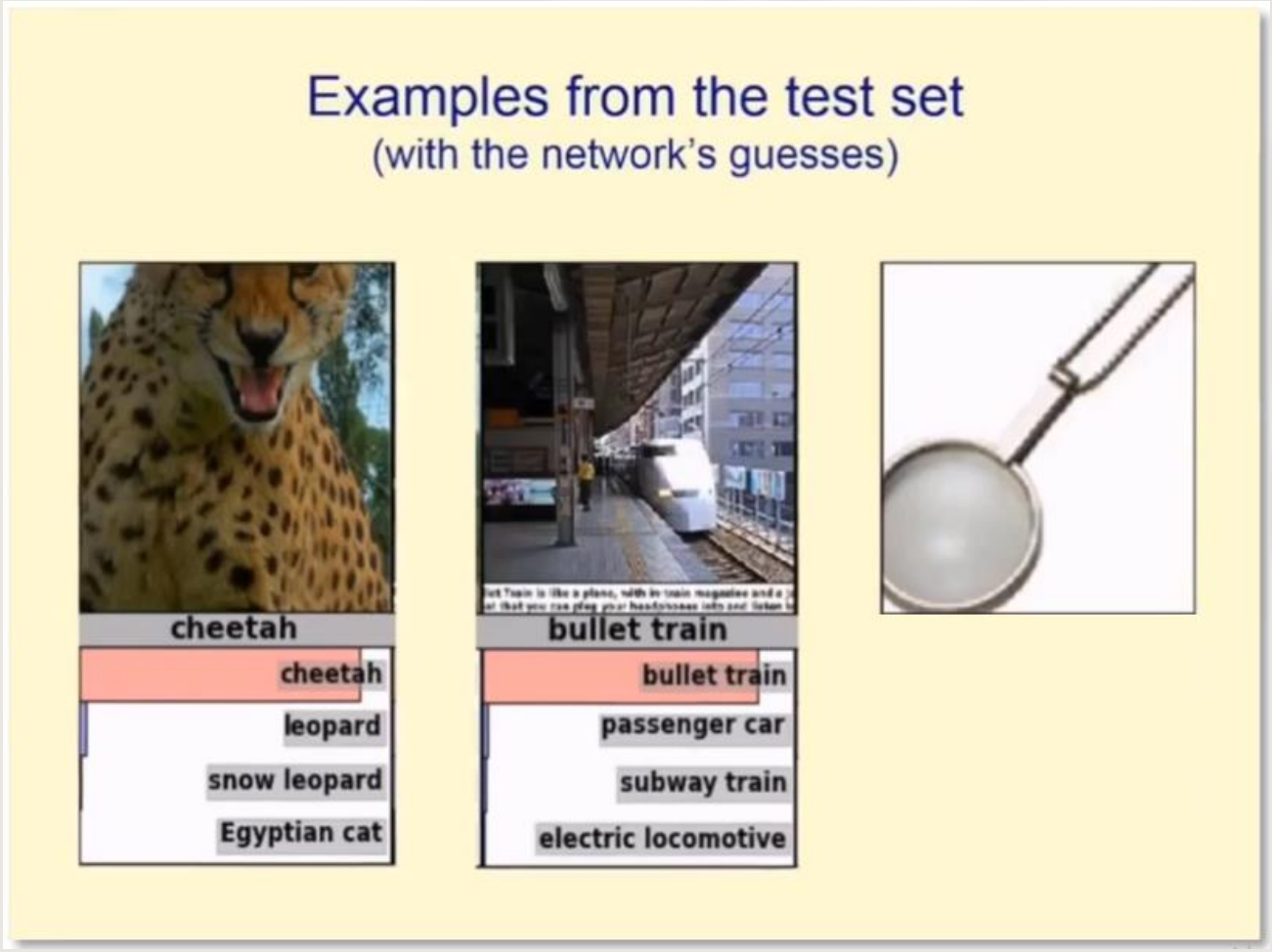
Sieci splotowe

Examples from the test set
(with the network's guesses)

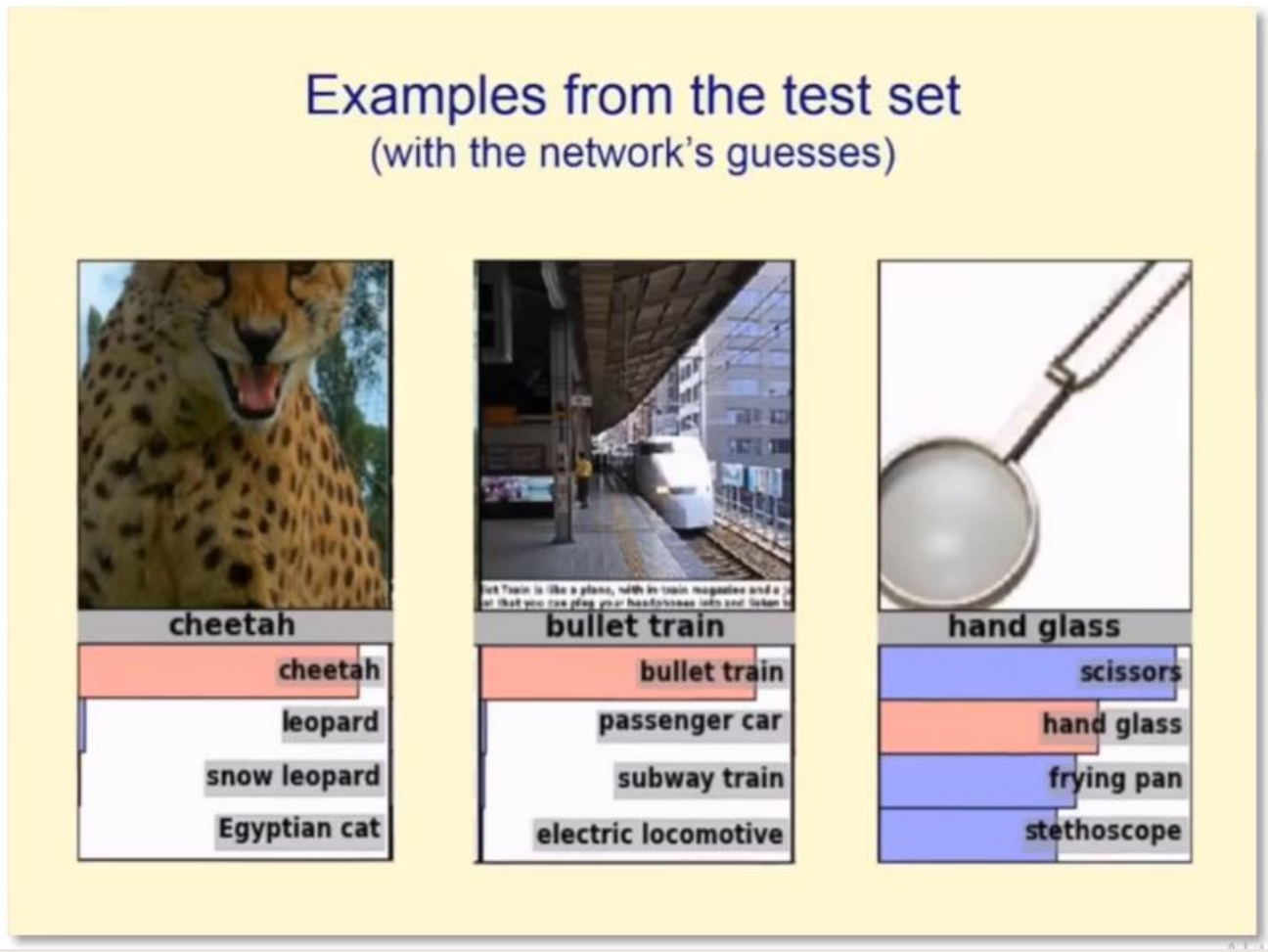
Sieci splotowe



Sieci splotowe



Sieci splotowe

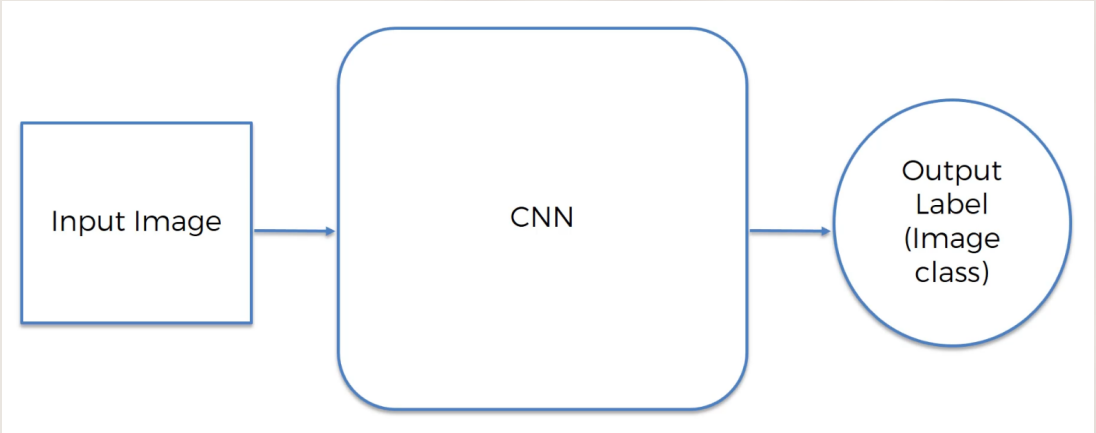


Sieci splotowe

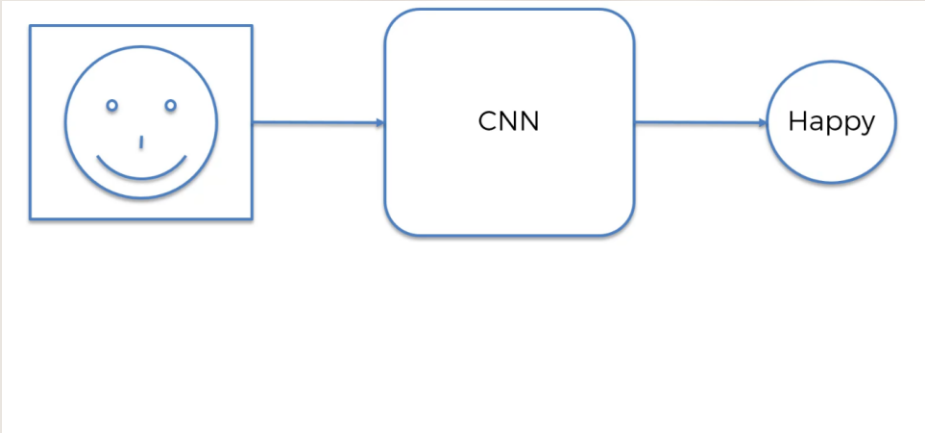
Sieci splotowe

Geoffrey Hinton. Yann LeCun

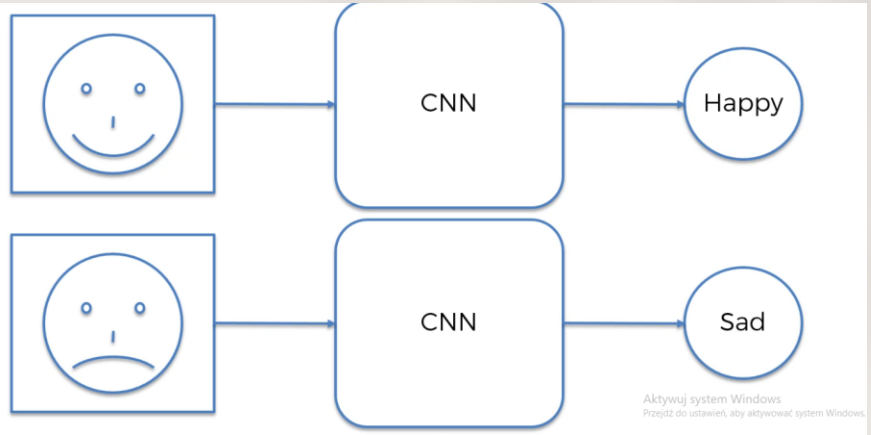
Sieci splotowe



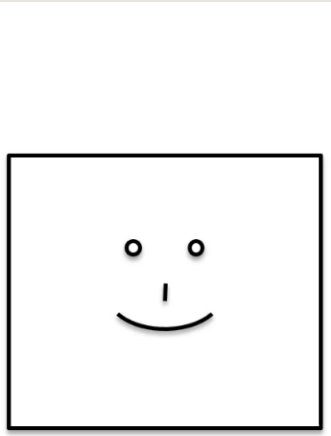
Sieci splotowe

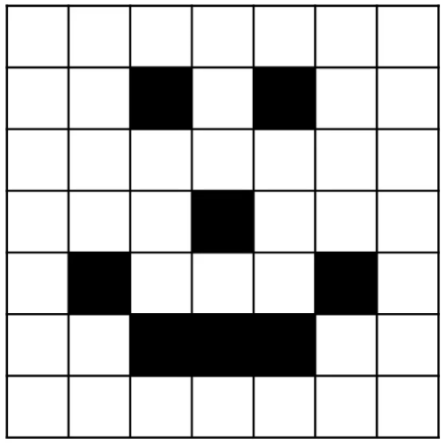


Sieci splotowe



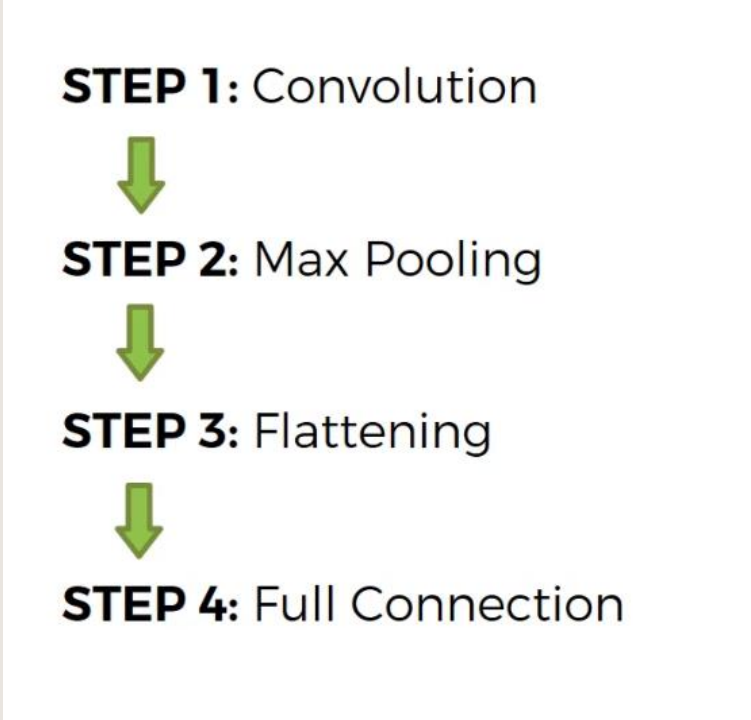
Sieci splotowe





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

Sieci splotowe

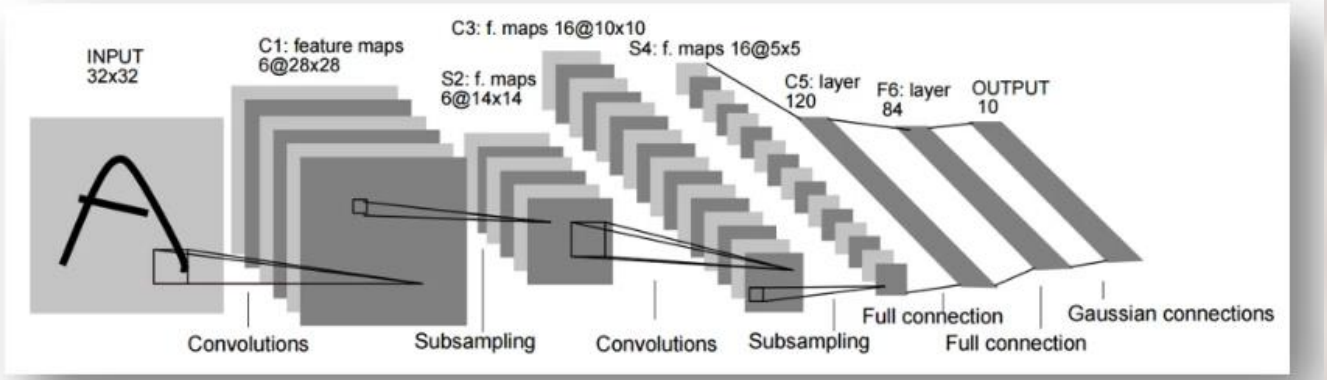


Sieci splotowe

Additional Reading:

Gradient-Based Learning Applied to Document Recognition

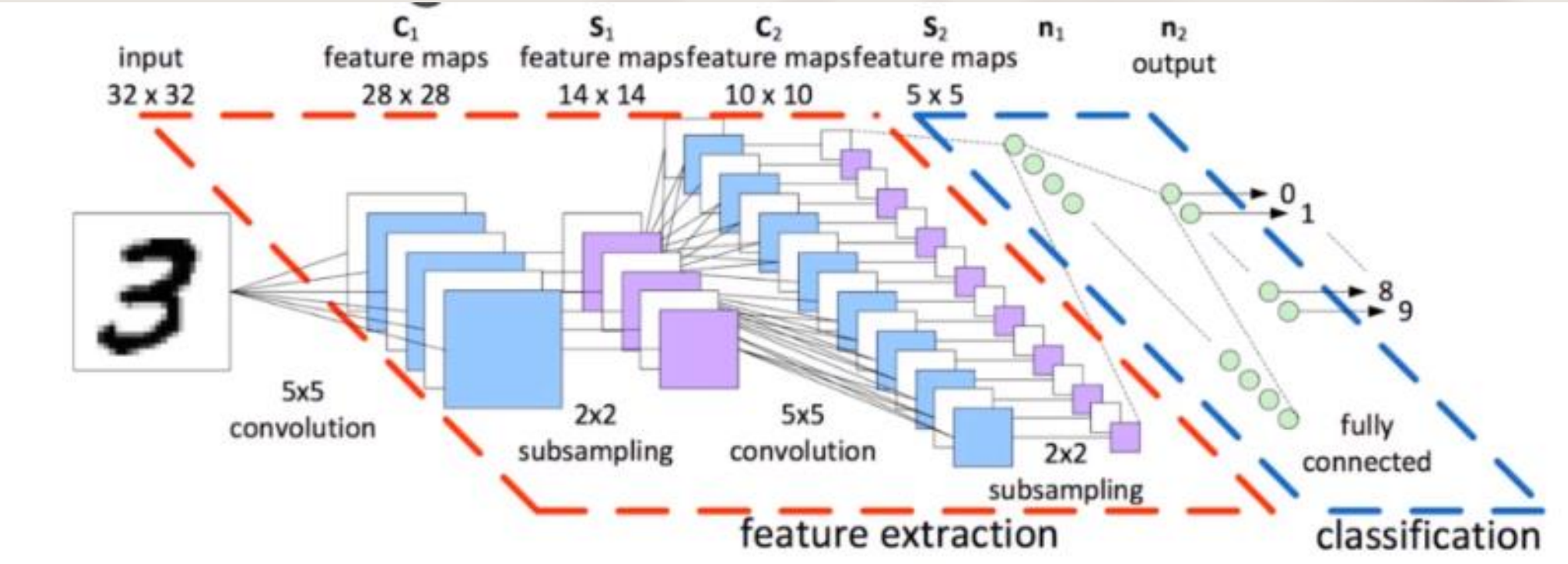
By Yann LeCun et al. (1998)



Link:

http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf

Sieci splotowe



Sieci splotowe

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

Sieci splotowe

Additional Reading:

Introduction to Convolutional Neural Networks

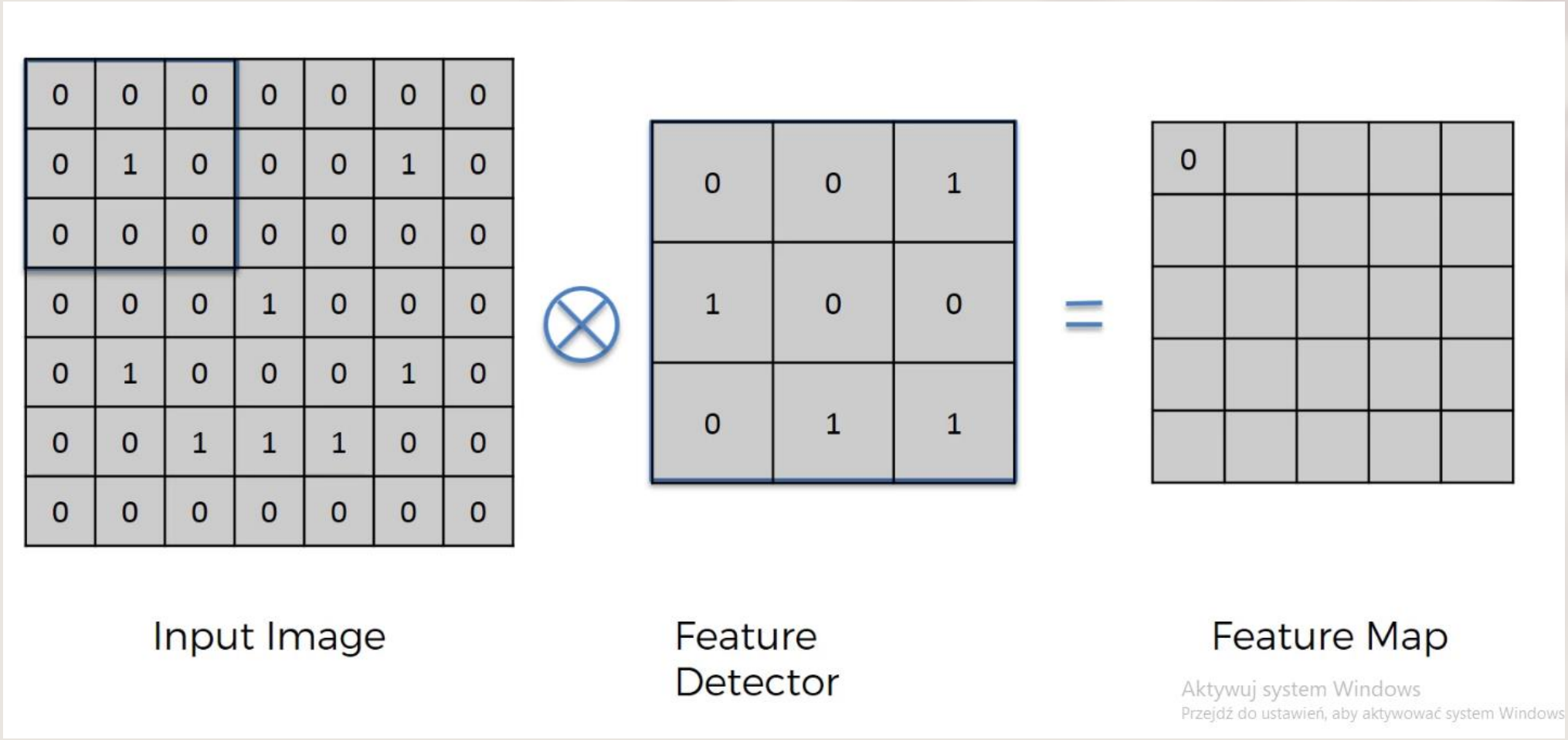
By Jianxin Wu (2017)

$$\begin{aligned}\frac{\partial z}{\partial(\mathrm{vec}(\boldsymbol{y})^T)}(F^T \otimes I) &= \left((F \otimes I)\frac{\partial z}{\partial\,\mathrm{vec}(\boldsymbol{y})}\right)^T \\ &= \left((F \otimes I)\,\mathrm{vec}\left(\frac{\partial z}{\partial Y}\right)\right)^T \\ &= \mathrm{vec}\left(I\frac{\partial z}{\partial Y}F^T\right)^T \\ &= \mathrm{vec}\left(\frac{\partial z}{\partial Y}F^T\right)^T,\end{aligned}$$

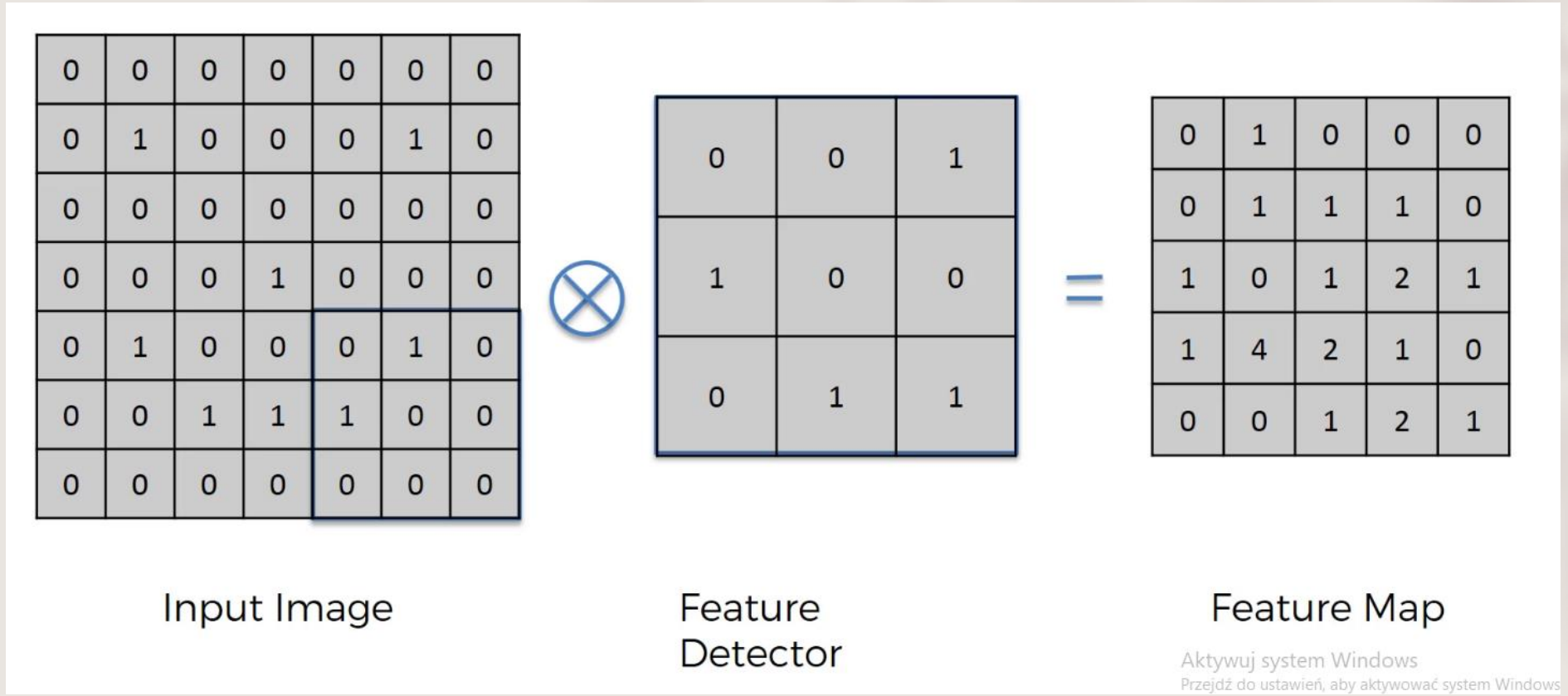
Link:

http://cs.nju.edu.cn/wujx/paper/CNN.pdf

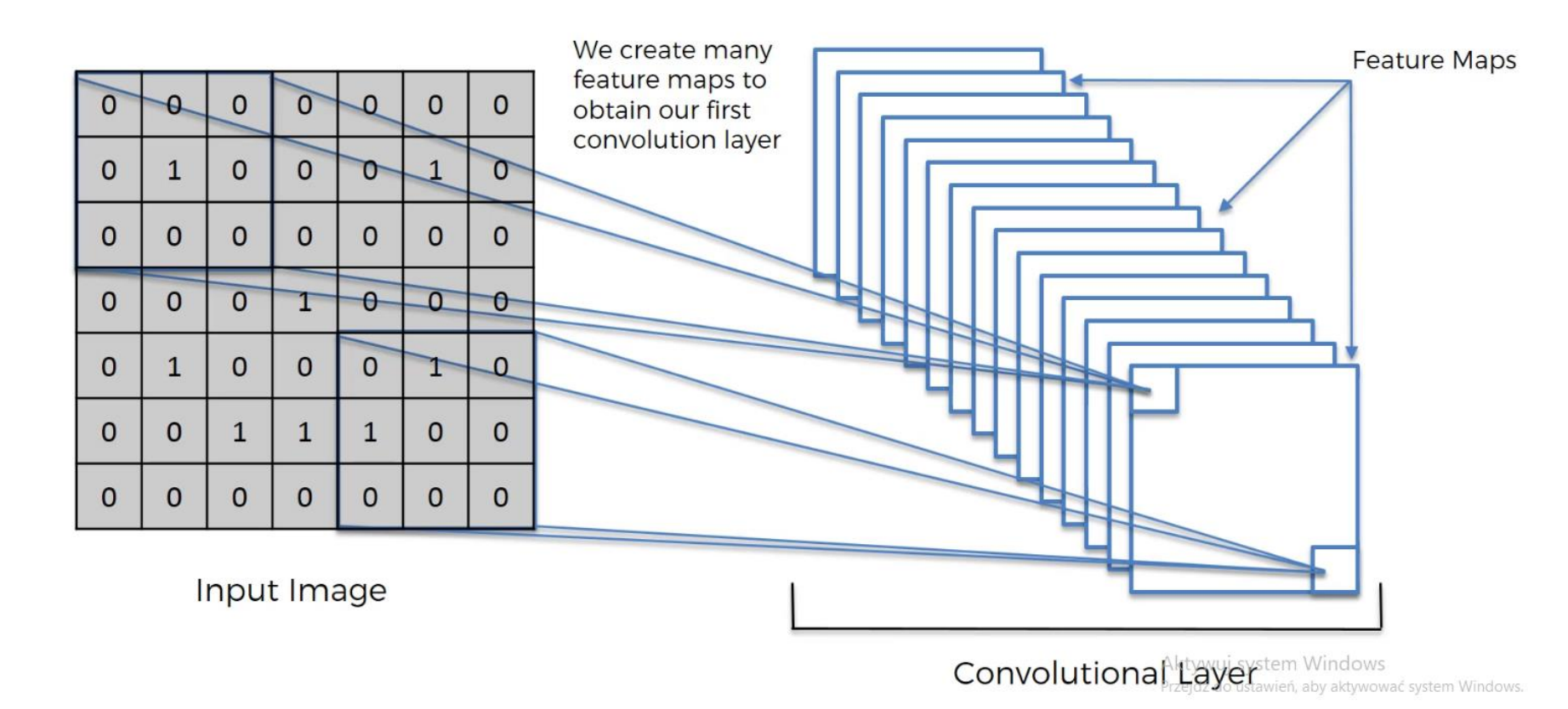
Sieci splotowe



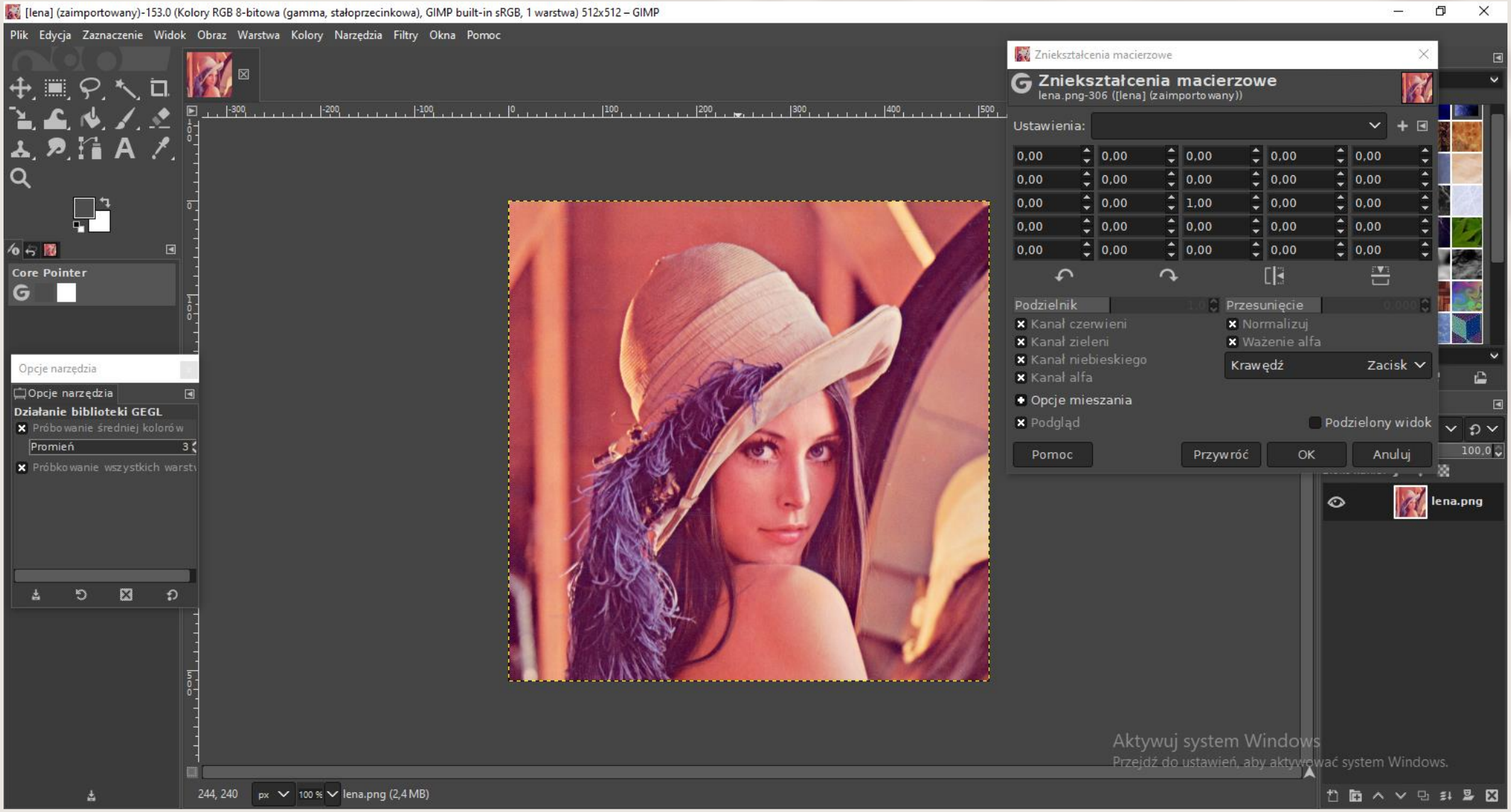
Sieci splotowe



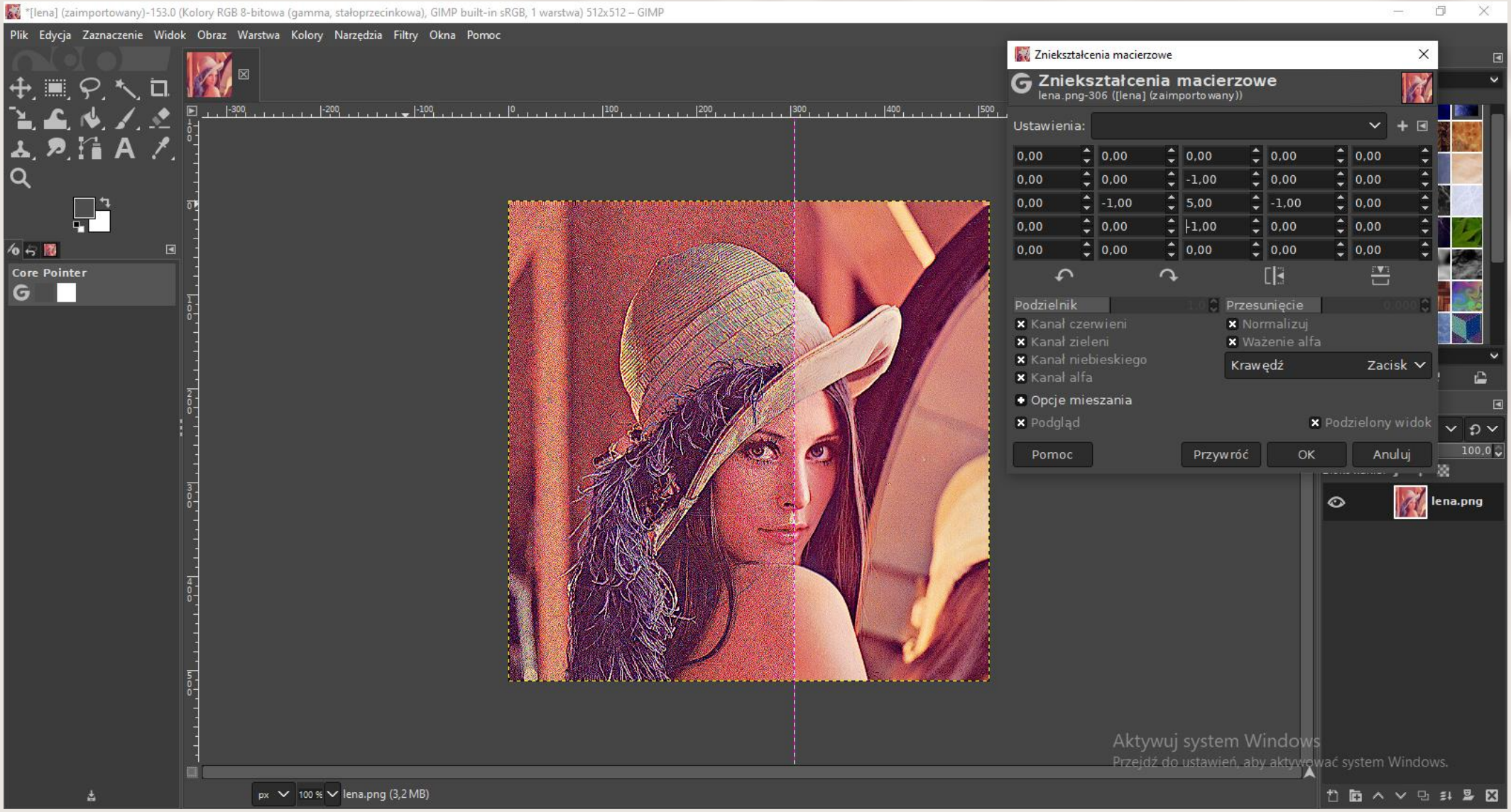
Sieci splotowe



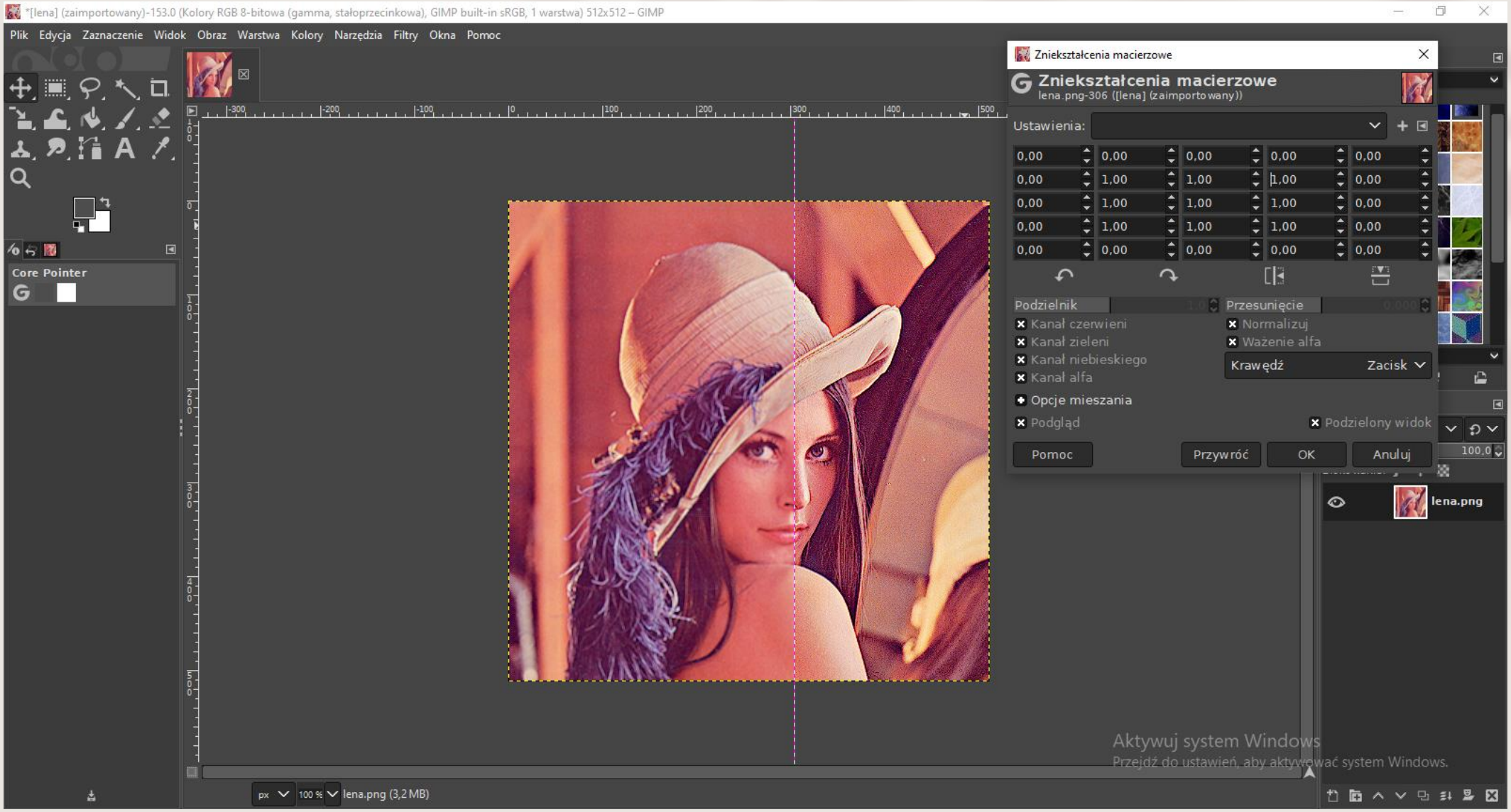
Sieci splotowe



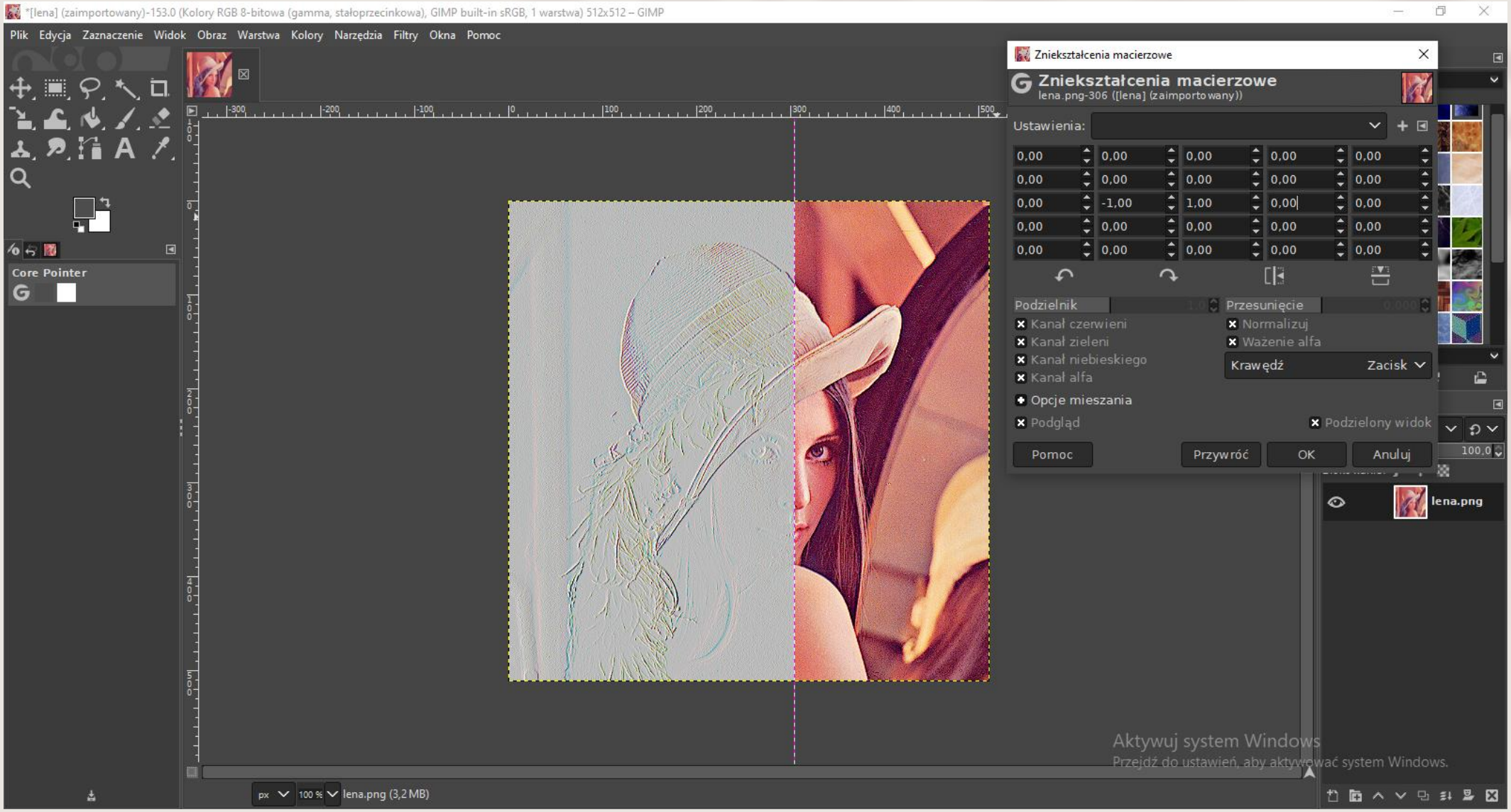
Sieci splotowe



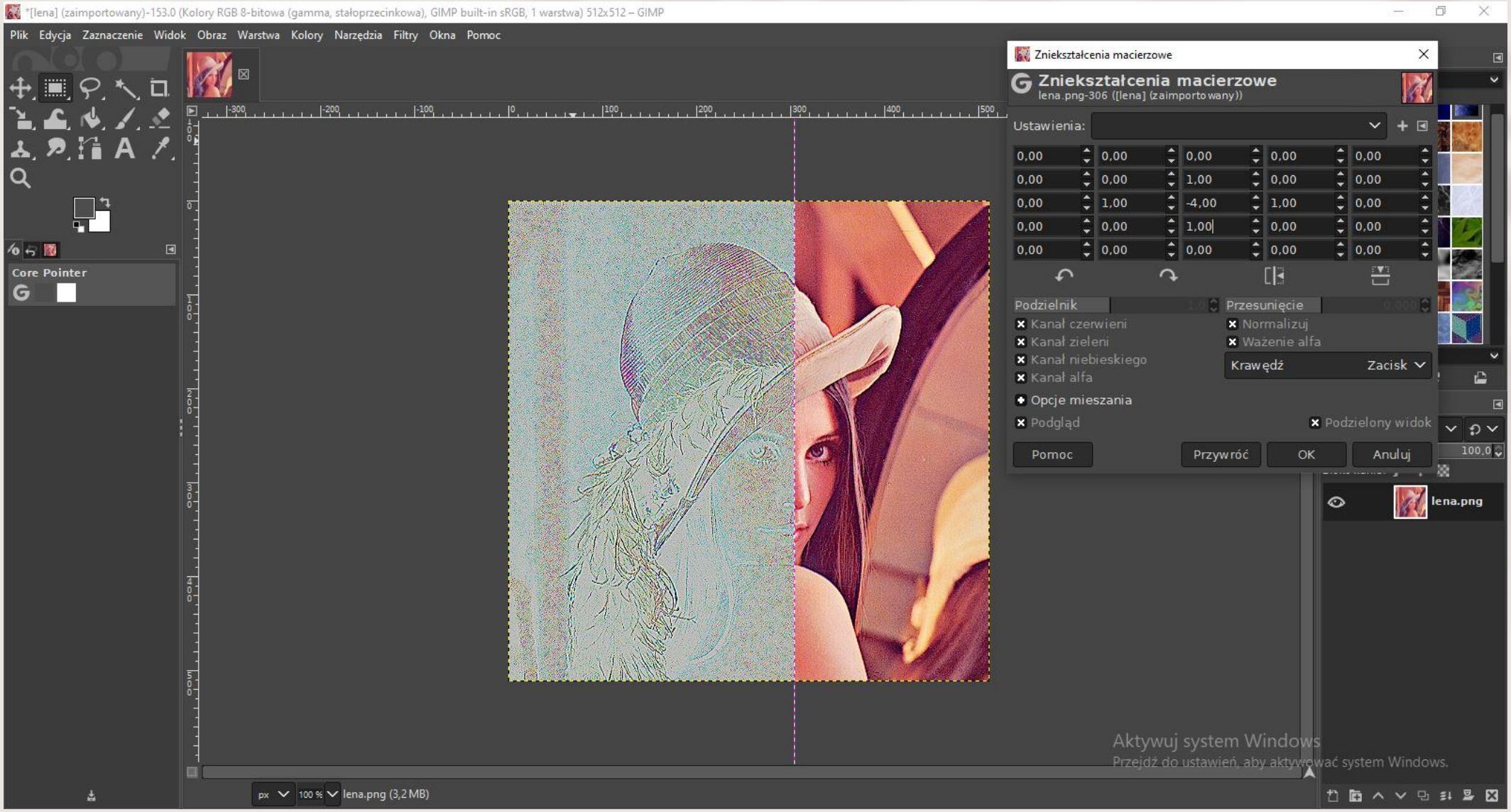
Sieci splotowe



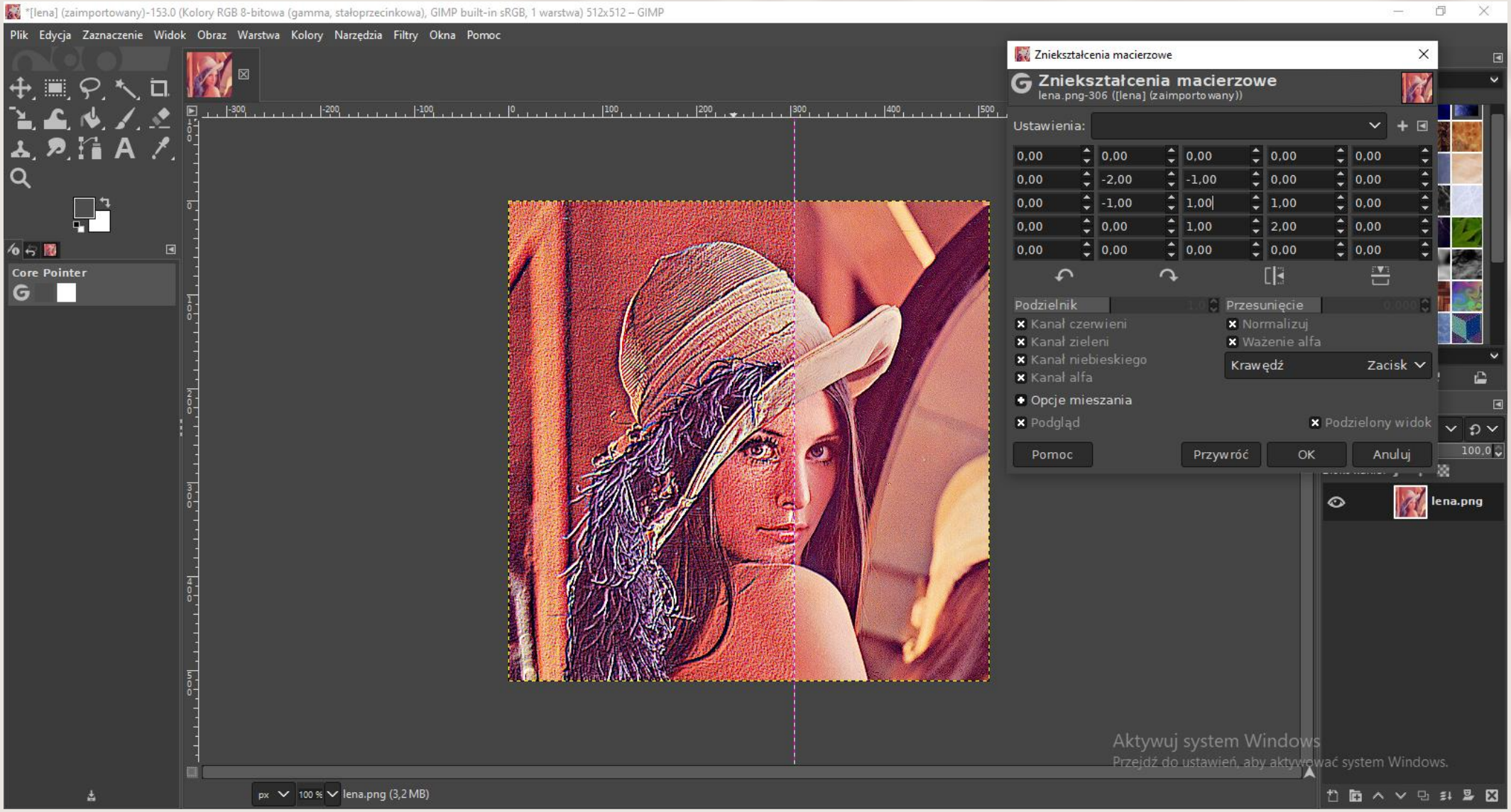
Sieci splotowe



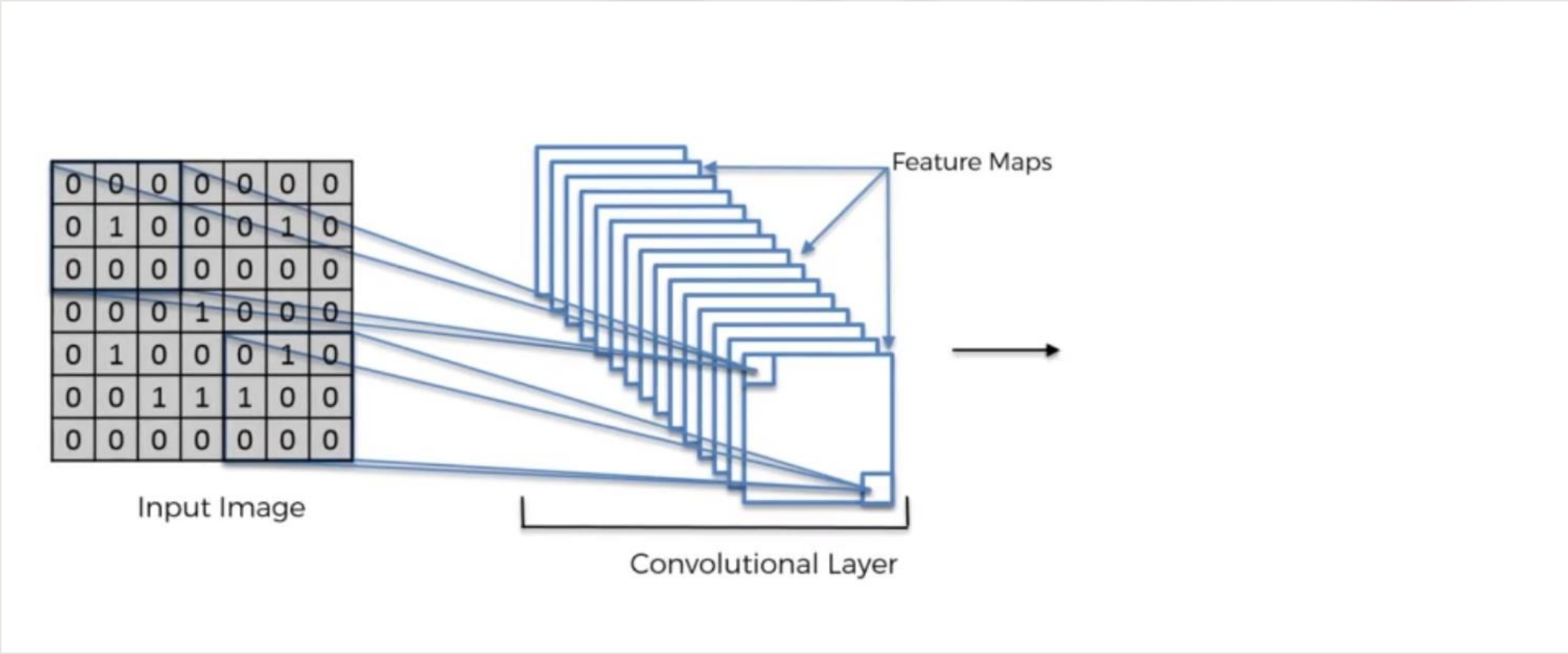
Sieci splotowe



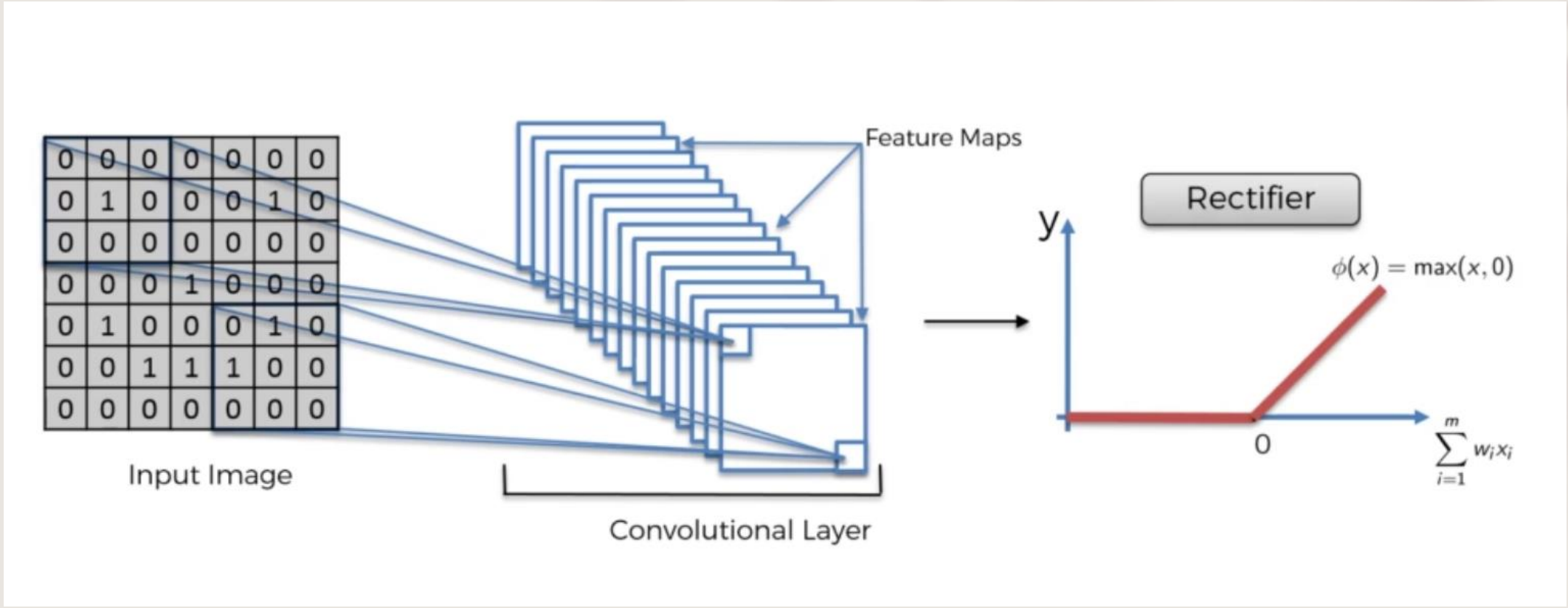
Sieci splotowe



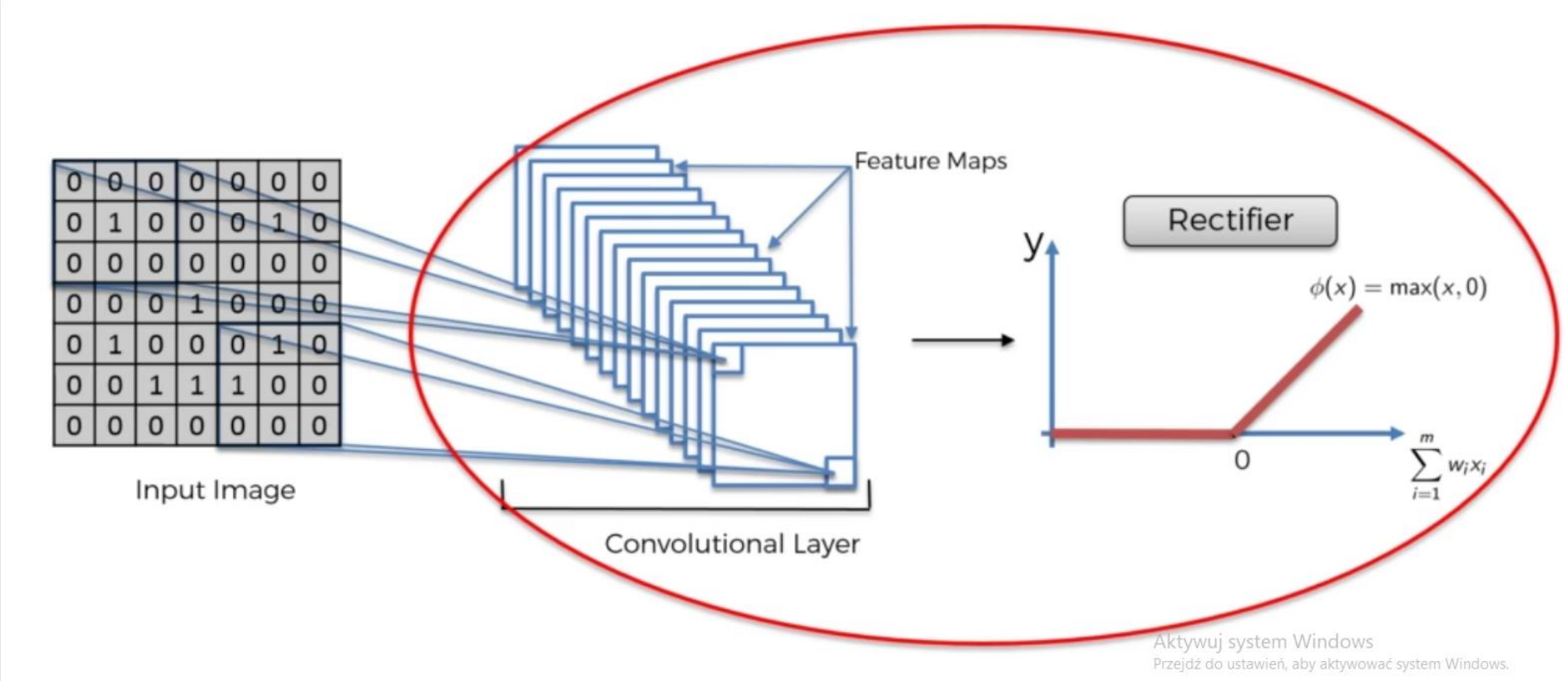
Sieci splotowe



Sieci splotowe

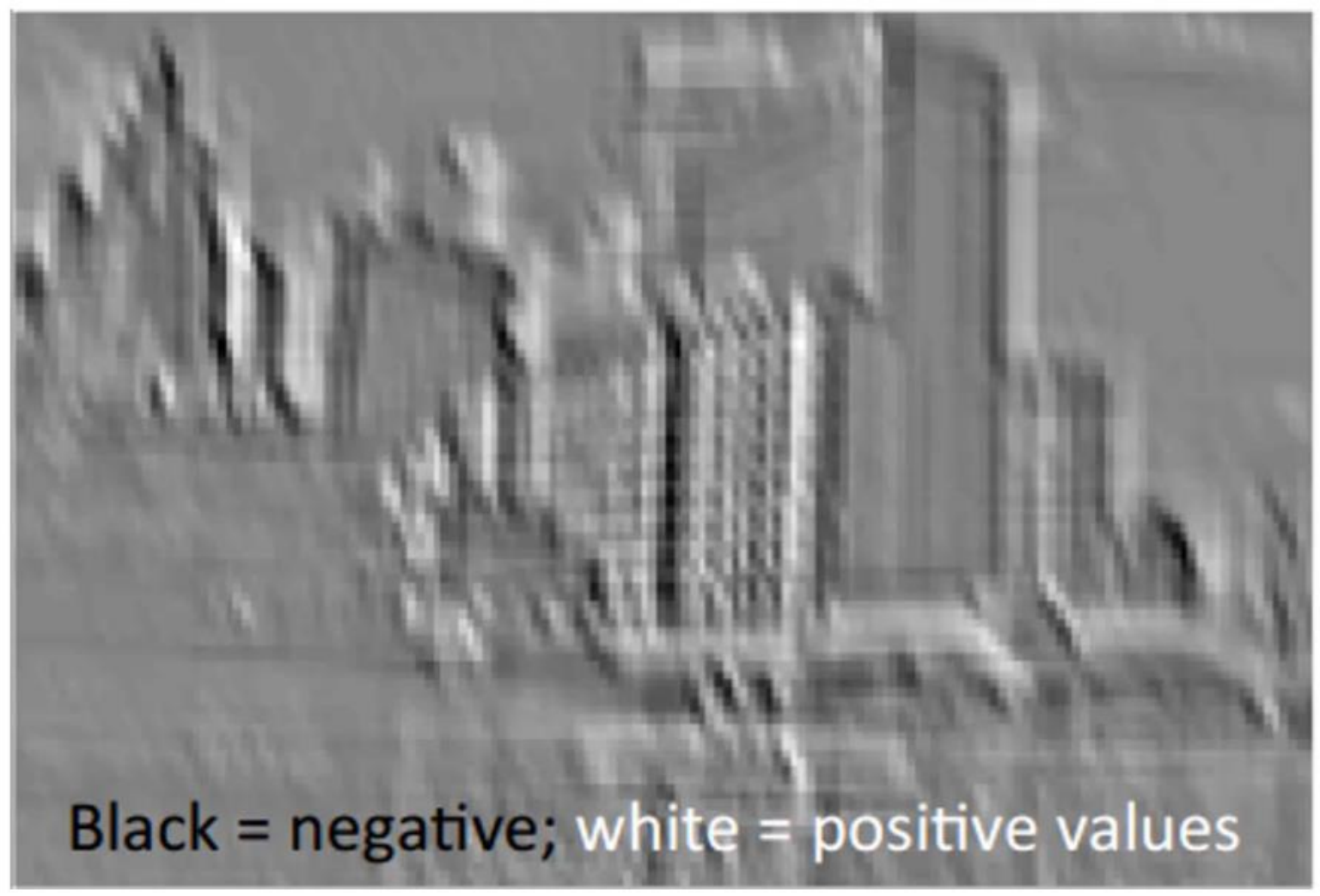


Sieci splotowe



Sieci splotowe

Sieci splotowe



Sieci splotowe

Sieci splotowe

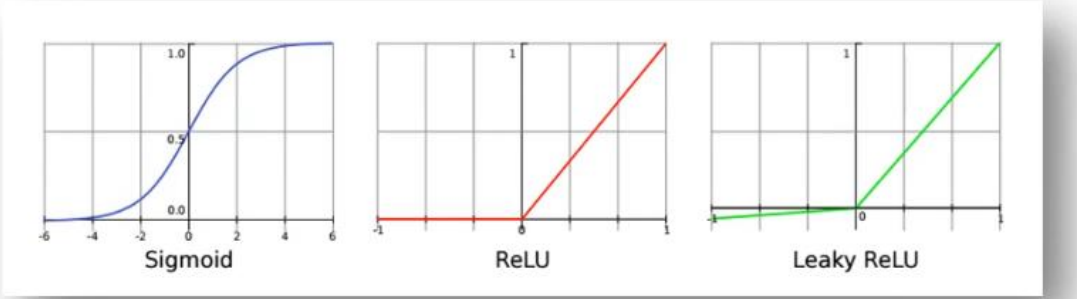
Additional Reading:

Understanding Convolutional Neural Networks with A Mathematical Model

By C.-C. Jay Kuo (2016)

Link:

https://arxiv.org/pdf/1609.04112.pdf



Sieci splotowe

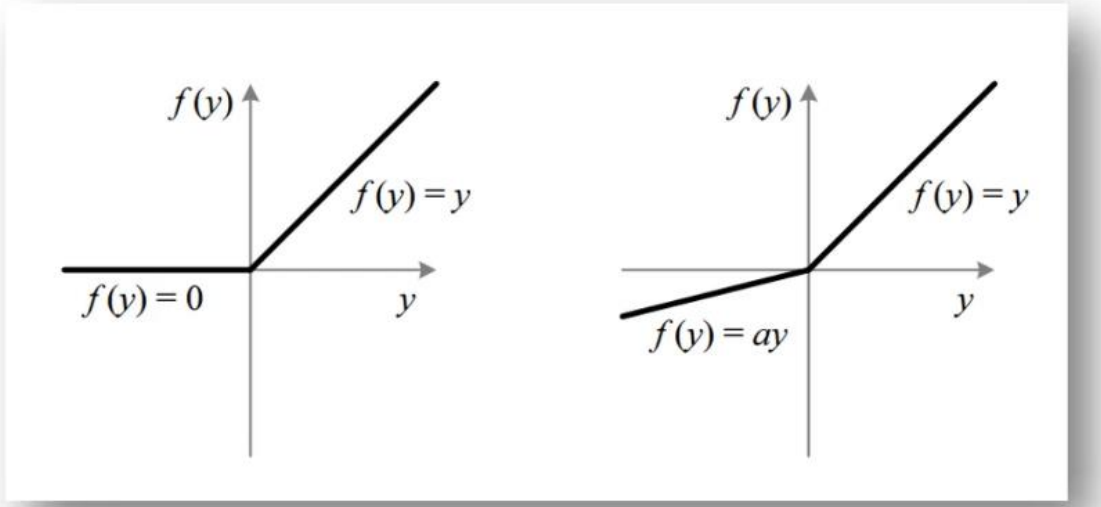
Additional Reading:

Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification

By Kaiming He et al. (2015)

Link:

https://arxiv.org/pdf/1502.01852.pdf



Sieci splotowe

Sieci splotowe

Sieci splotowe

Sieci splotowe

Sieci splotowe

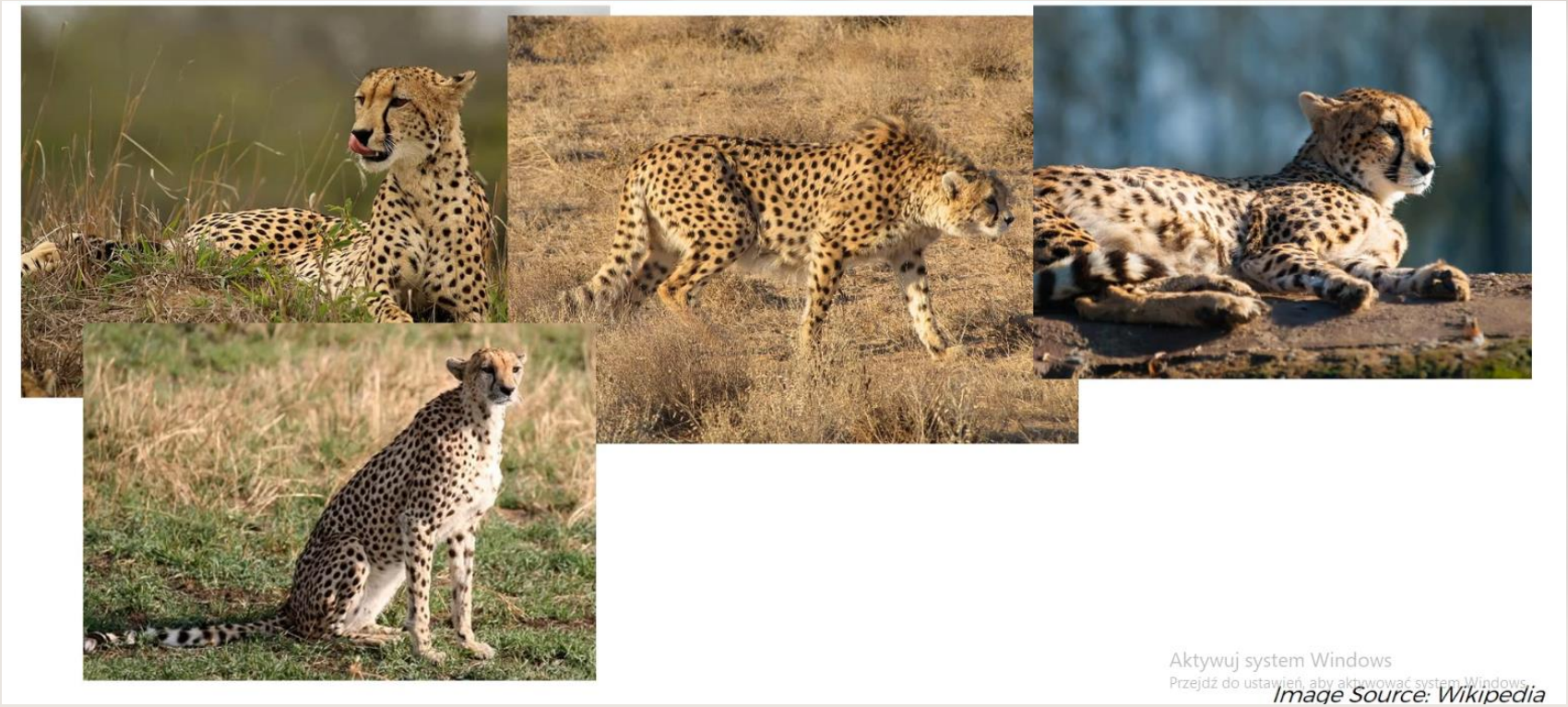


Image Source: Wikipedia

Sieci splotowe

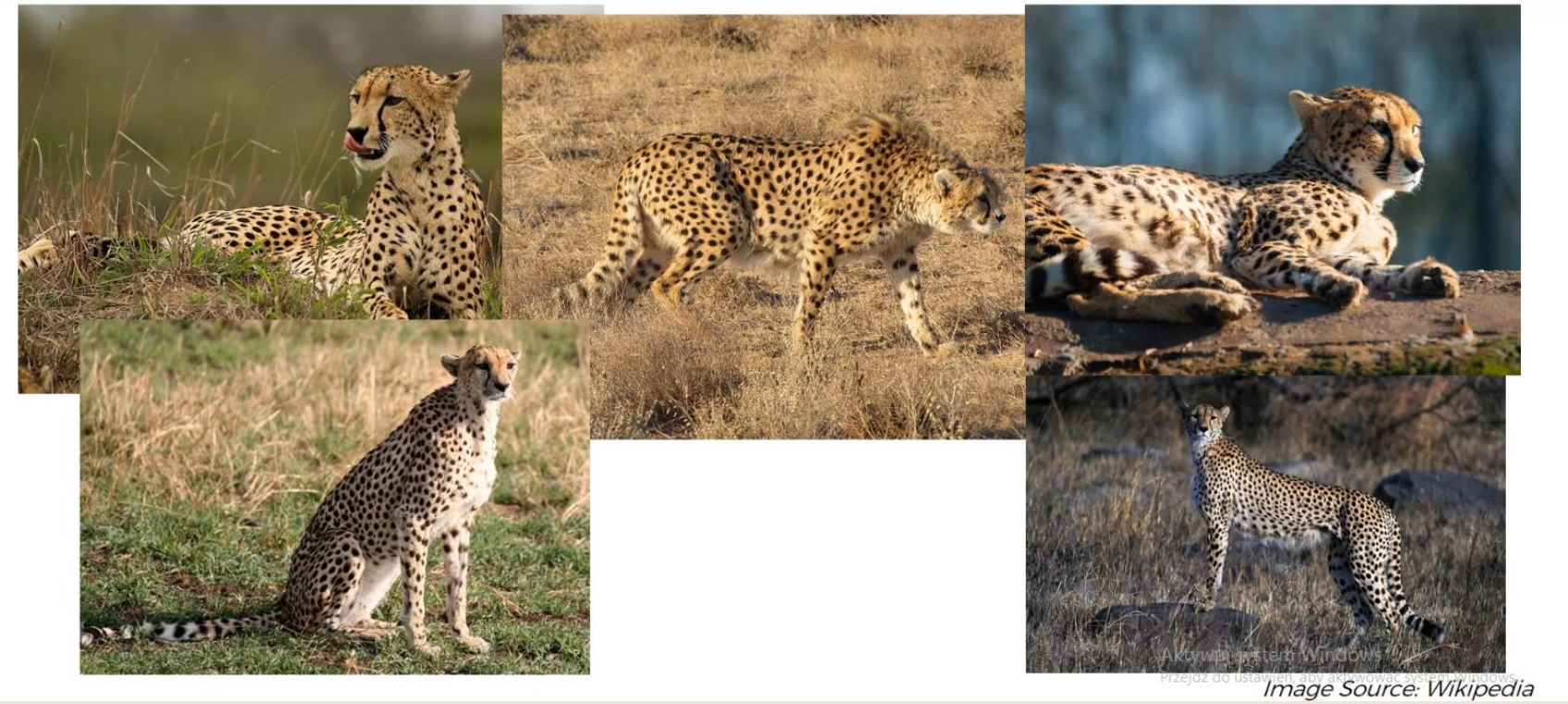


Image Source: Wikipedia

Sieci splotowe

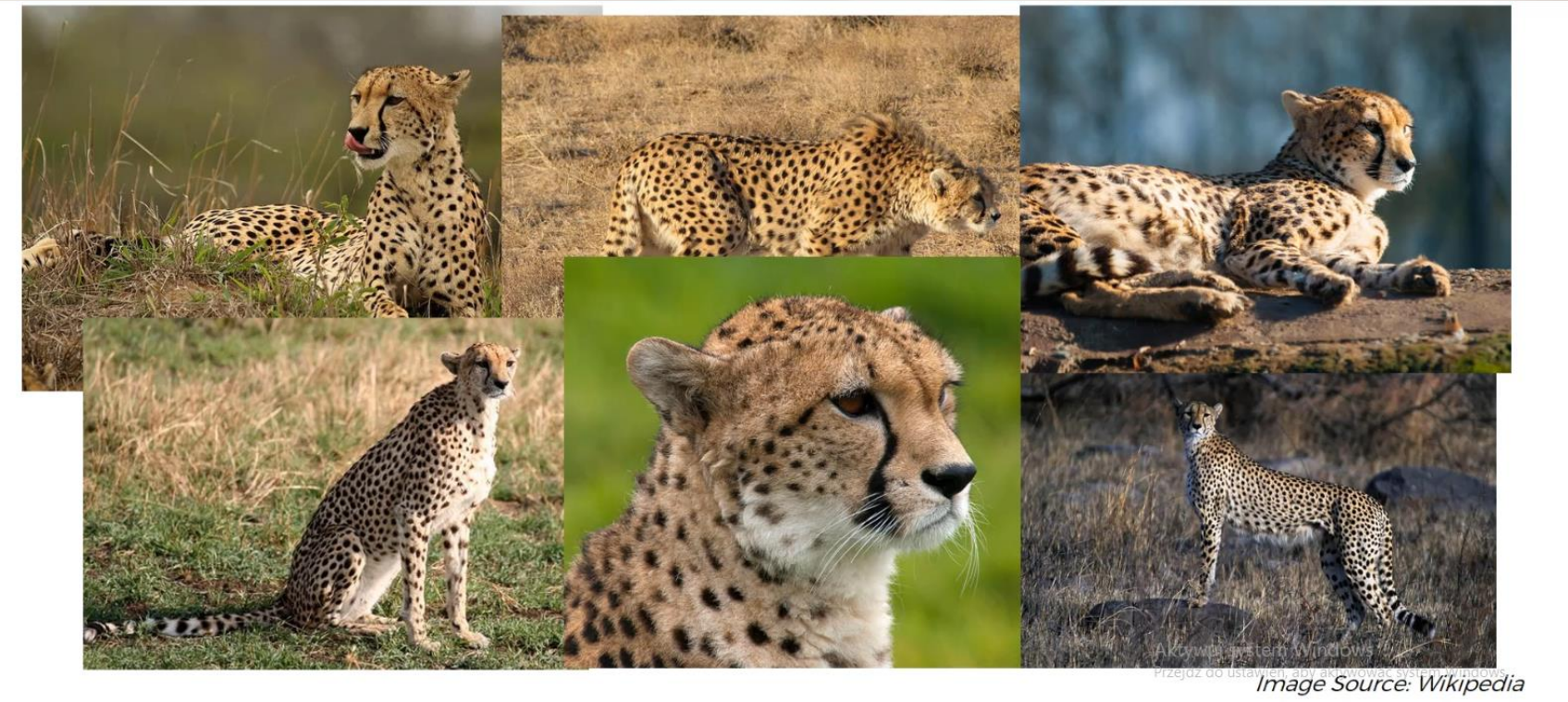


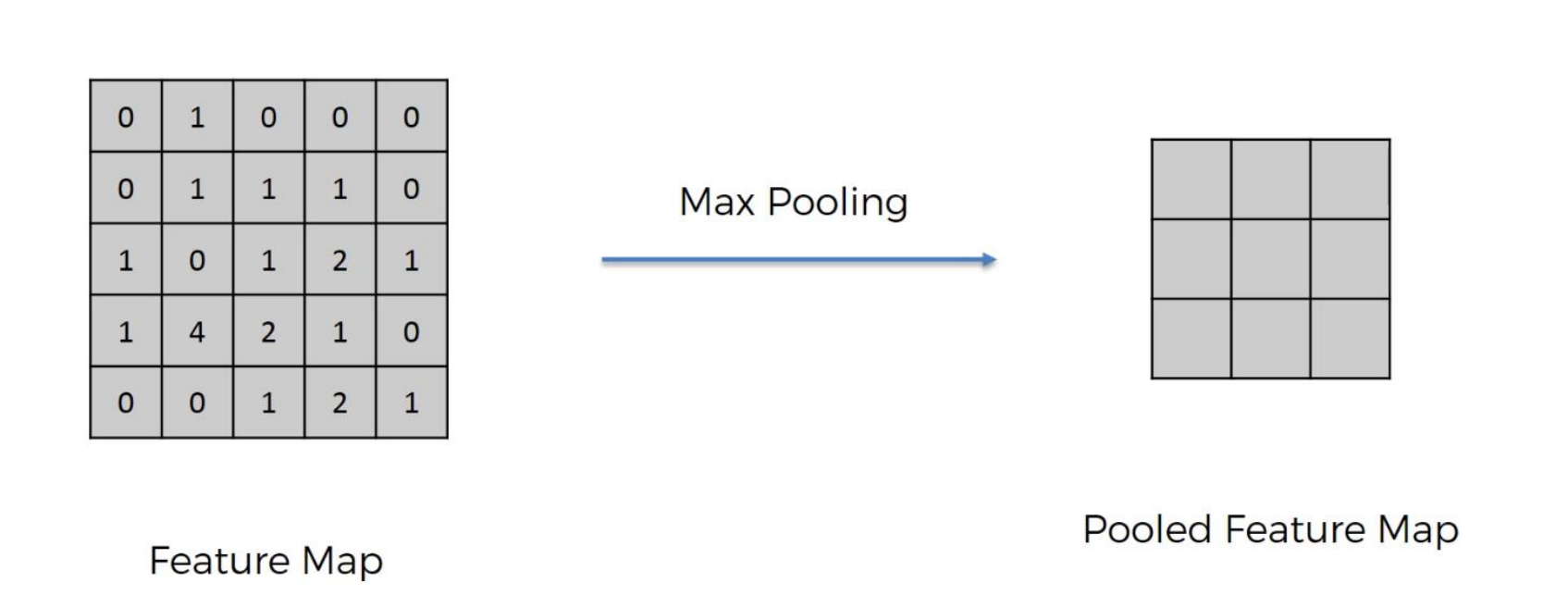
Image Source: Wikipedia

Sieci splotowe

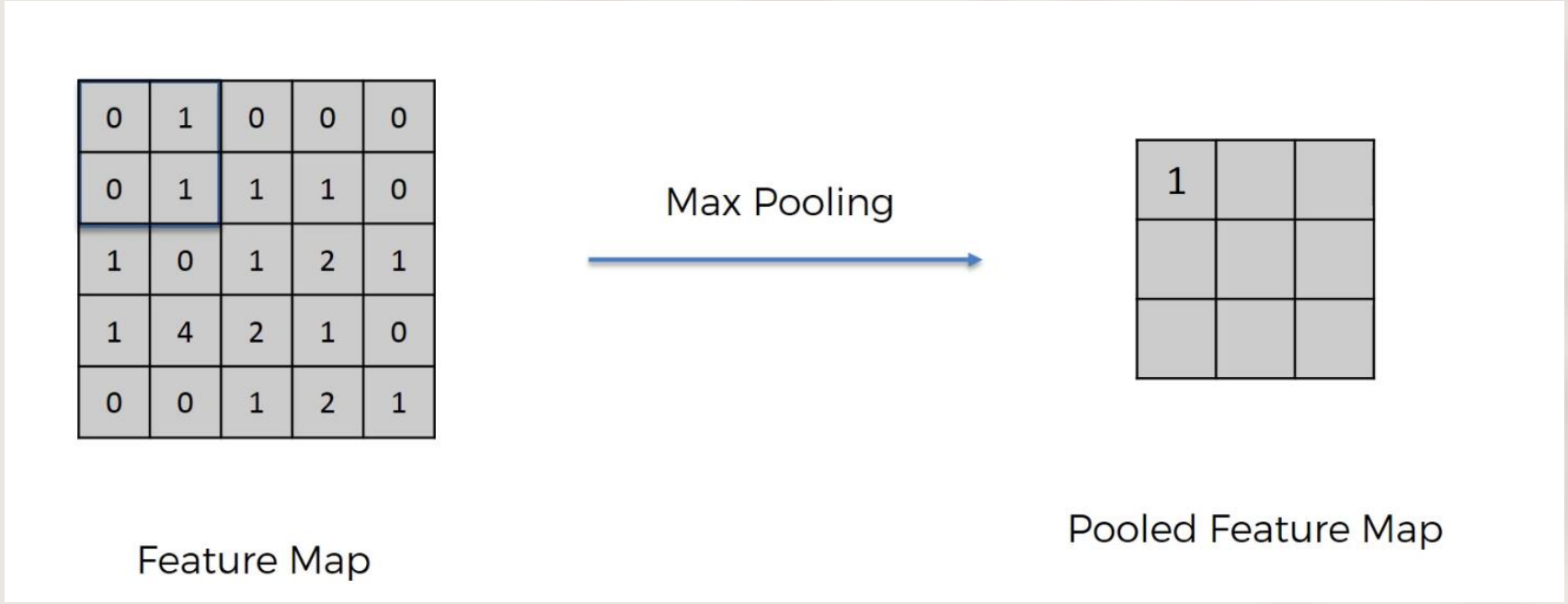
0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Feature Map

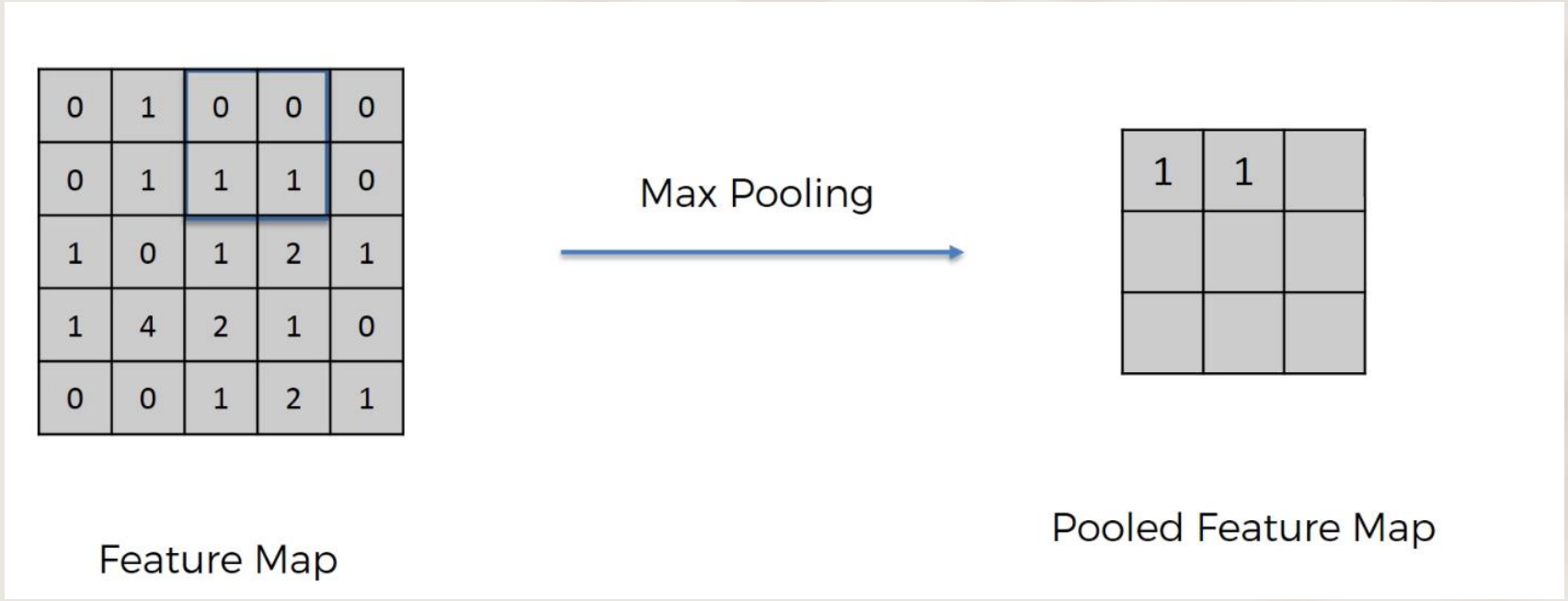
Sieci splotowe



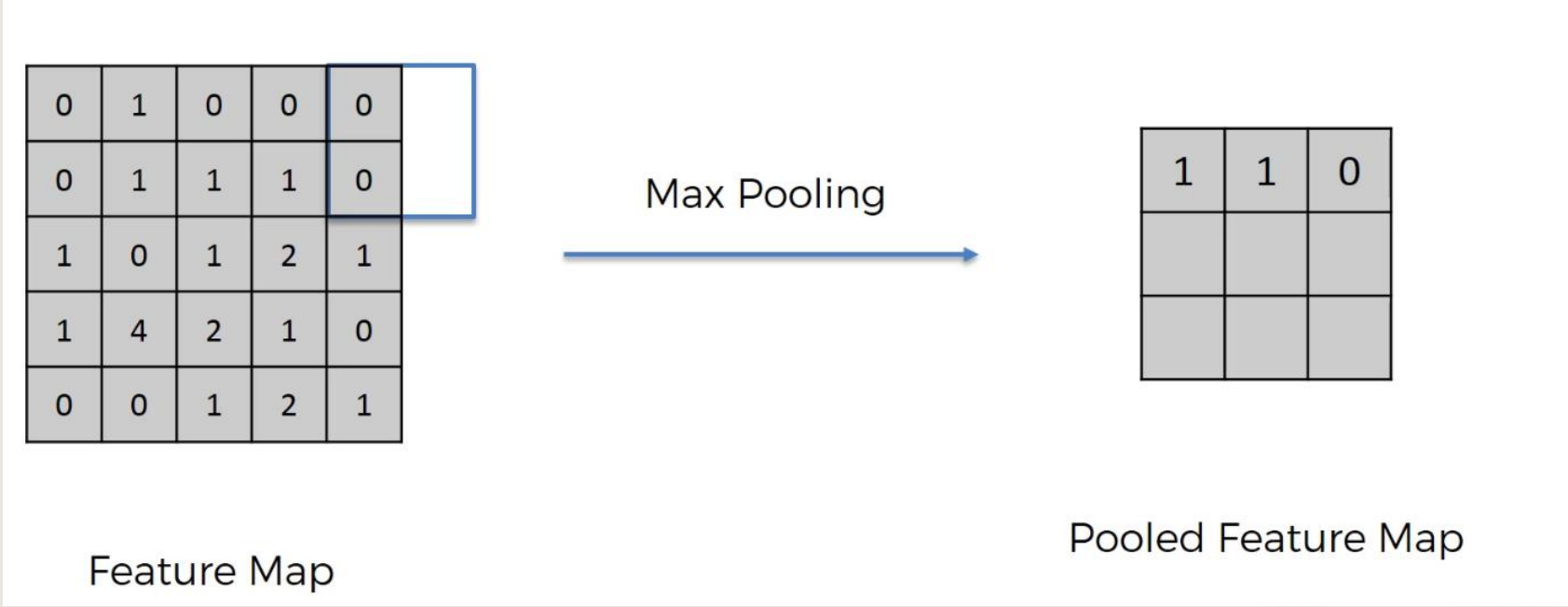
Sieci splotowe



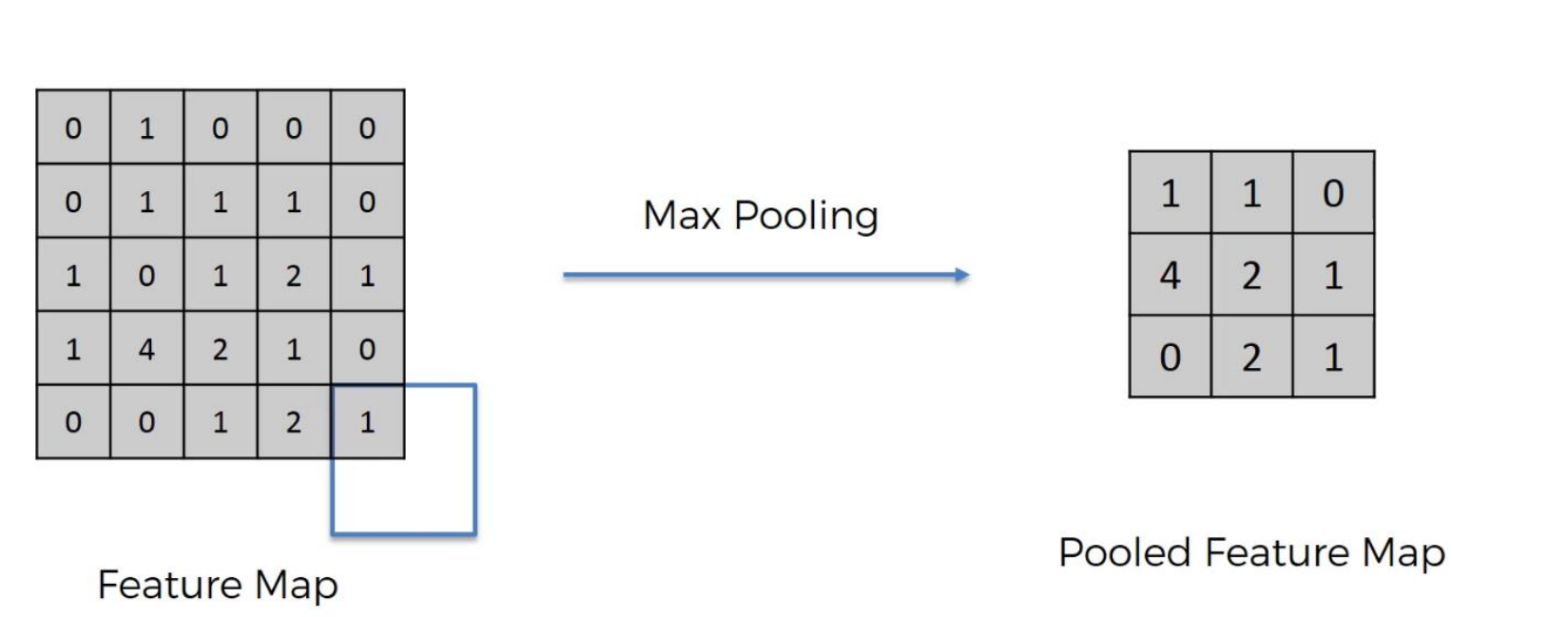
Sieci splotowe



Sieci splotowe



Sieci splotowe



Sieci splotowe

Additional Reading:

Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition

By Dominik Scherer et al. (2010)

Link:

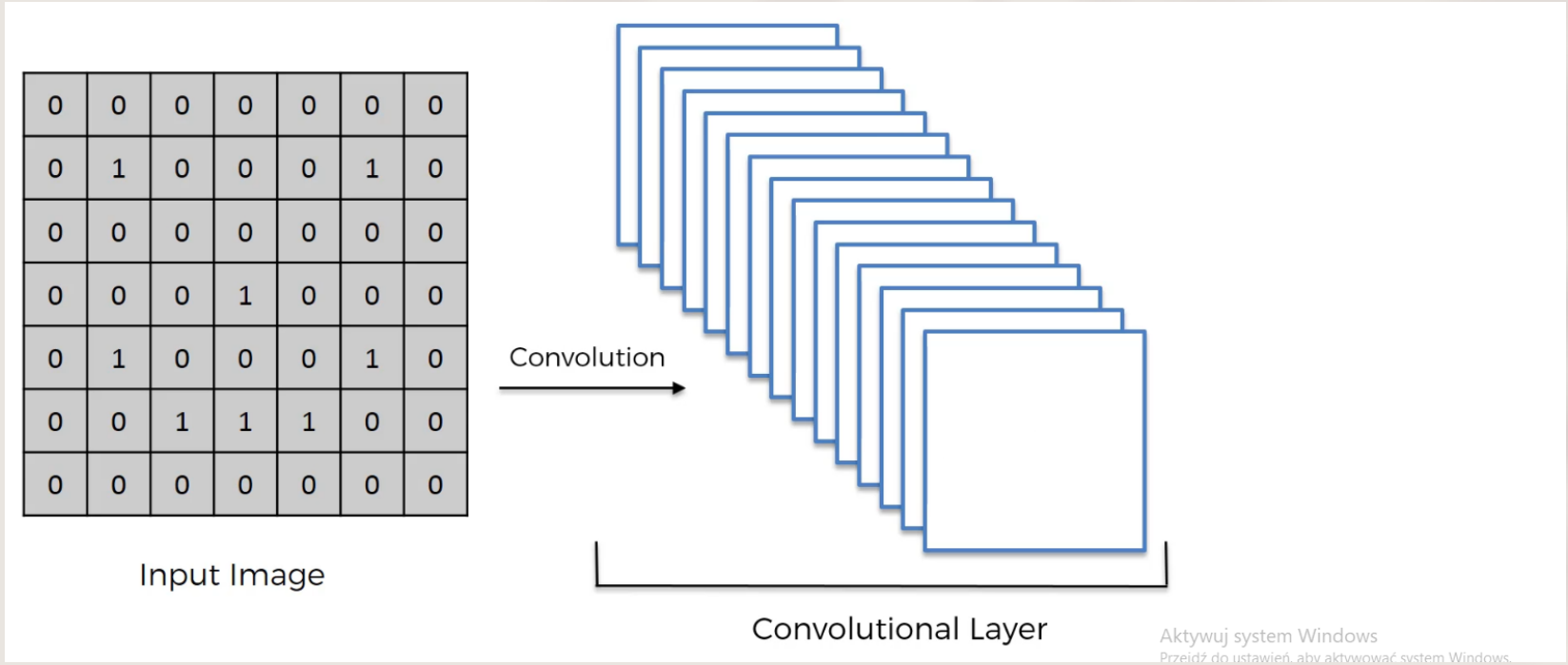
http://ais.uni-bonn.de/papers/icann2010_maxpool.pdf

Sieci splotowe

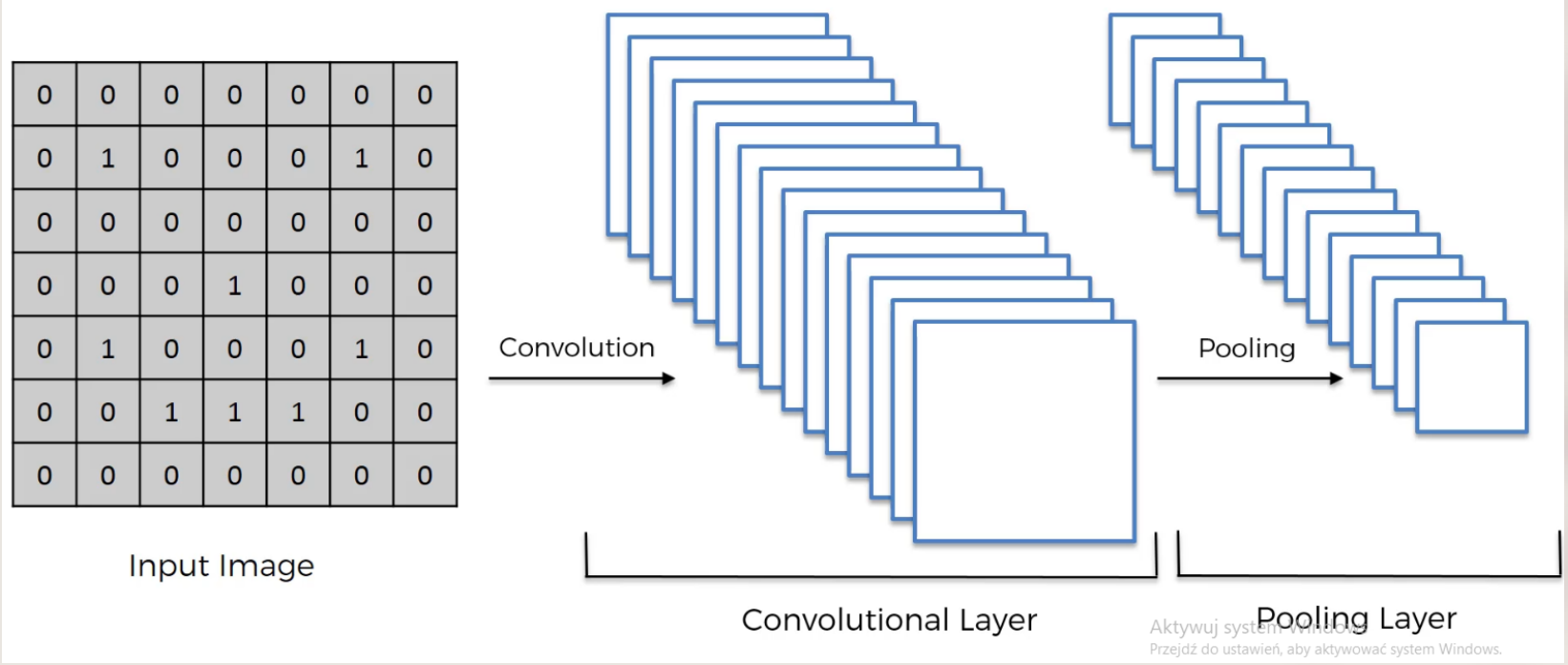
0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Input Image

Sieci splotowe



Sieci splotowe

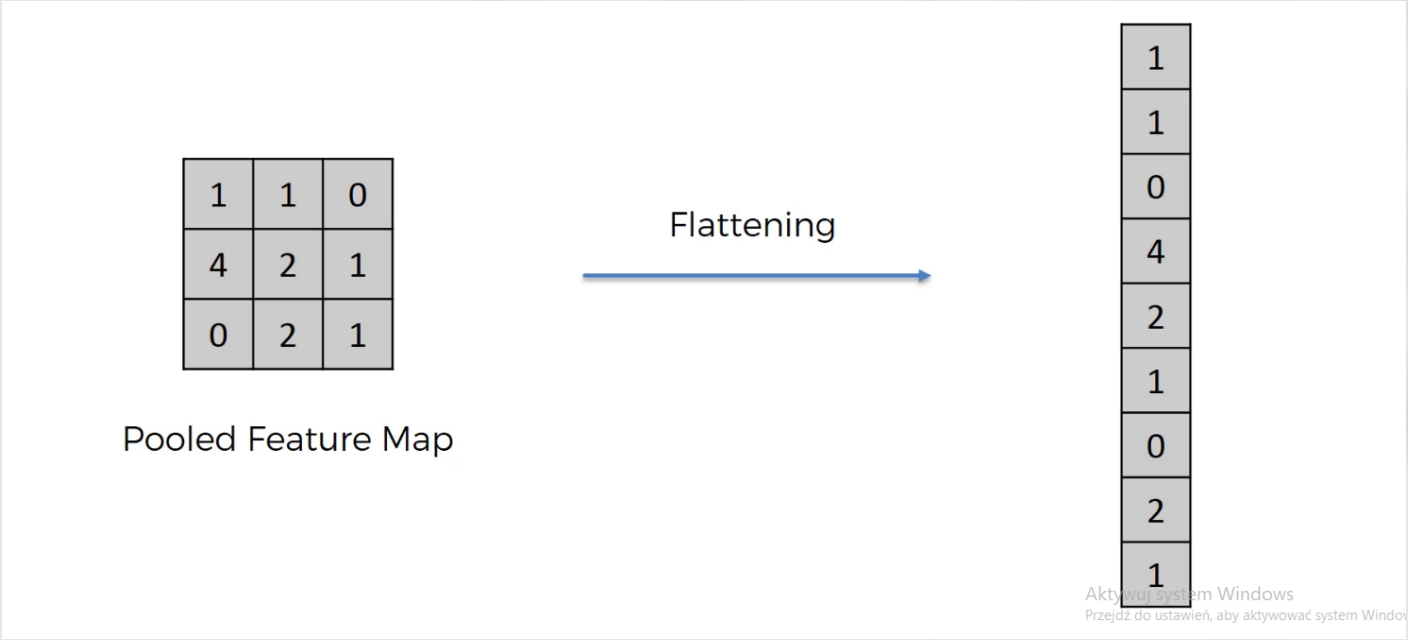


Sieci splotowe

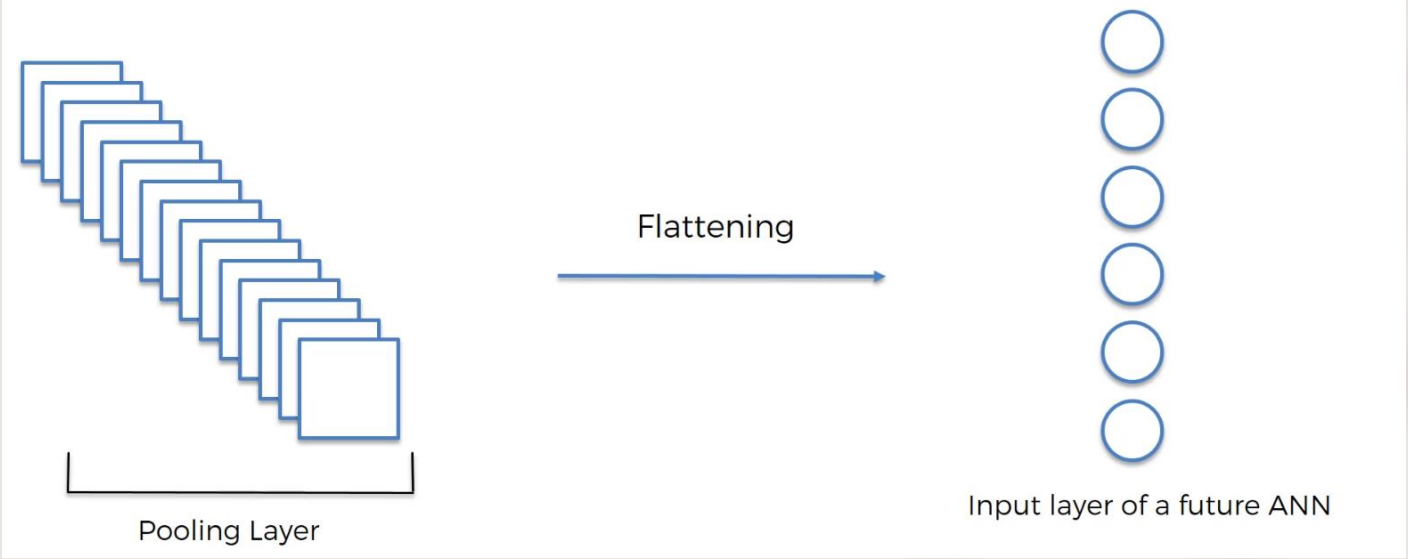
1	1	0
4	2	1
0	2	1

Pooled Feature Map

Sieci splotowe

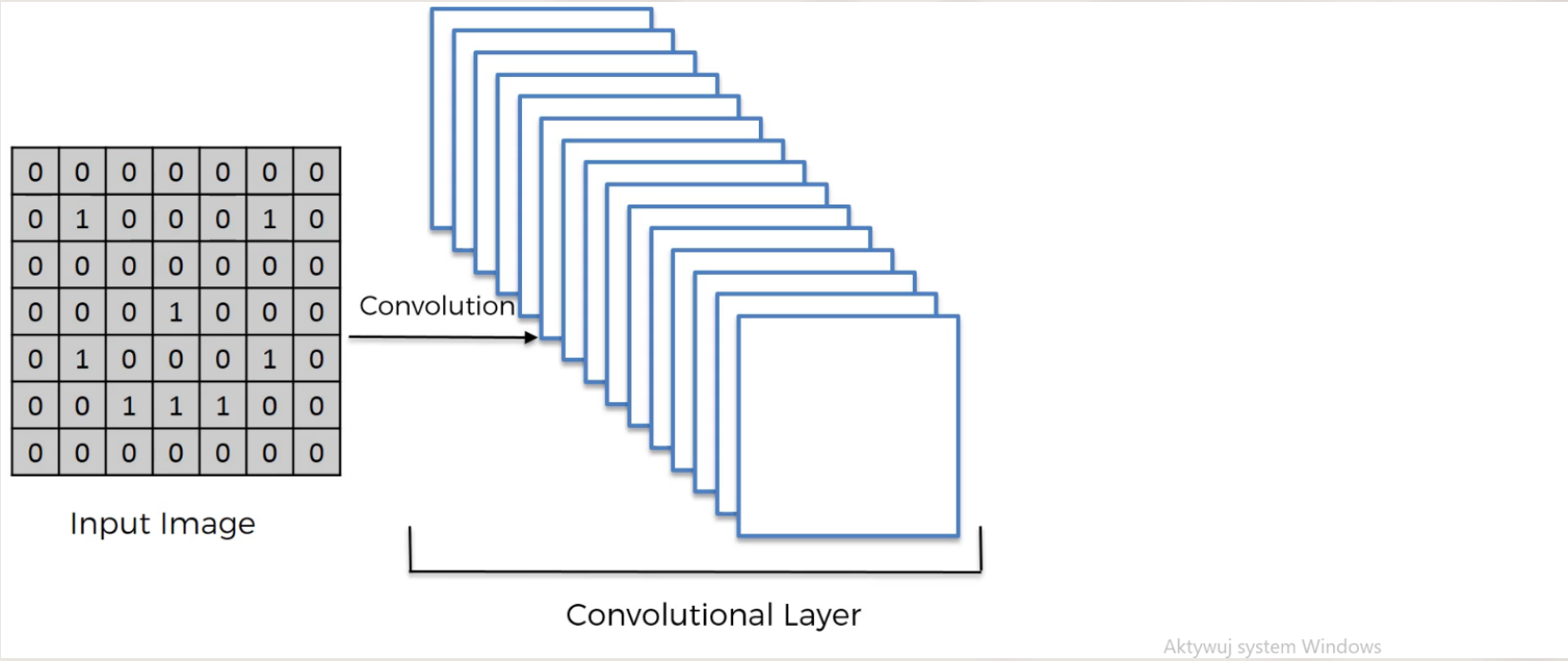


Sieci splotowe

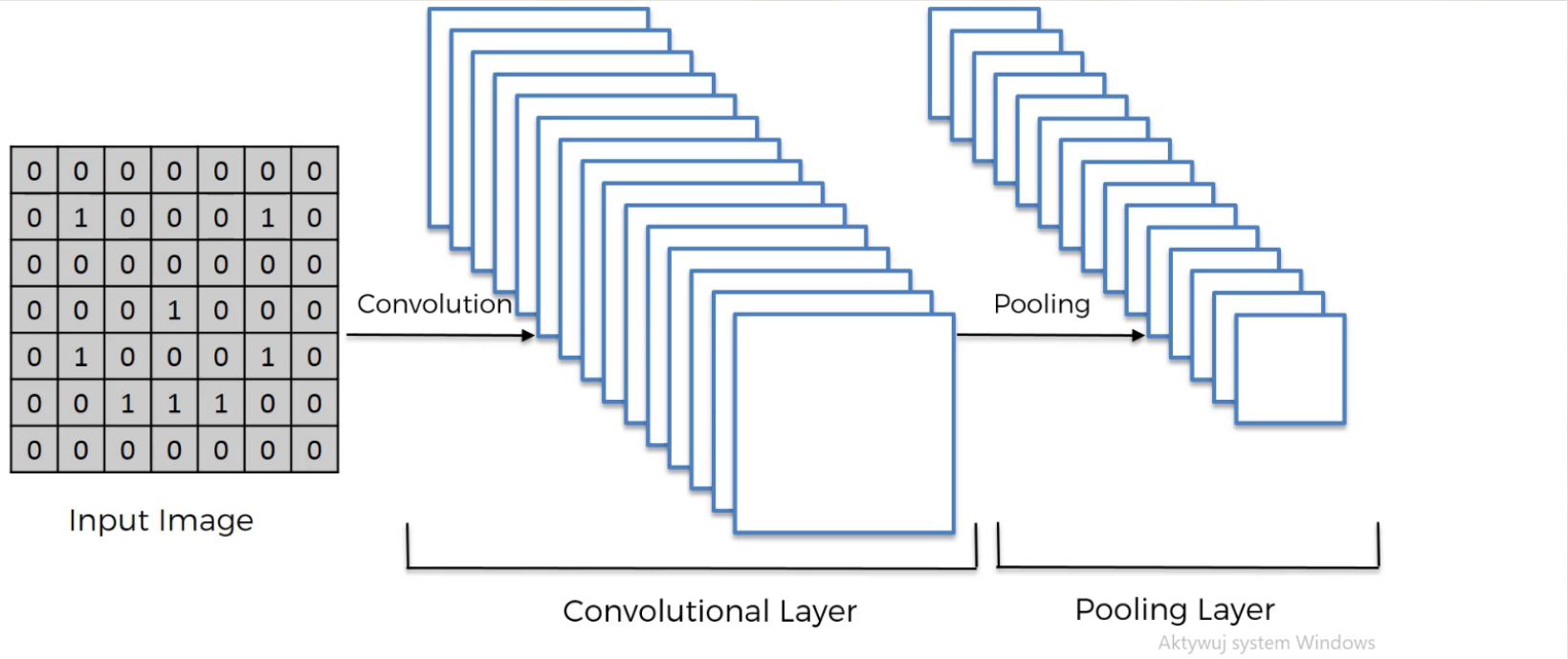


Sieci splotowe

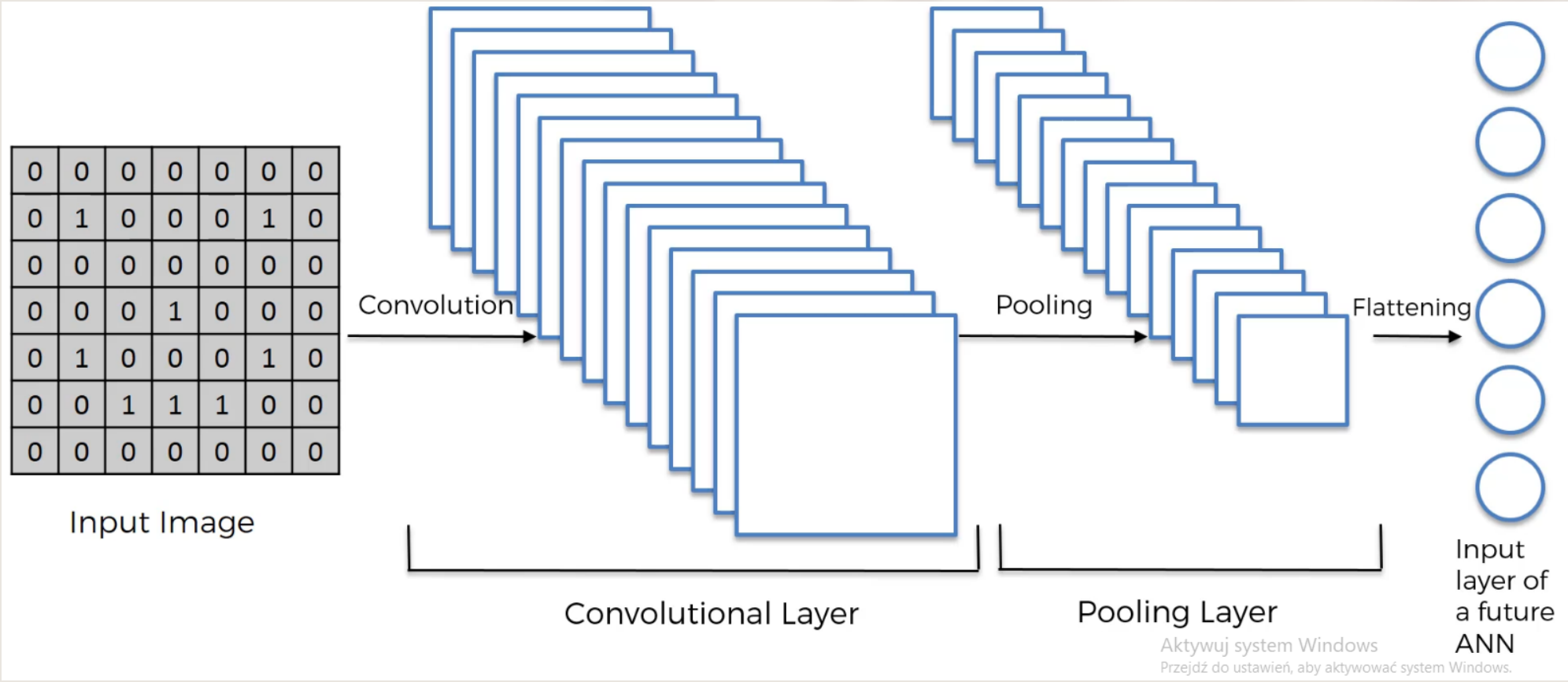
Sieci splotowe



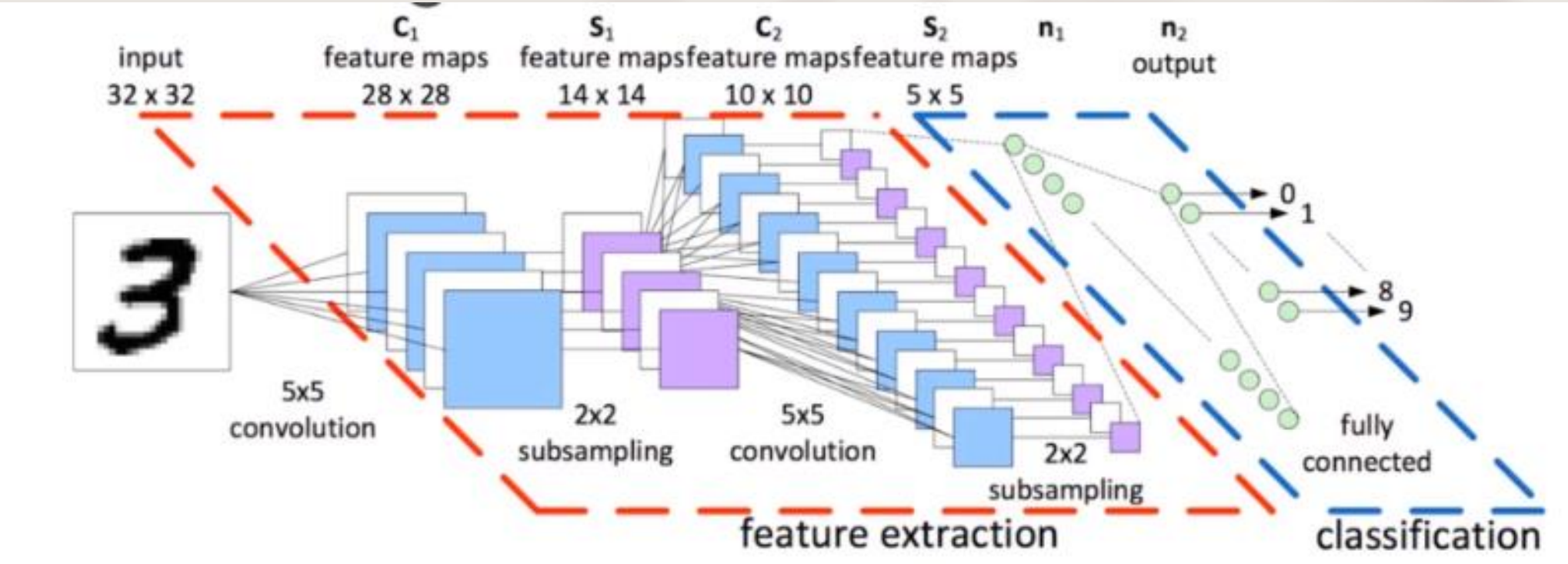
Sieci splotowe



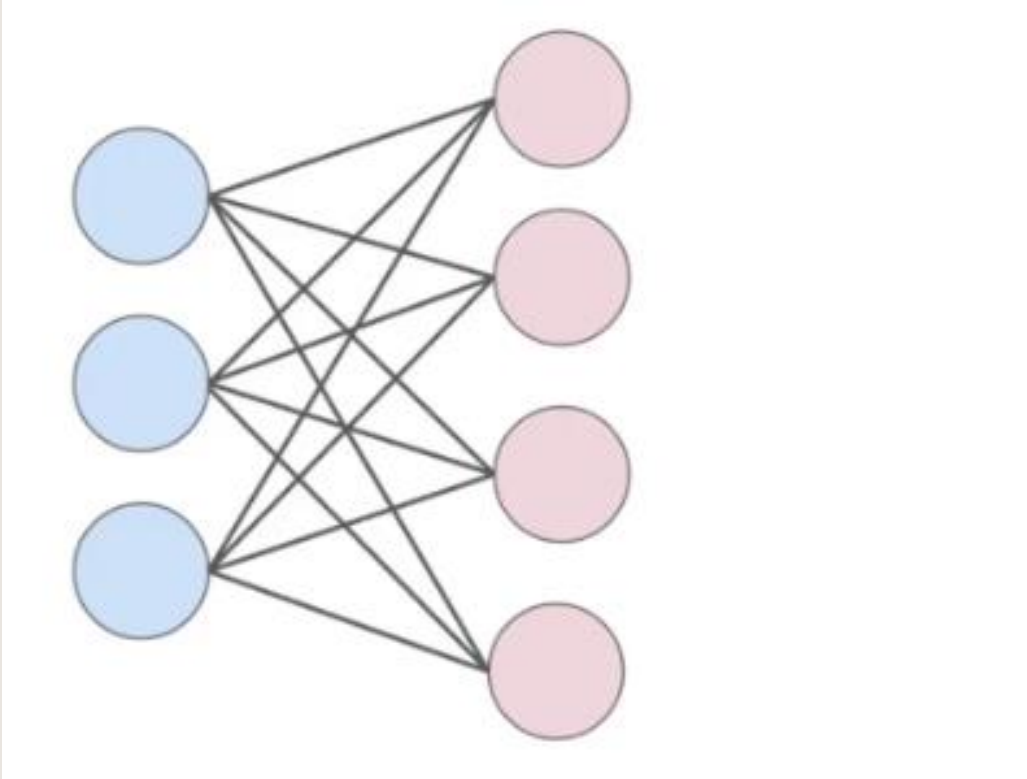
Sieci splotowe



Sieci splotowe

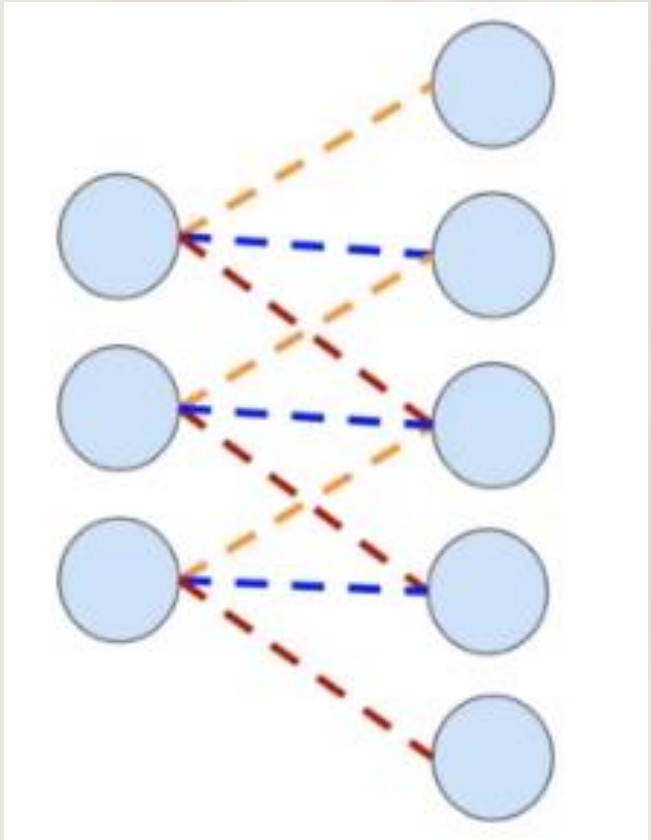


Sieci splotowe



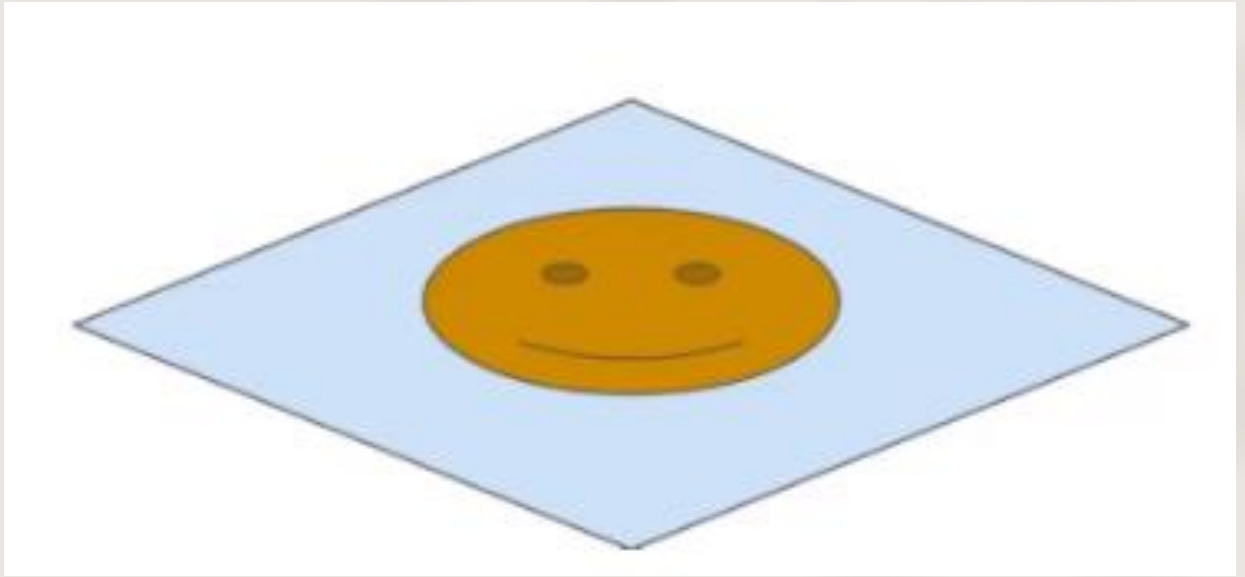
Warstwa gęsta

Sieci splotowe



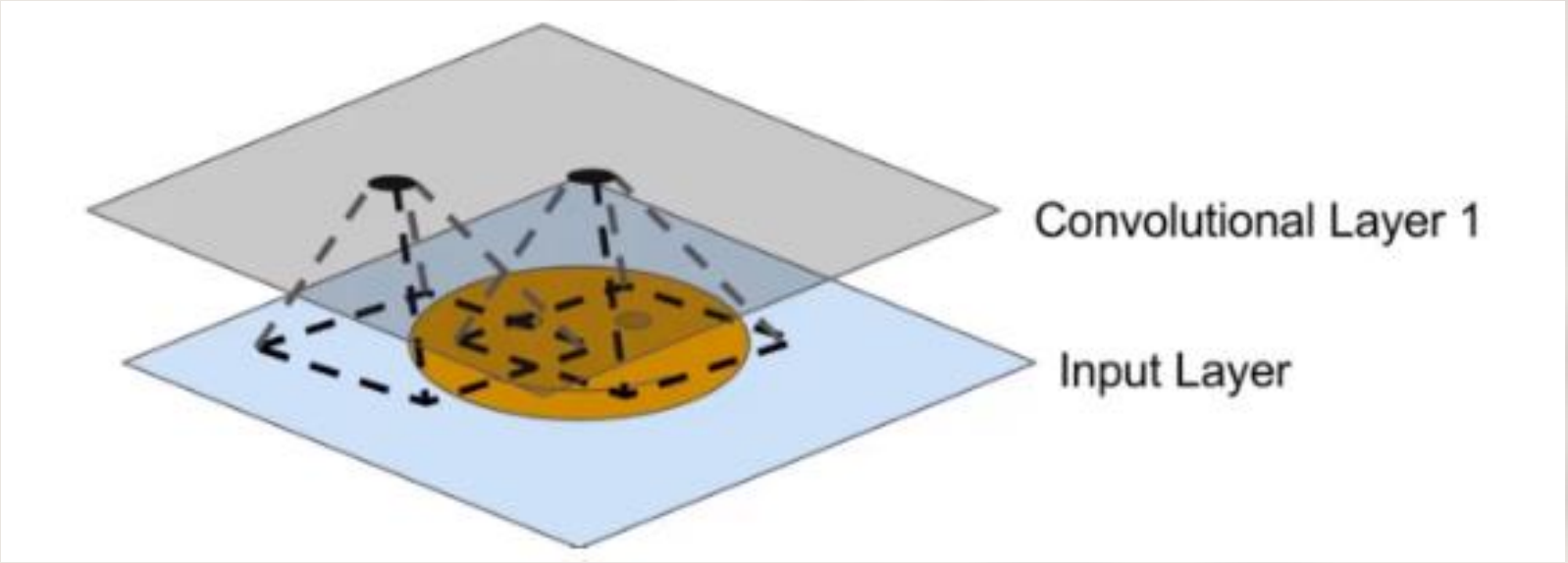
Warstwa splotowa

Sieci splotowe



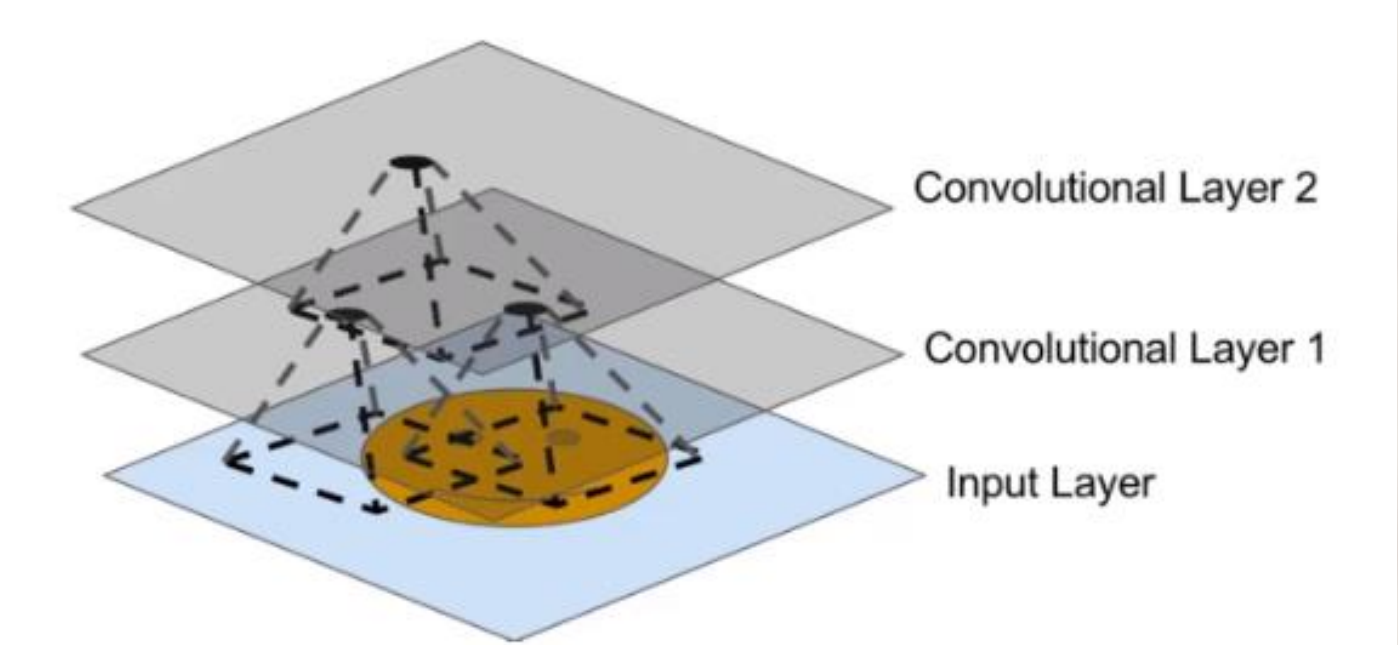
Warstwa wejściowa

Sieci splotowe



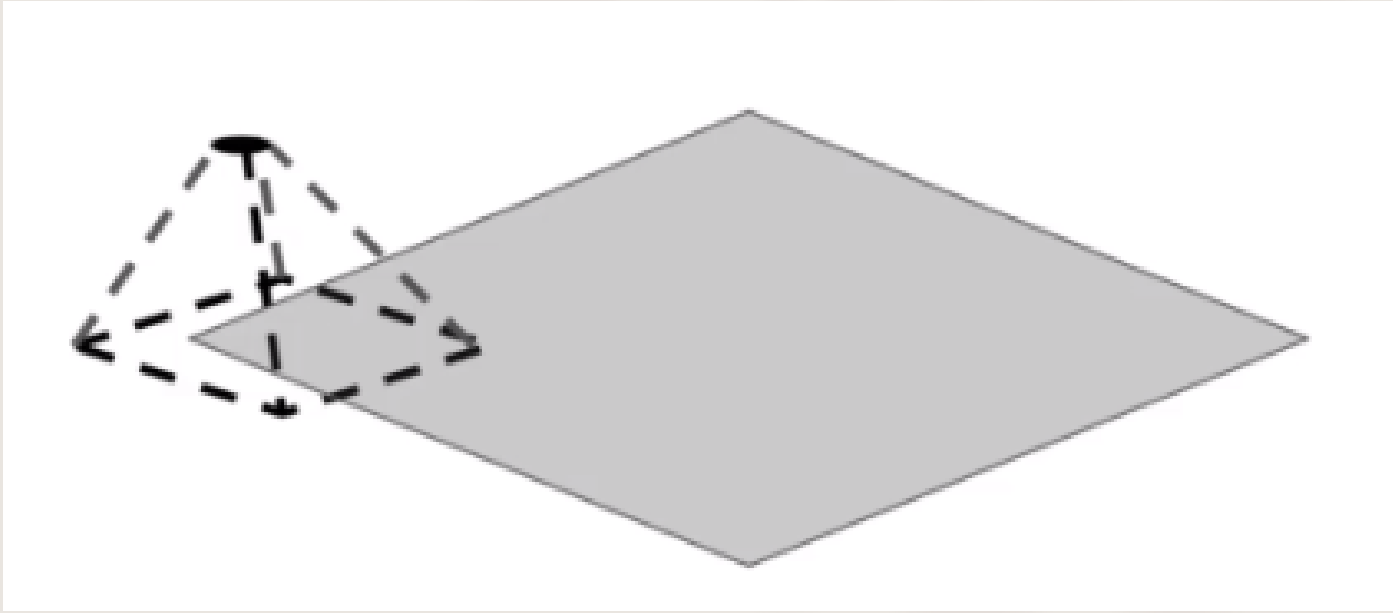
I warstwa splotowa

Sieci splotowe



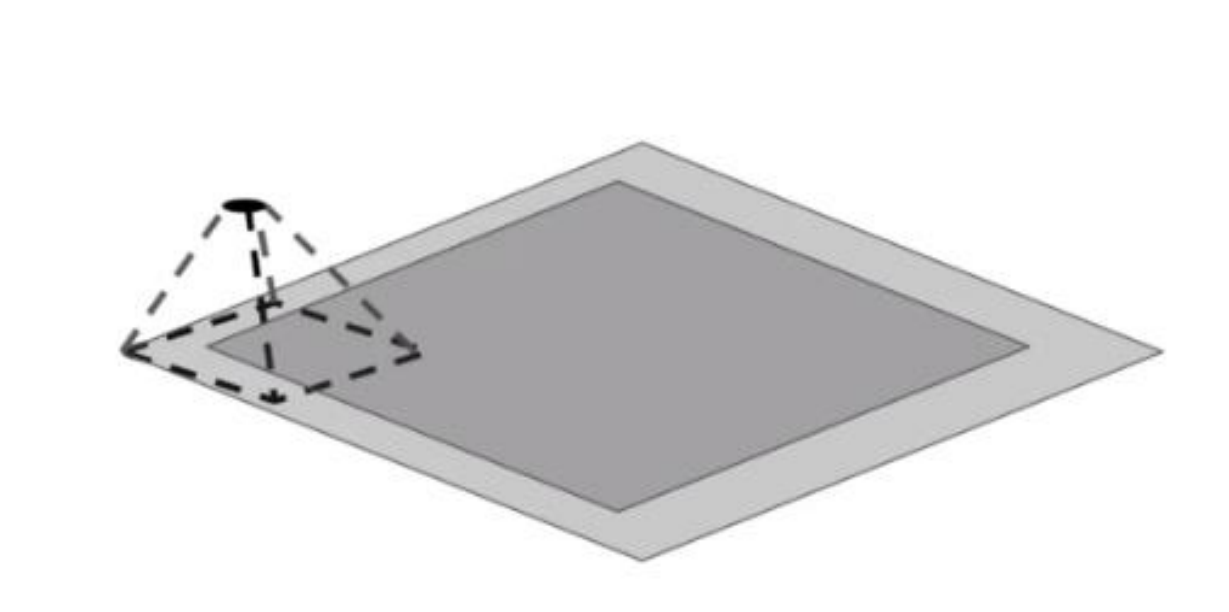
II warstwa splotowa

Sieci splotowe



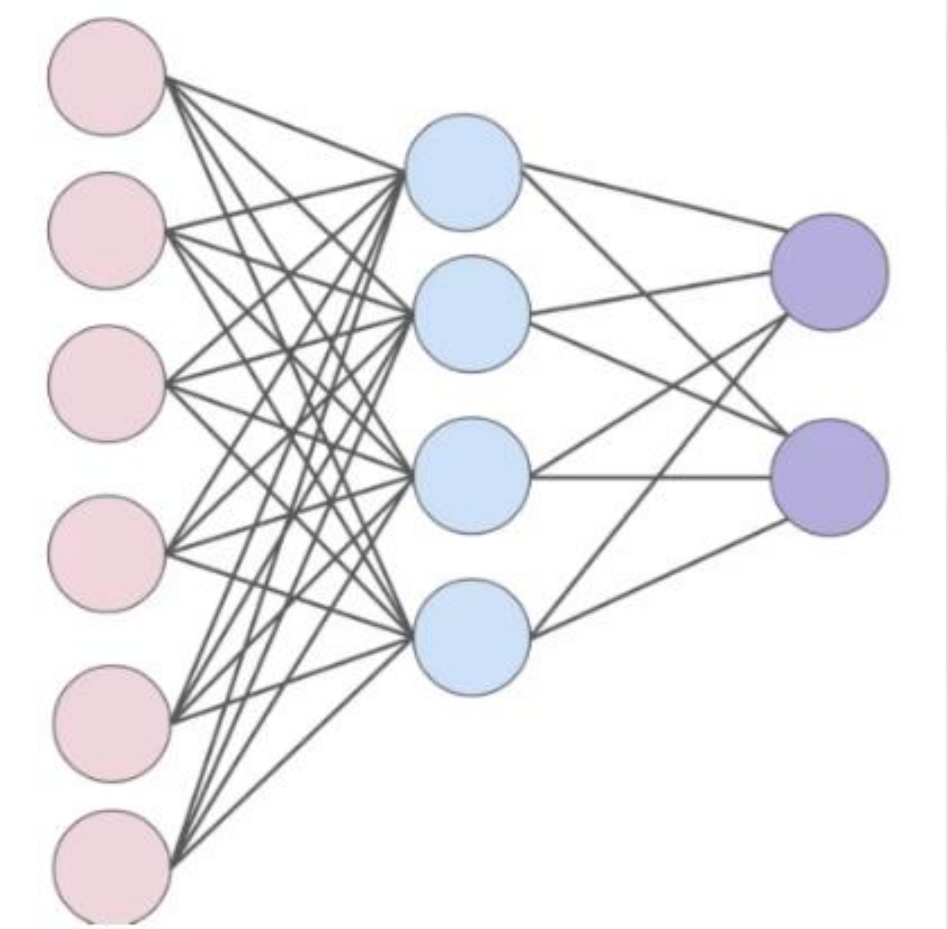
Padding

Sieci splotowe

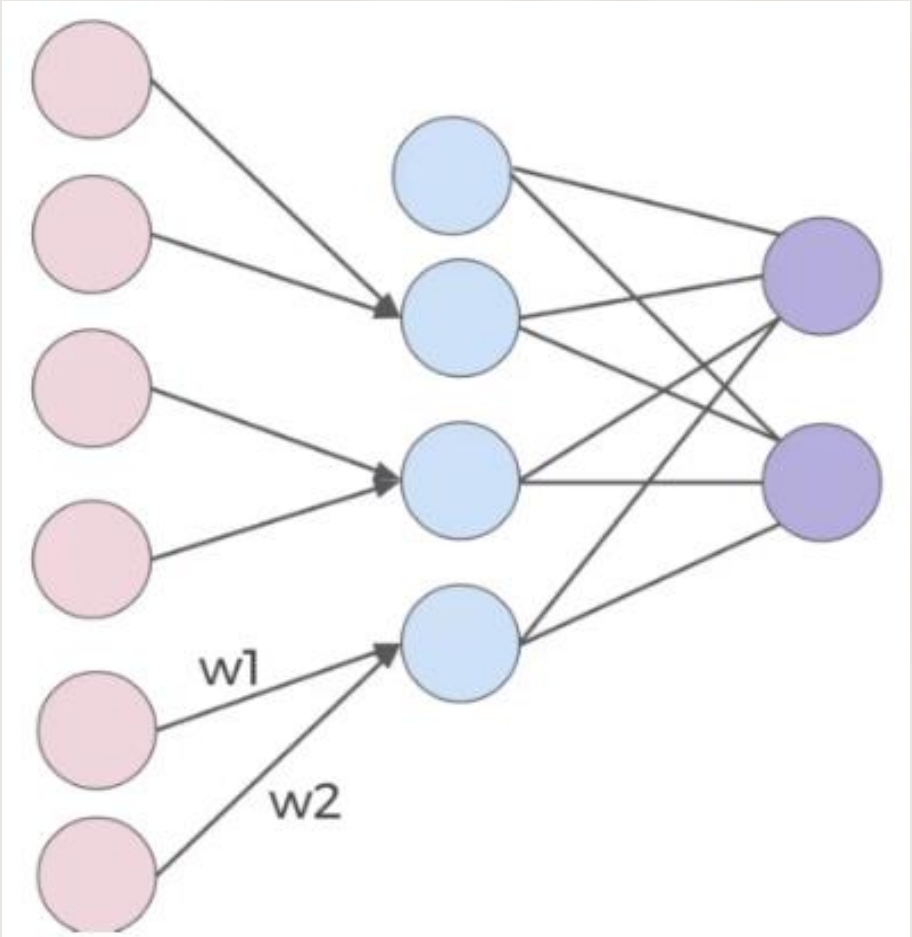


Padding

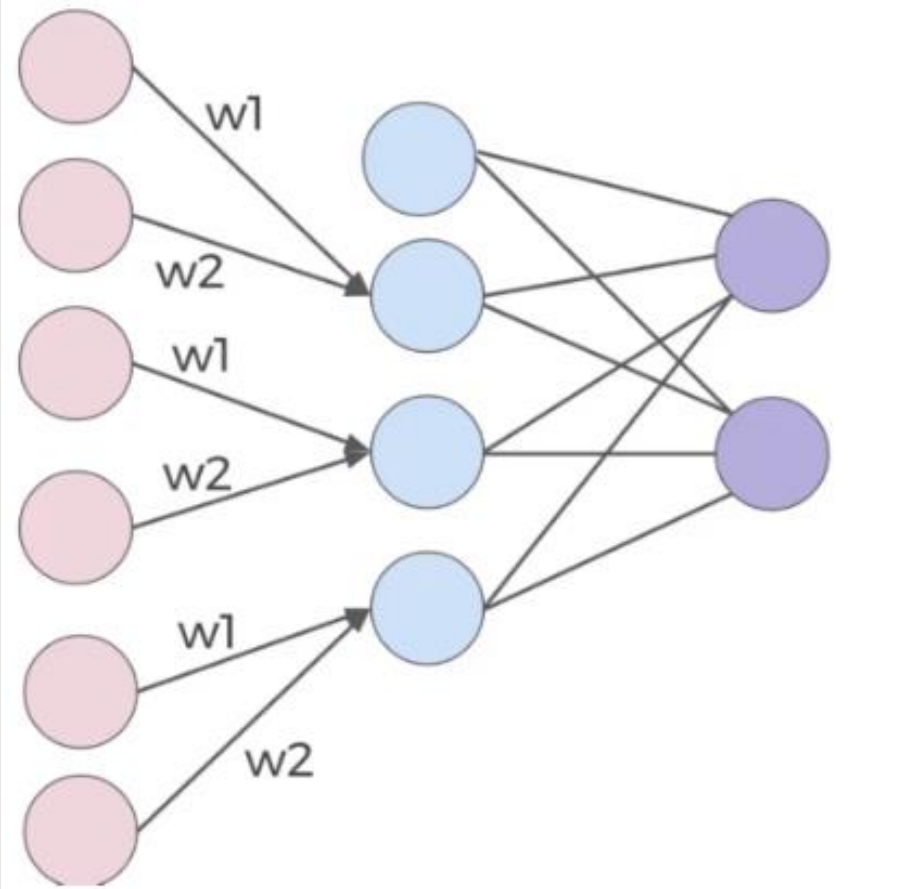
Warstwa splotowa



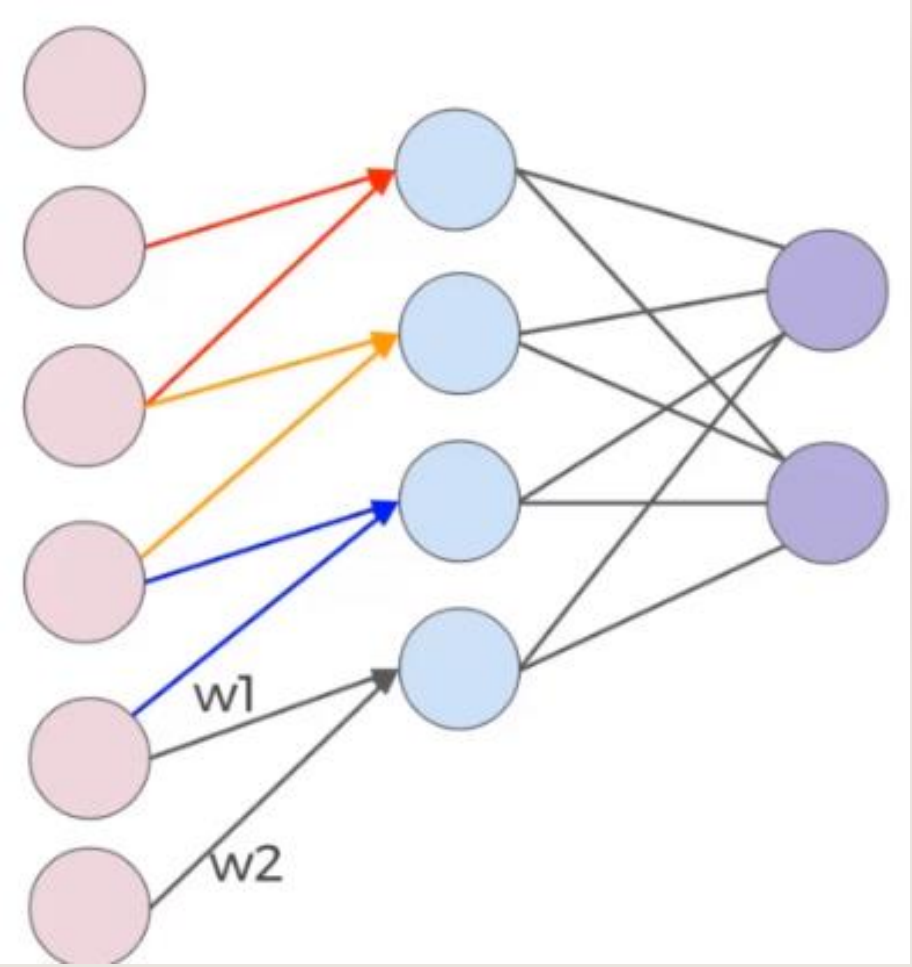
Warstwa splotowa



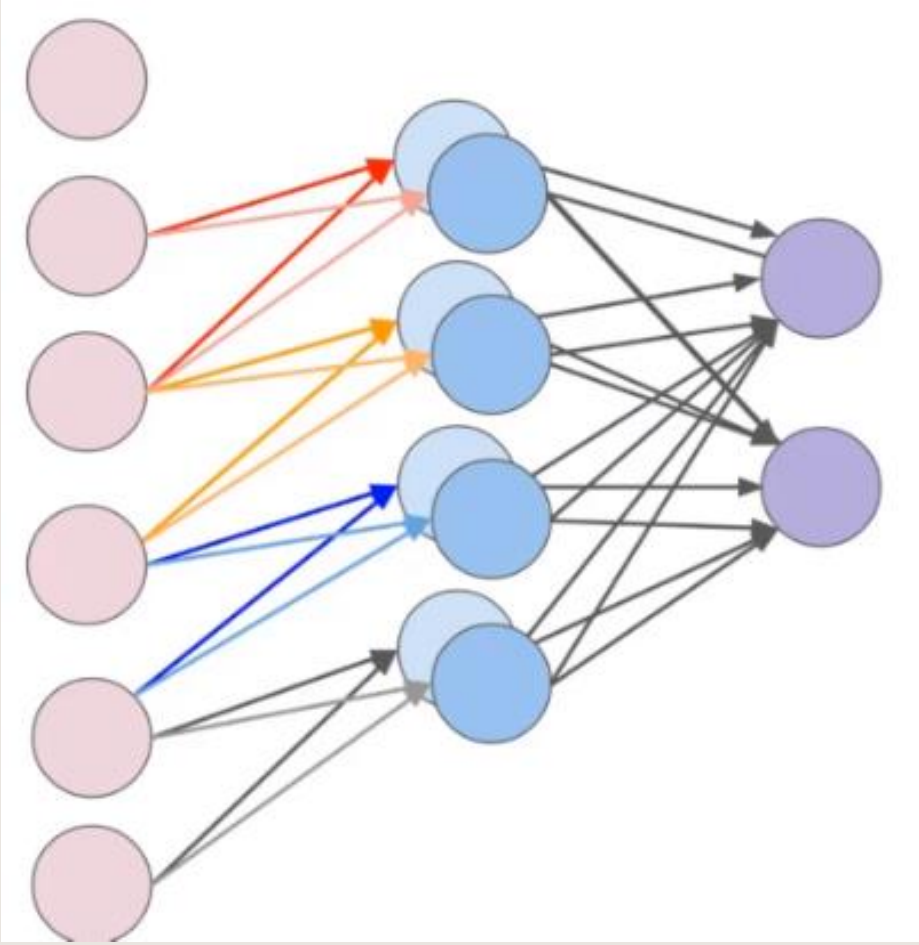
Warstwa splotowa



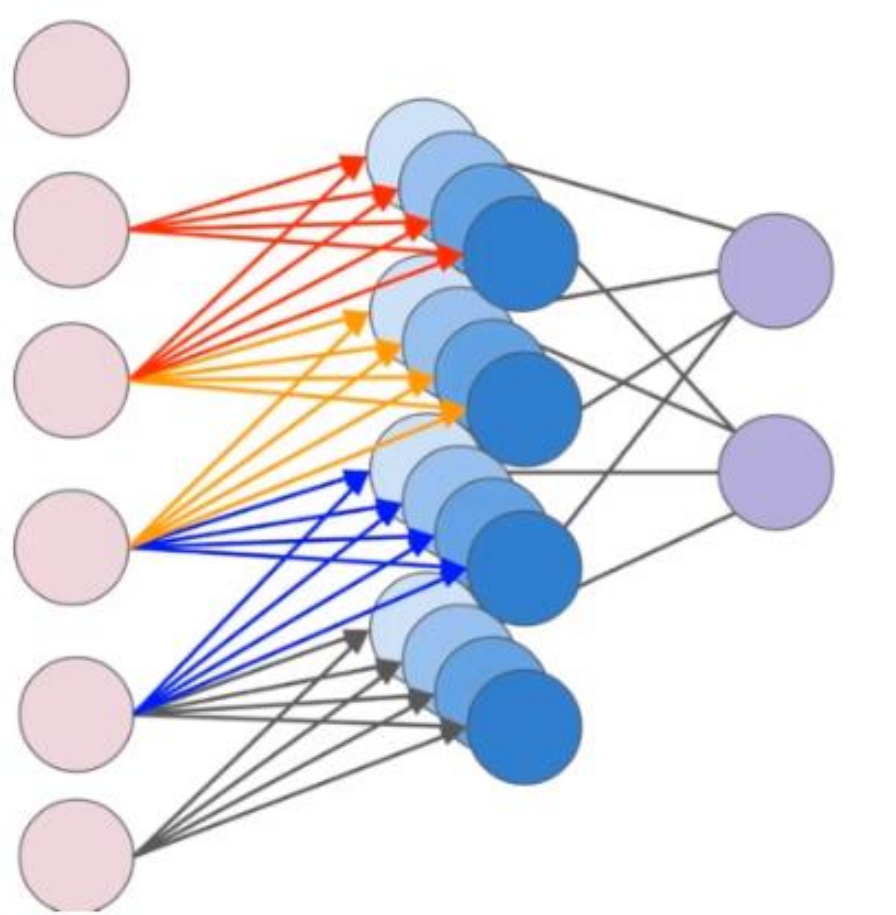
Warstwa splotowa



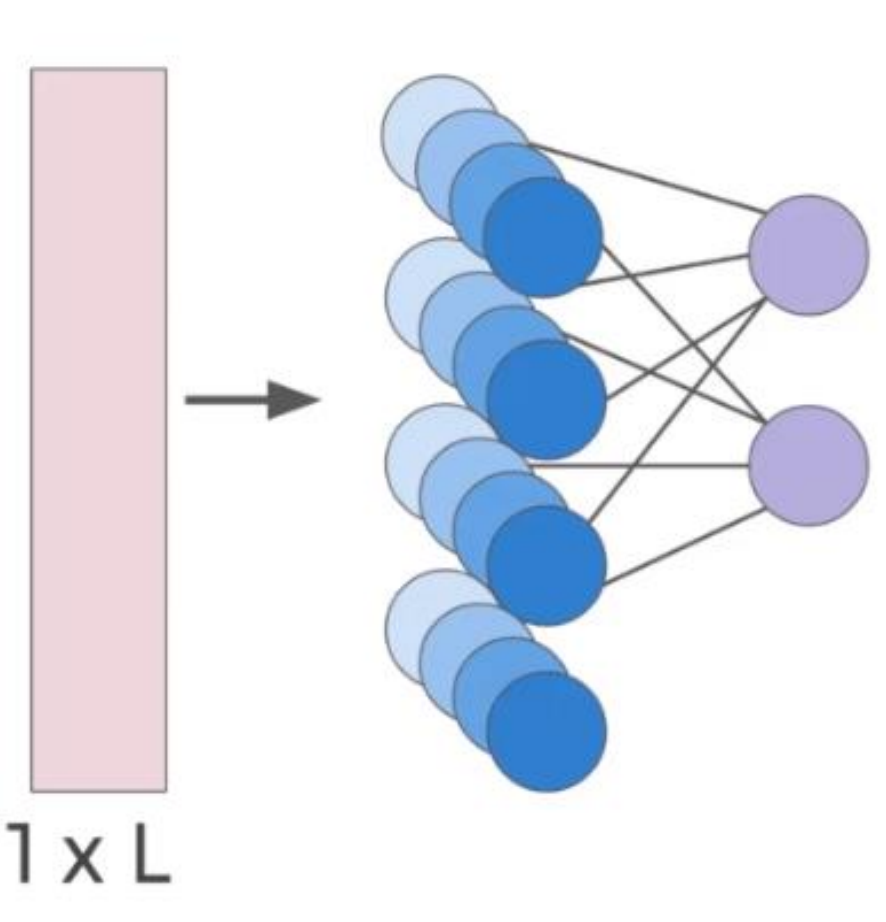
Warstwa splotowa



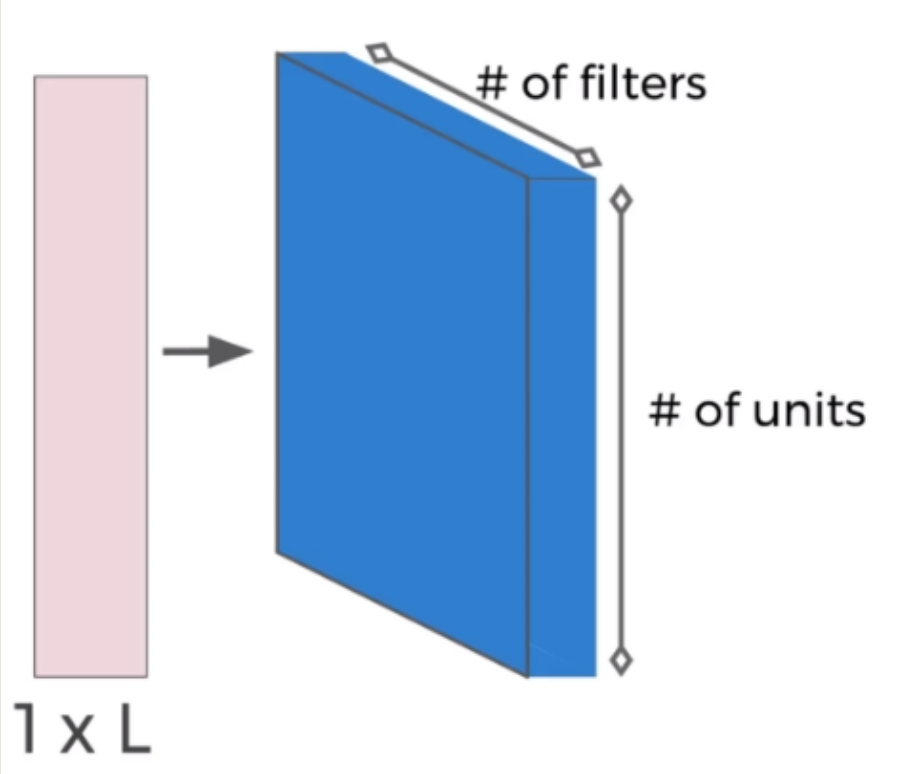
Warstwa splotowa



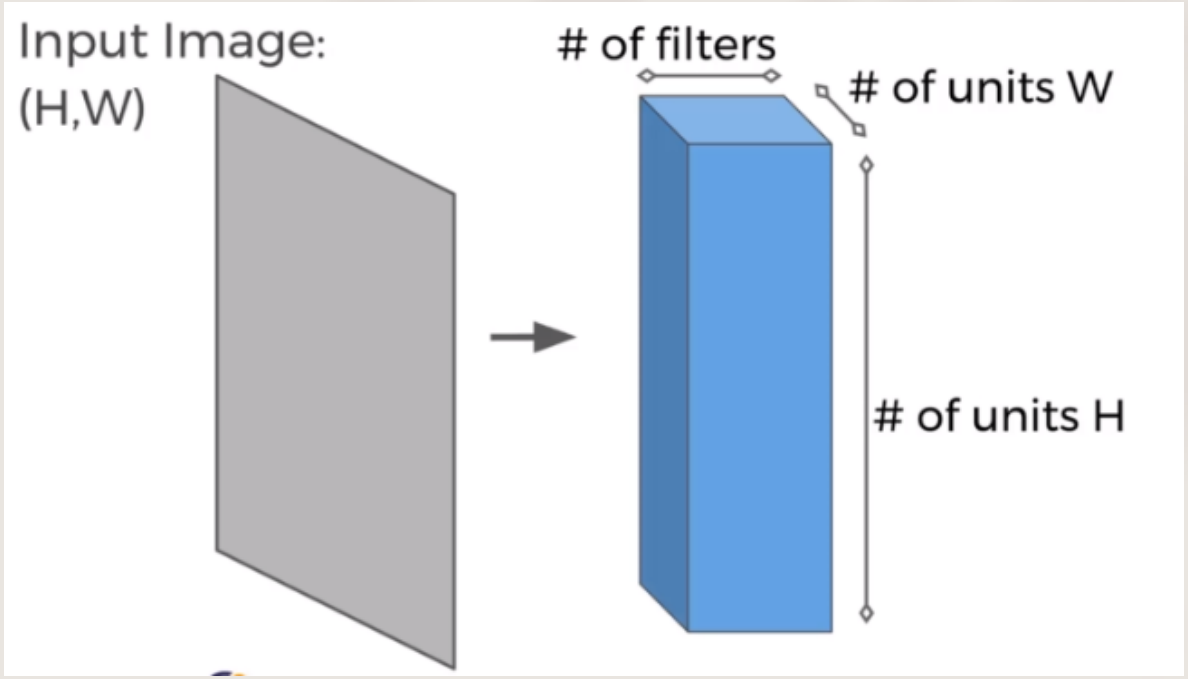
Warstwa splotowa



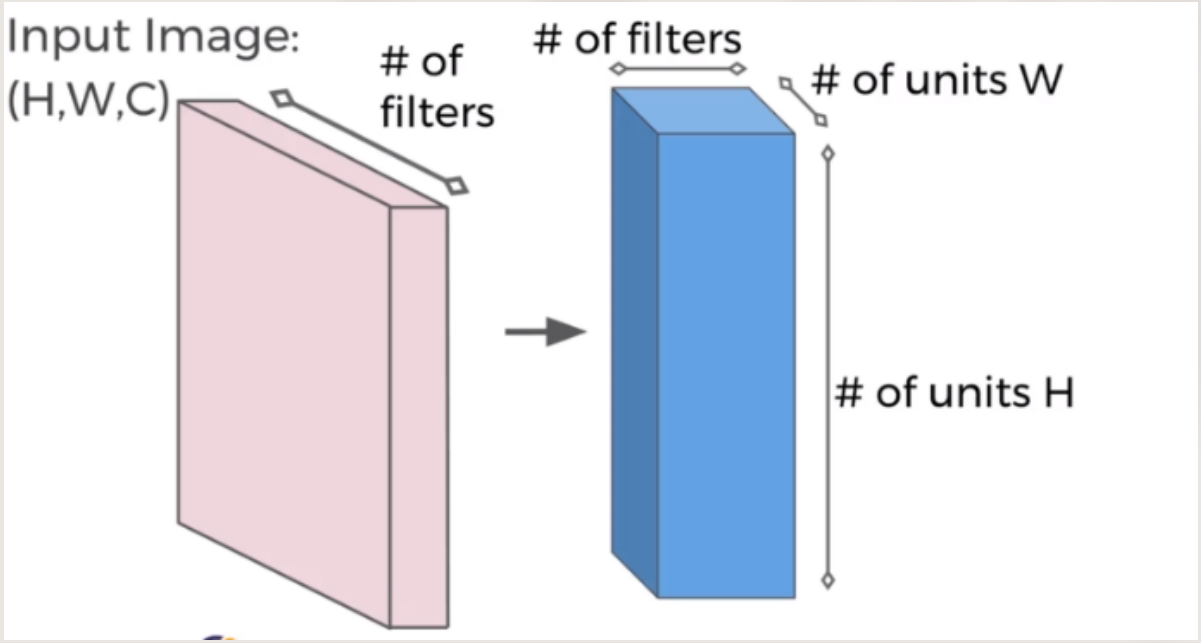
Warstwa splotowa



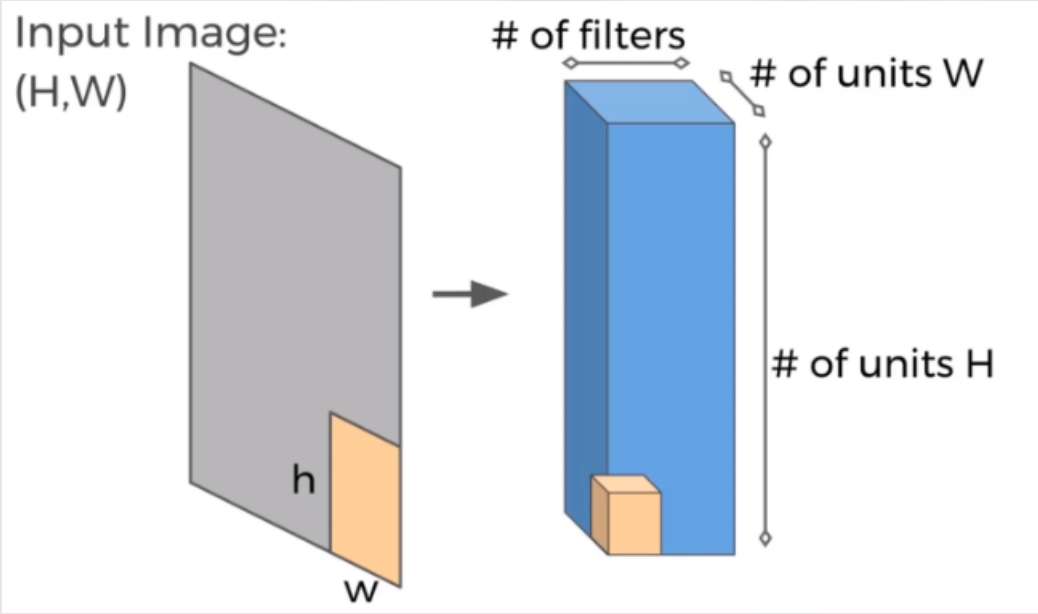
Warstwa splotowa



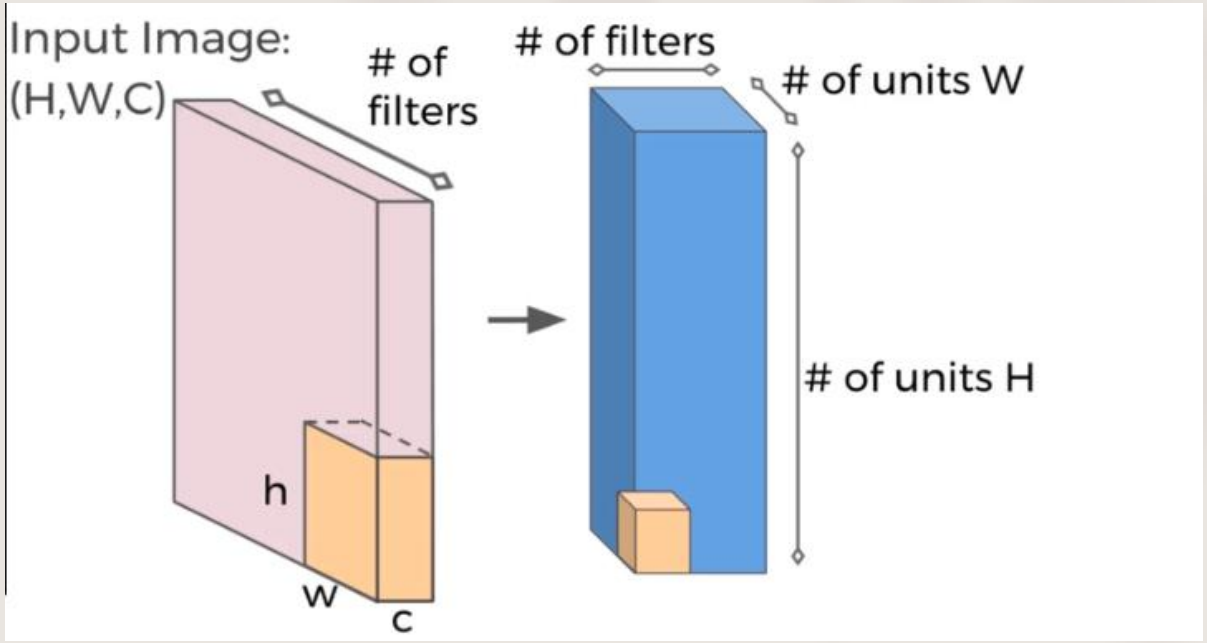
Warstwa splotowa



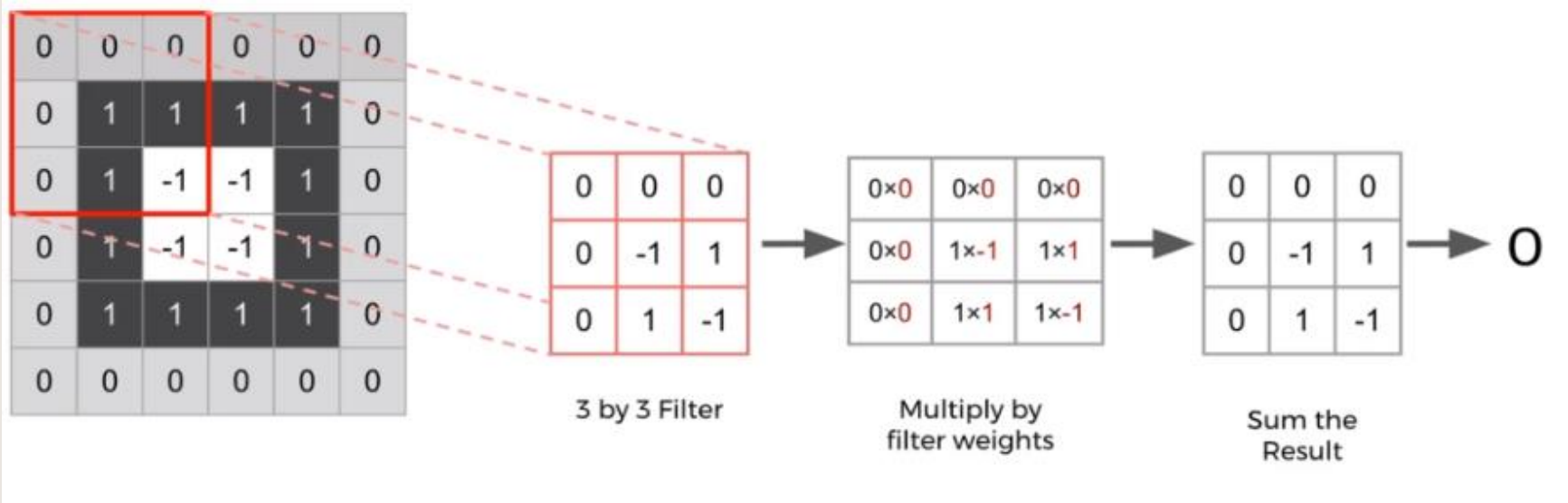
Warstwa splotowa



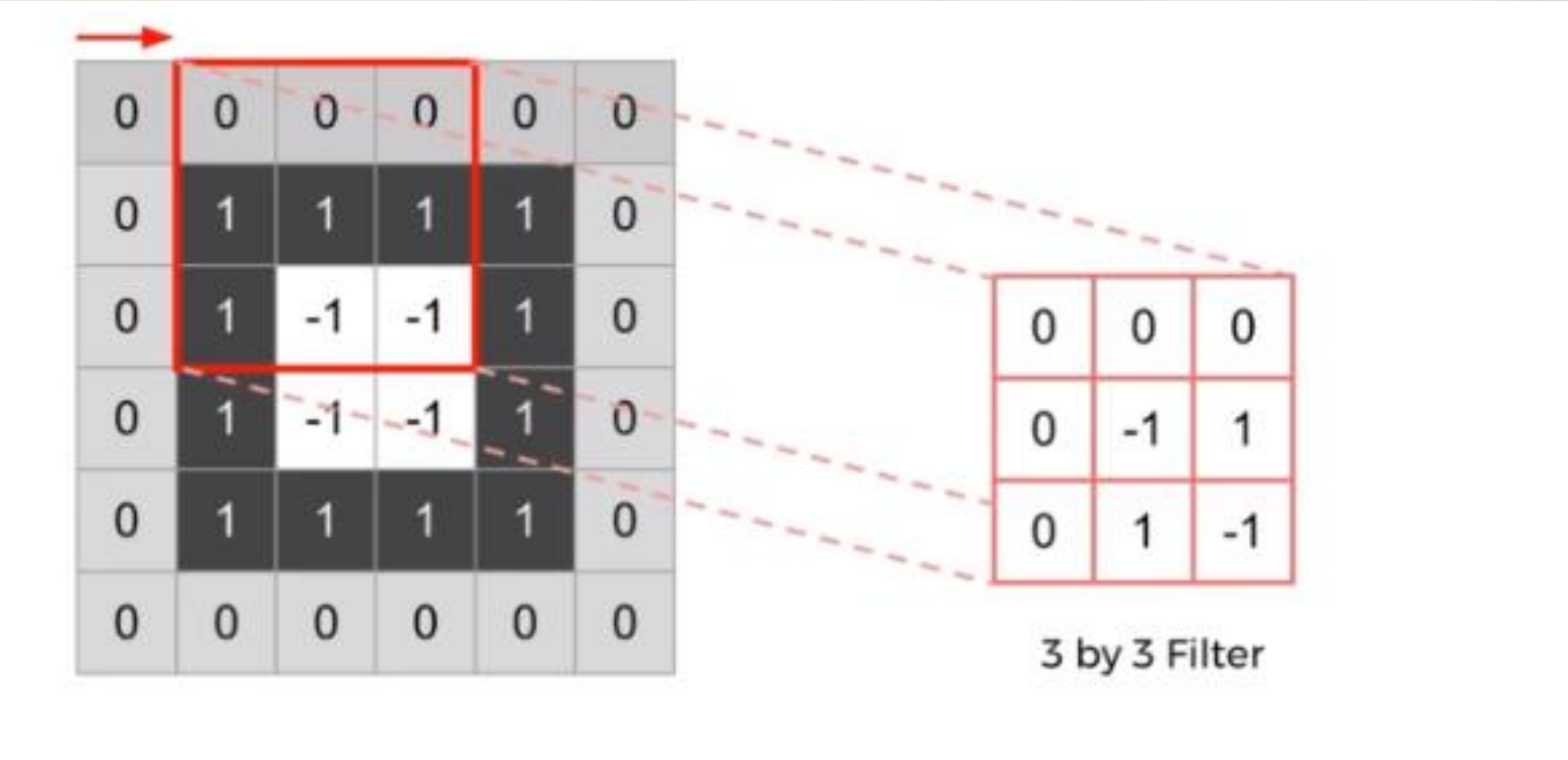
Warstwa splotowa



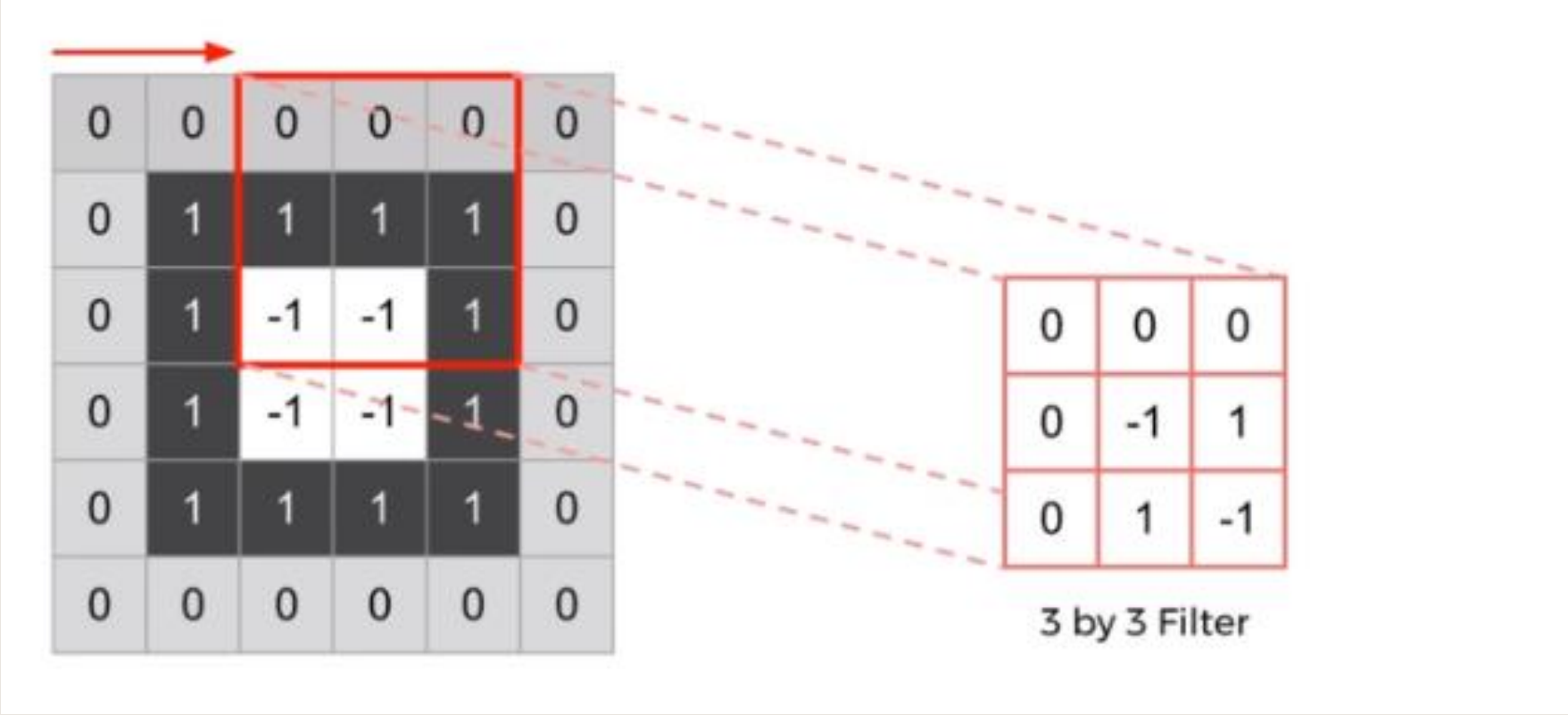
Warstwa splotowa



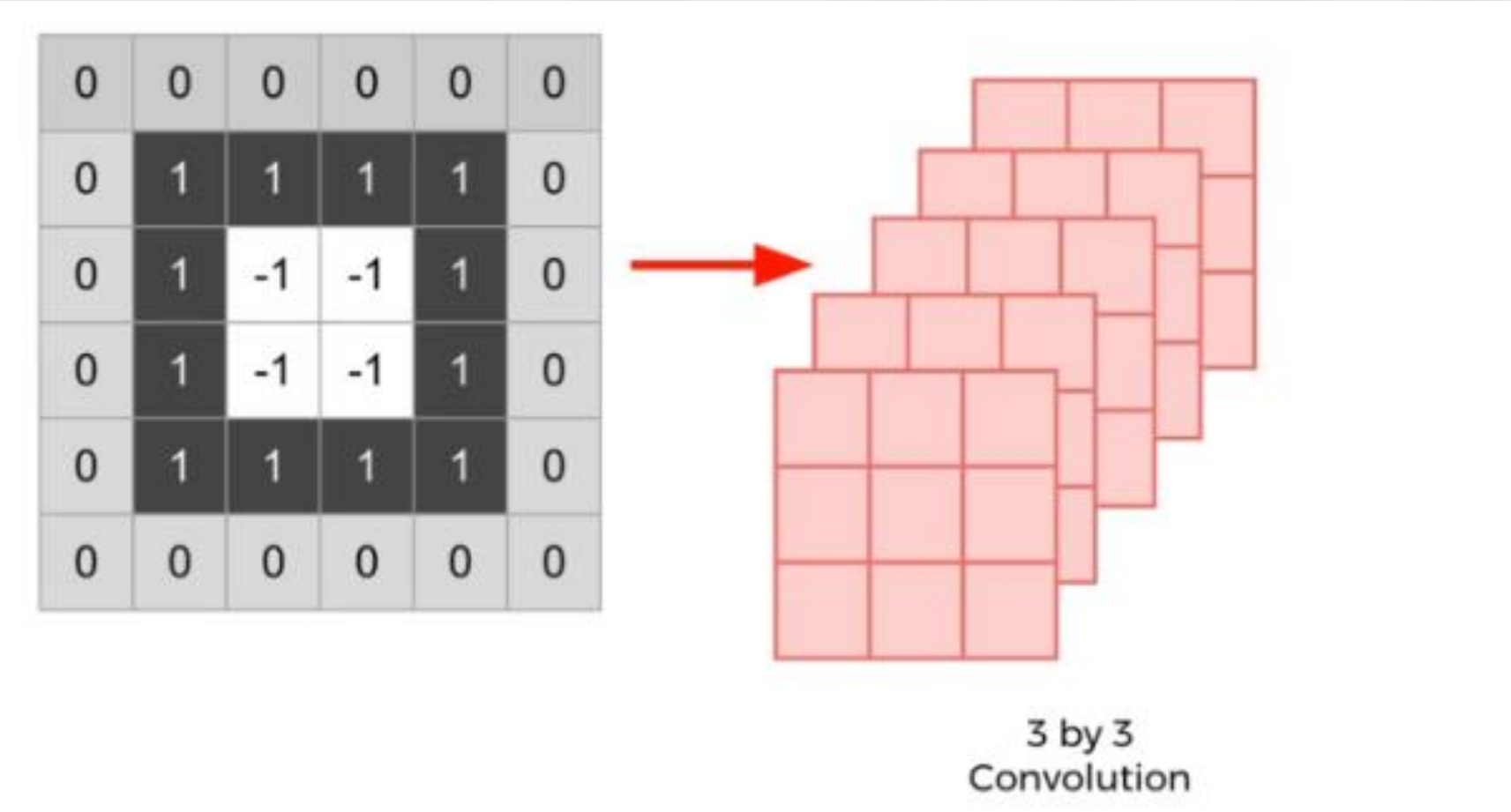
Warstwa splotowa



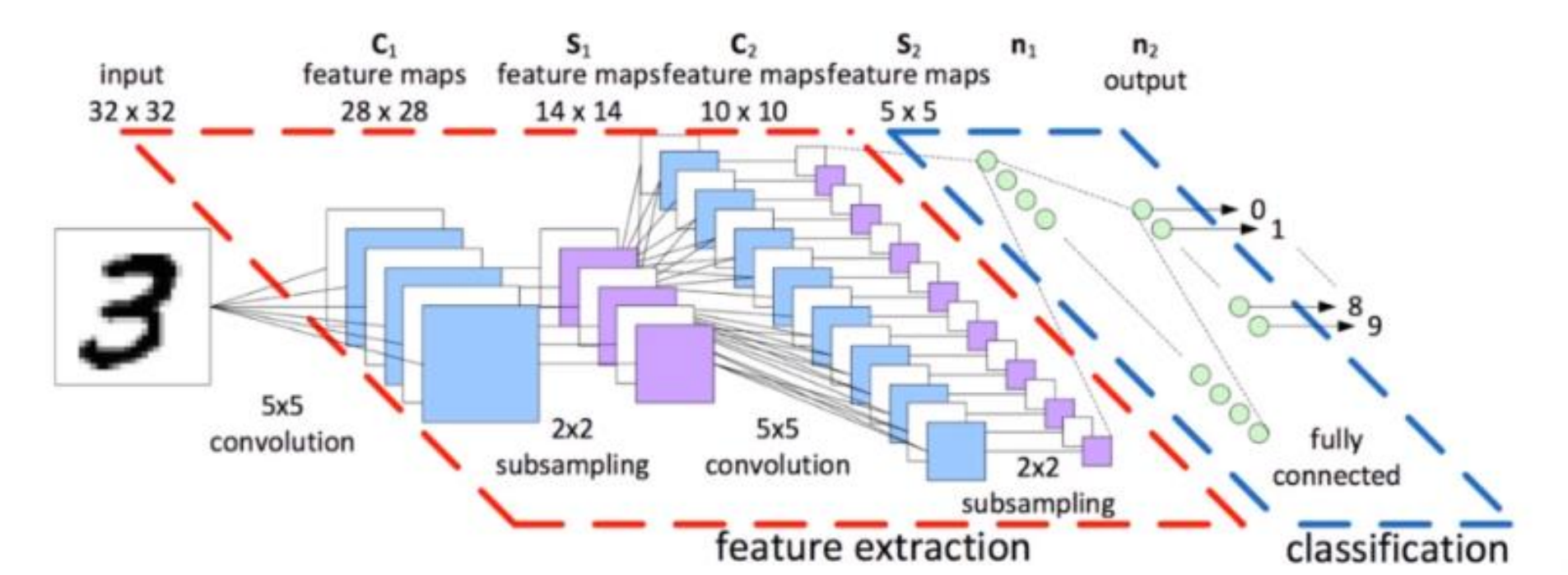
Warstwa splotowa



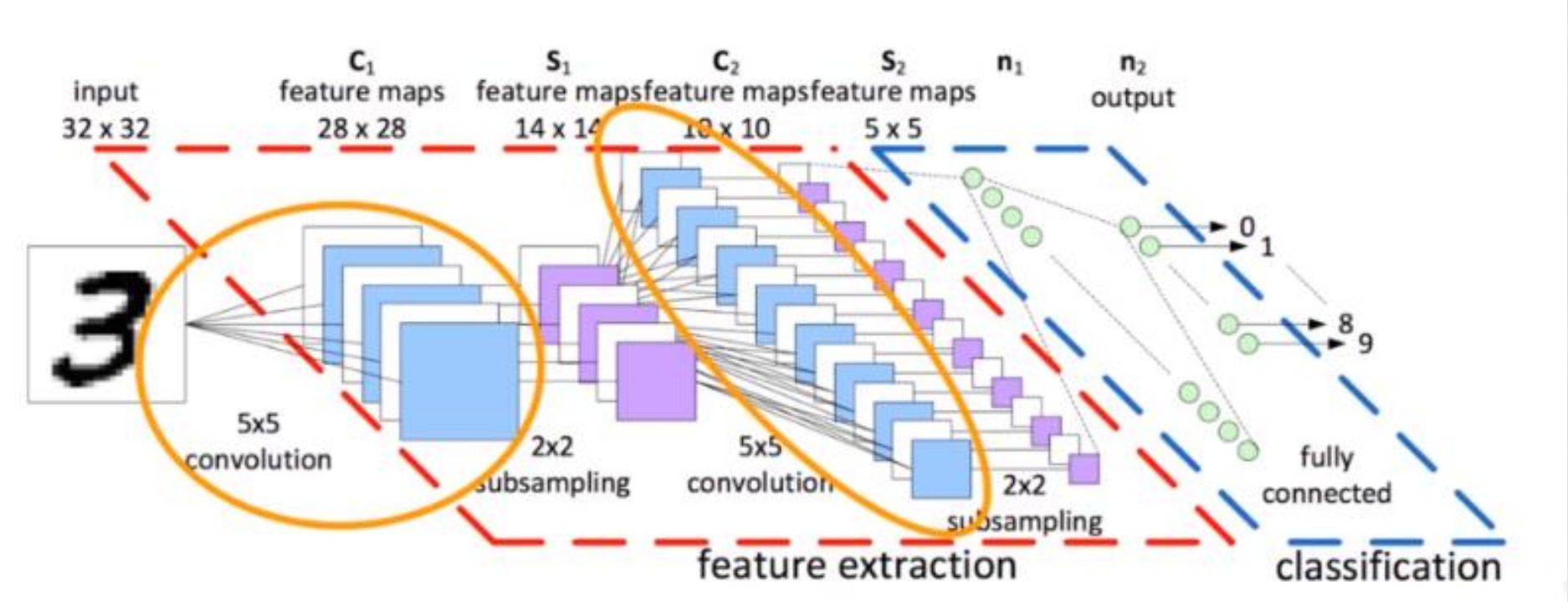
Warstwa splotowa



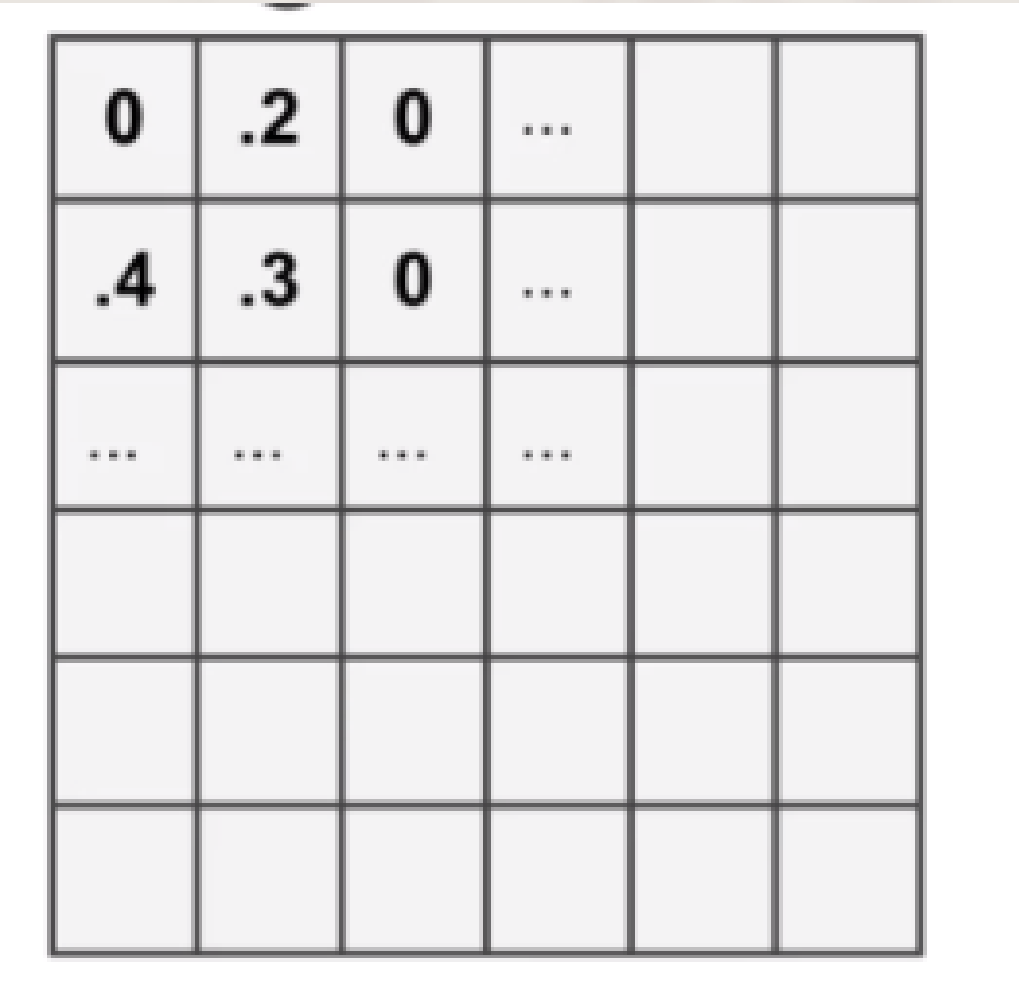
Warstwa splotowa



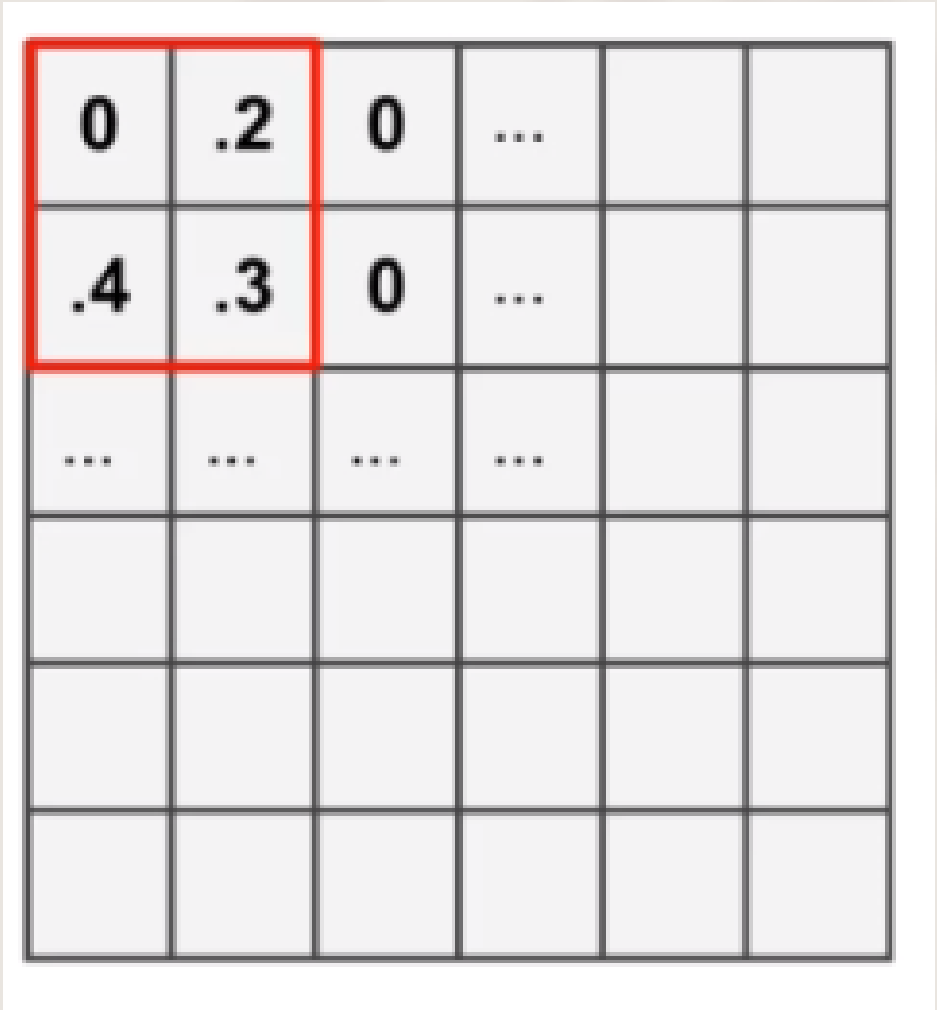
Warstwa splotowa



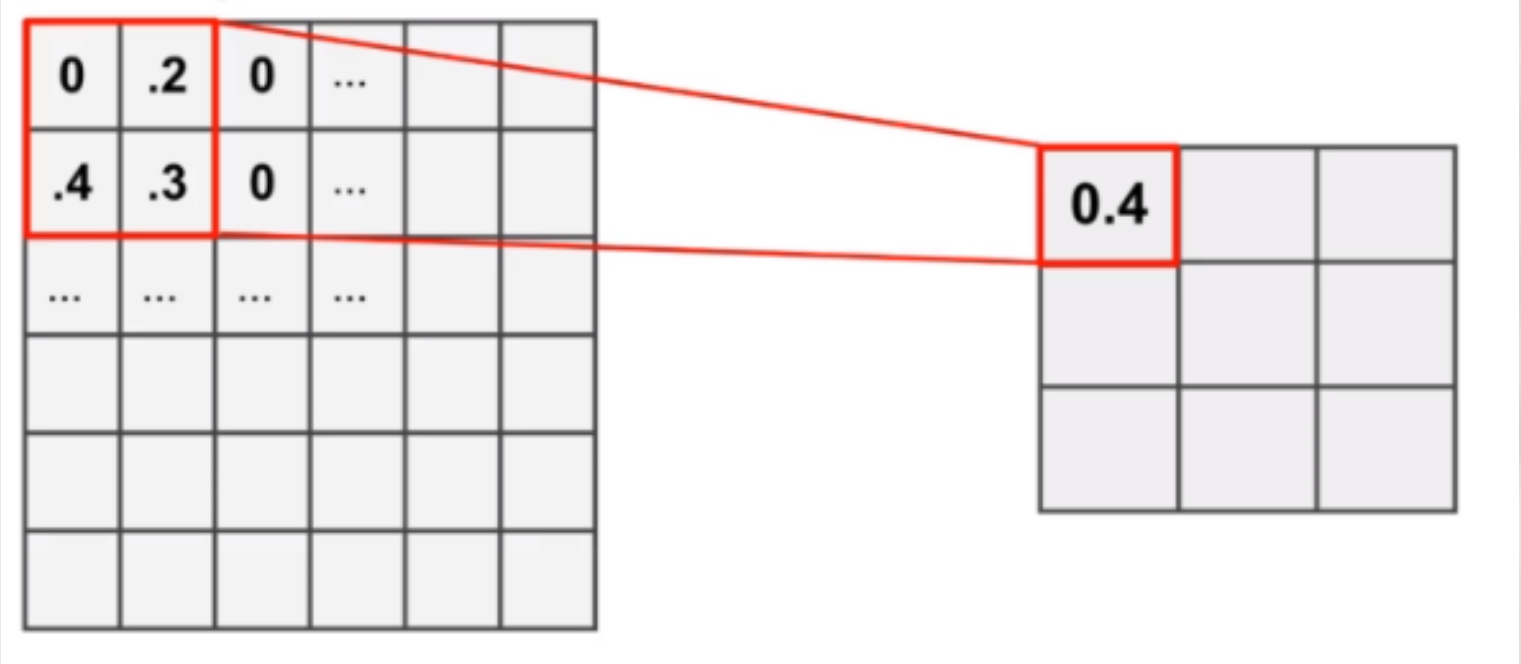
Warstwa łącząca



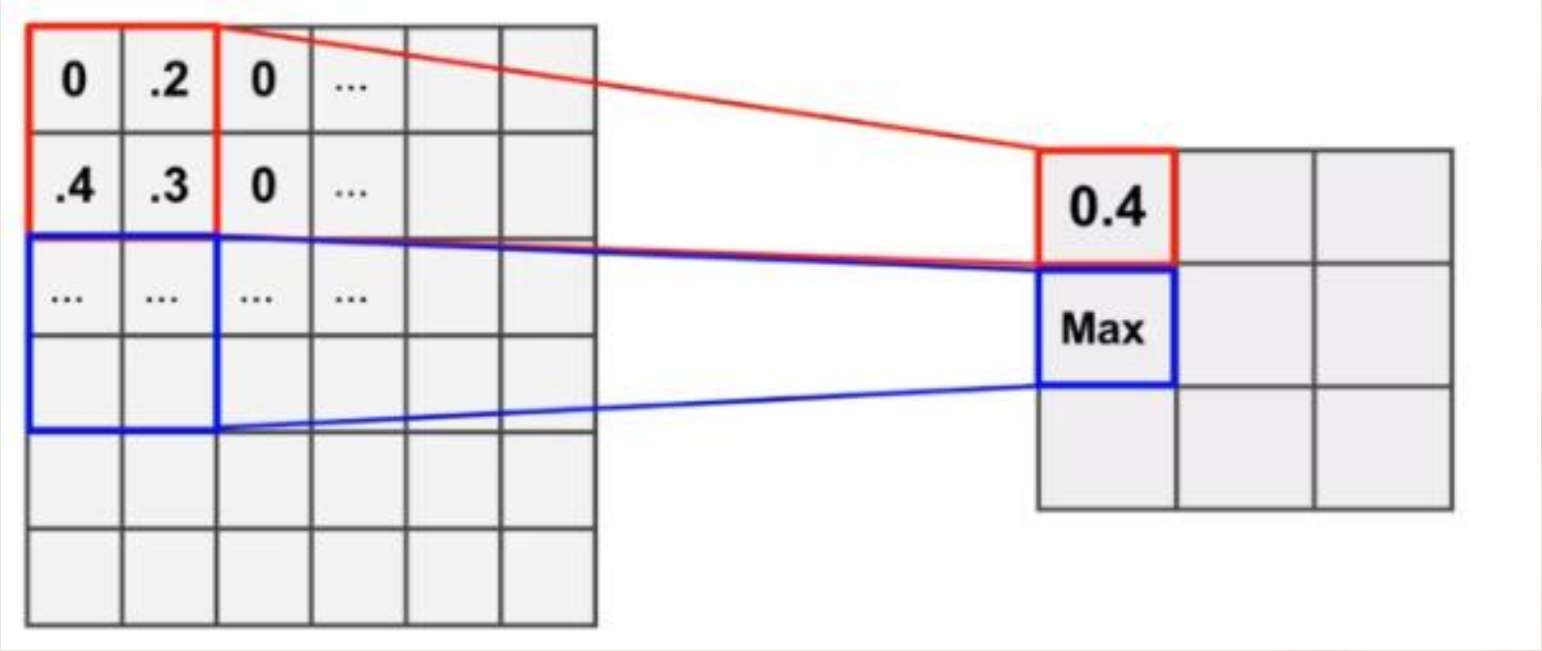
Warstwa łącząca



Warstwa łącząca

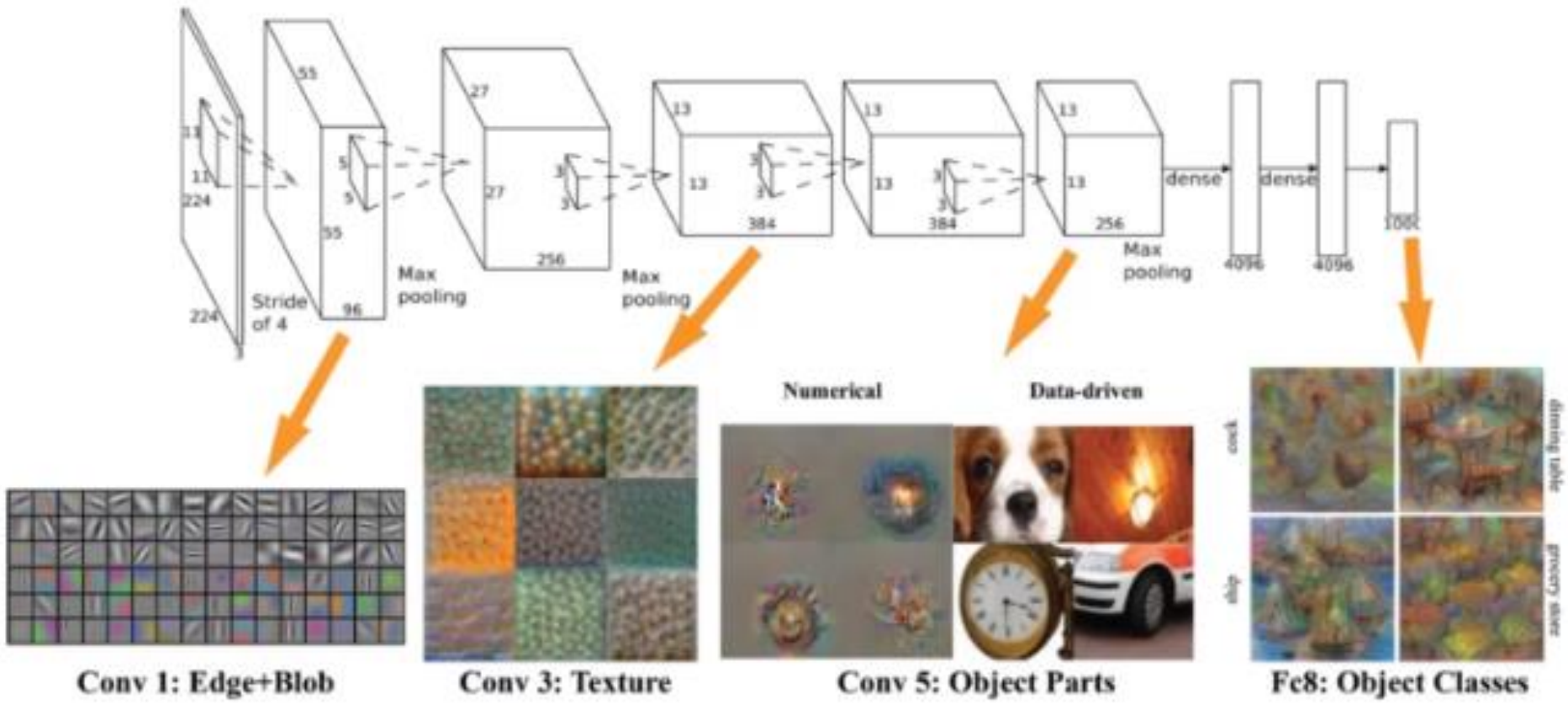


Warstwa łącząca



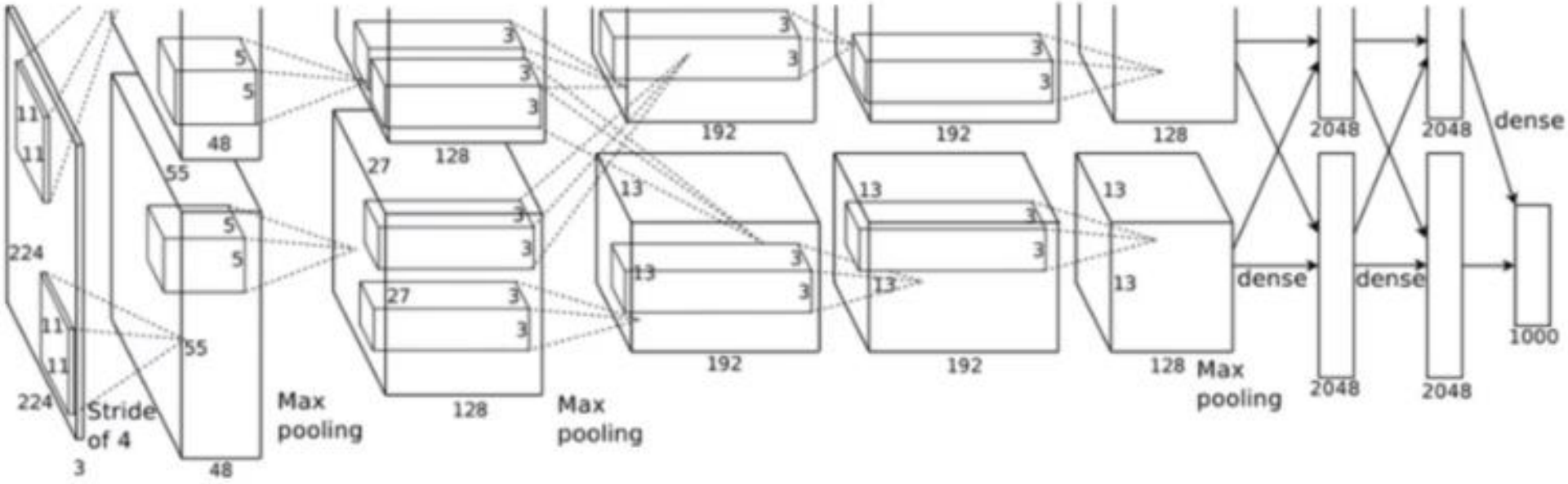
Sieci splotowe

- AlexNet



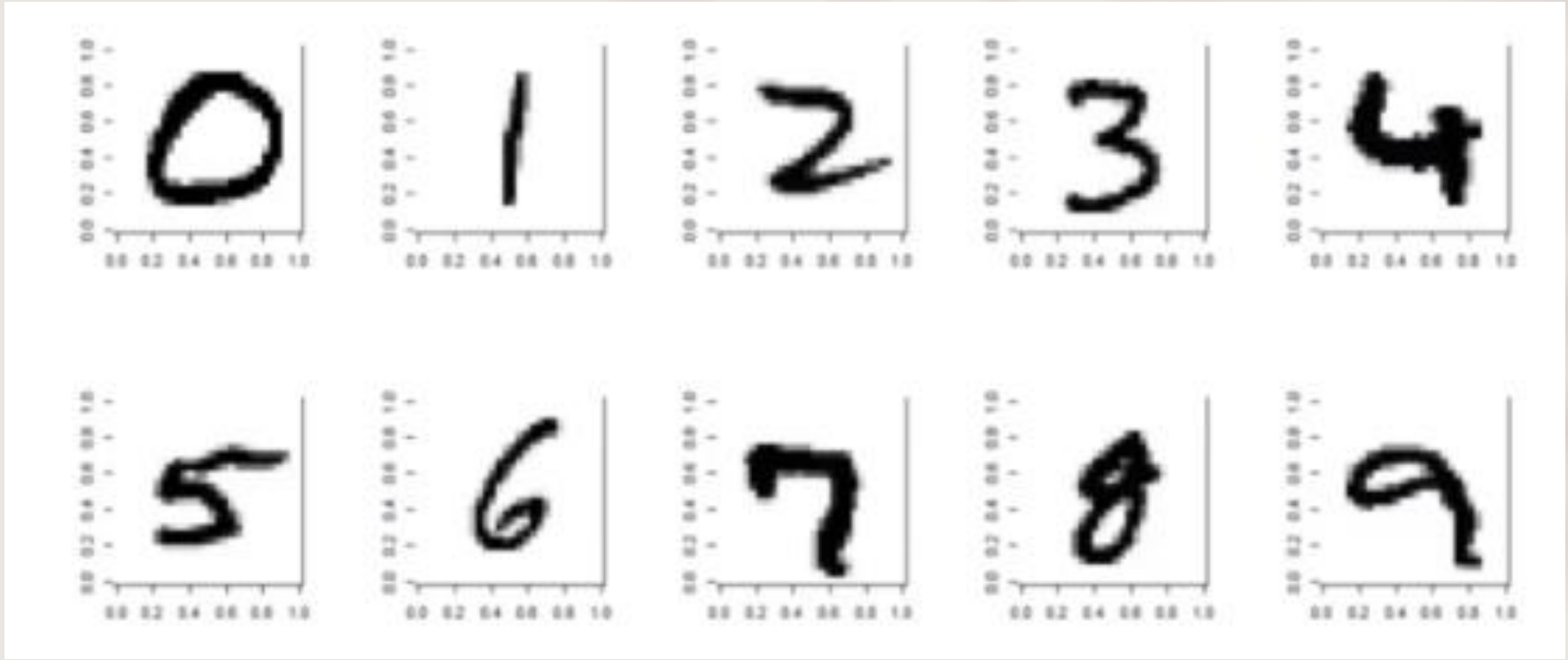
Sieci splotowe

- AlexNet

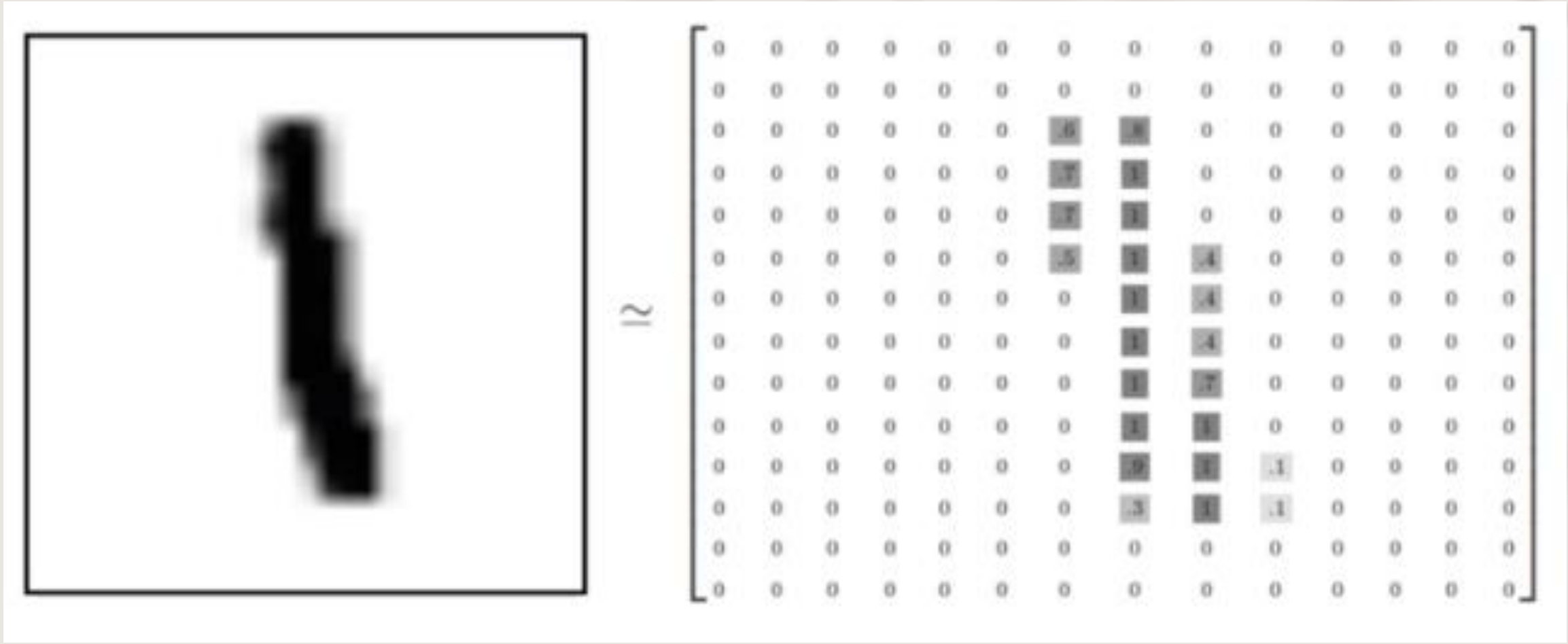


MNIST

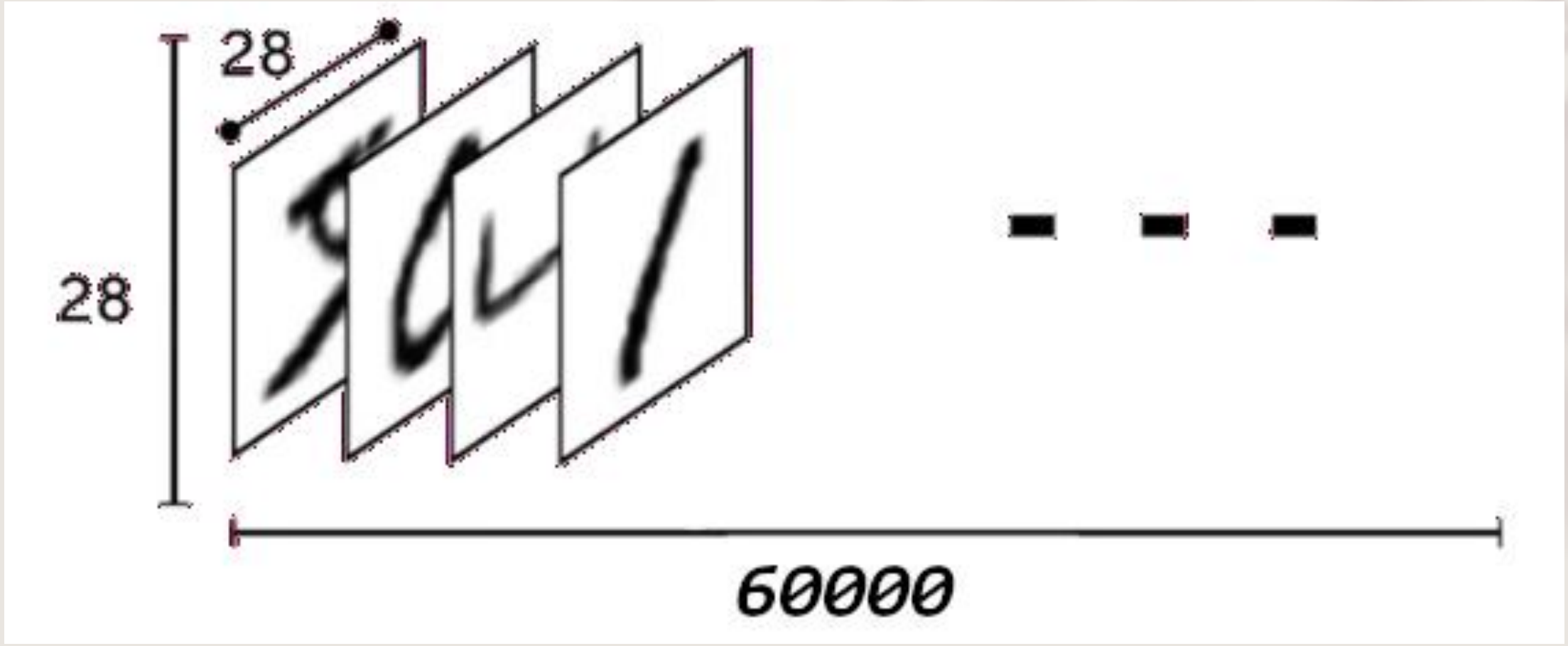
MNIST



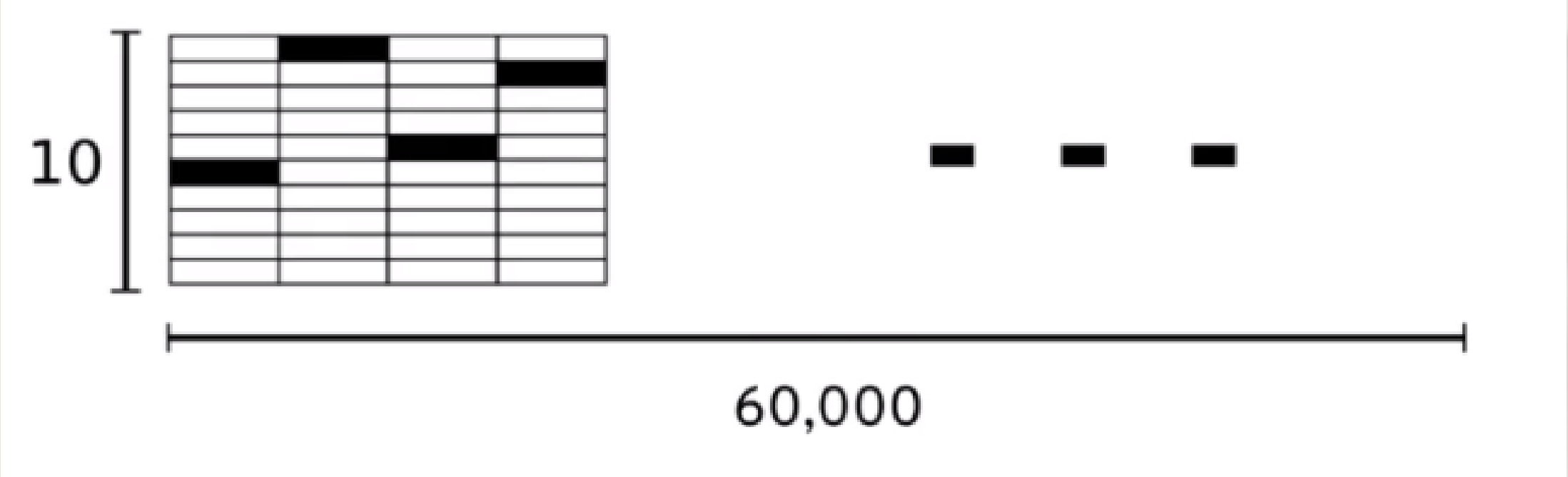
MNIST



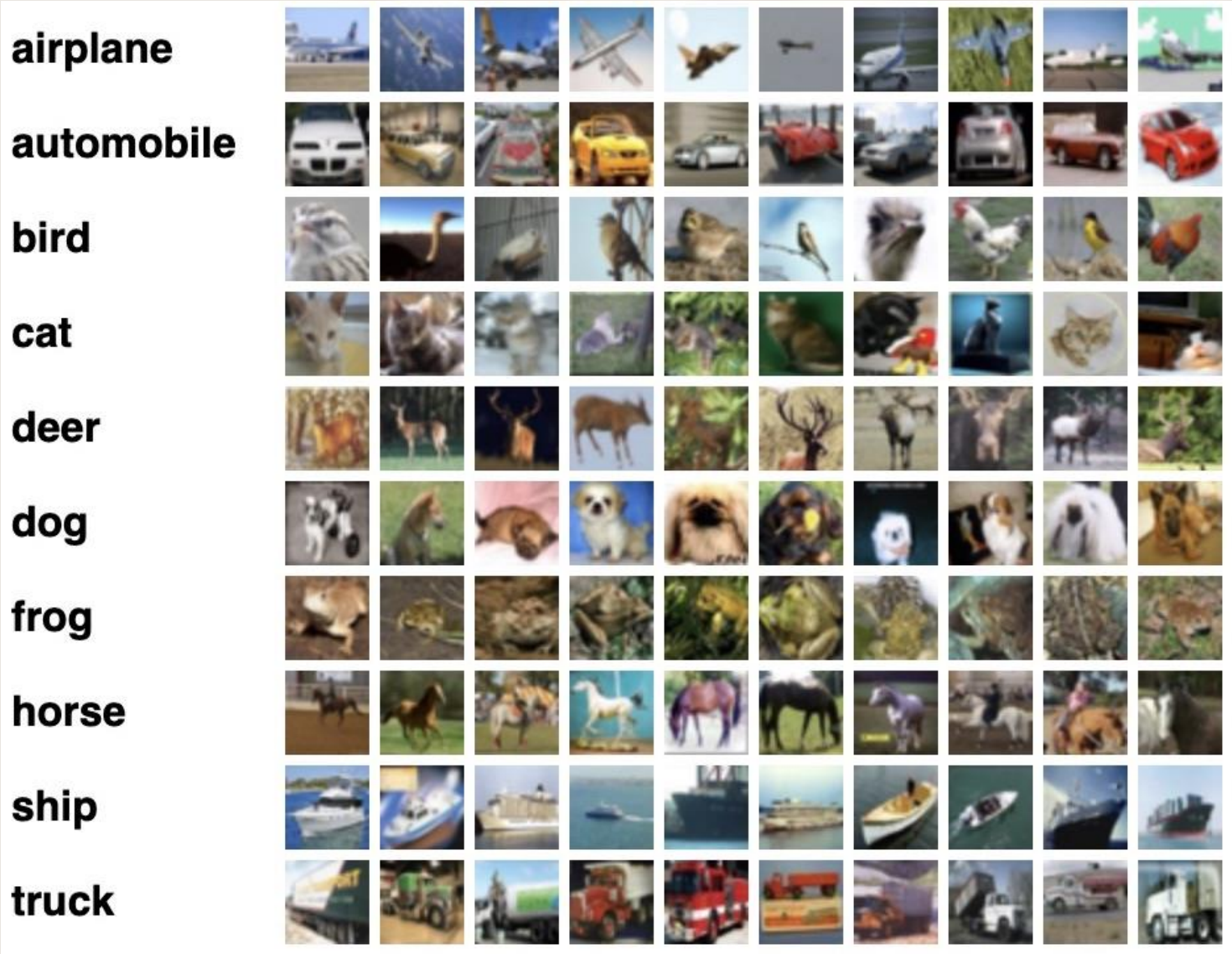
MNIST - feature



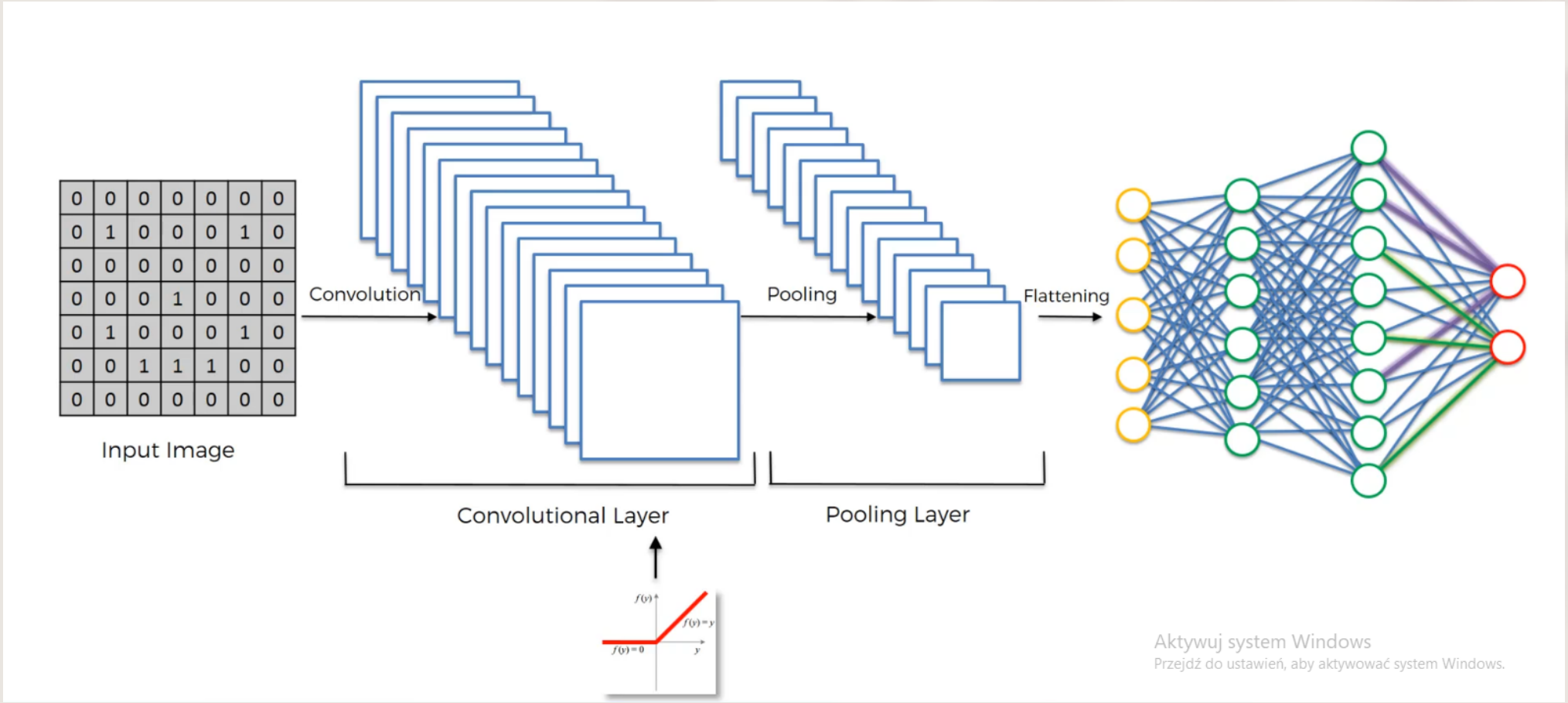
MNIST - targets



CIFAR-10


airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

Sieci splotowe

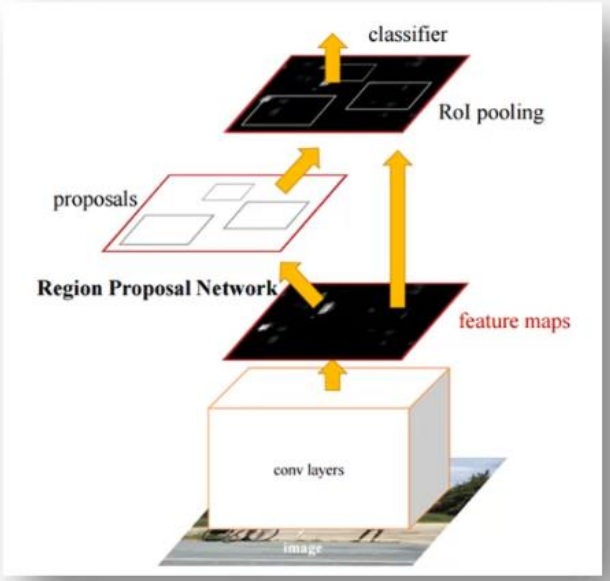


Sieci splotowe

Additional Reading:

The 9 Deep Learning Papers You Need To Know About (Understanding CNNs Part 3)

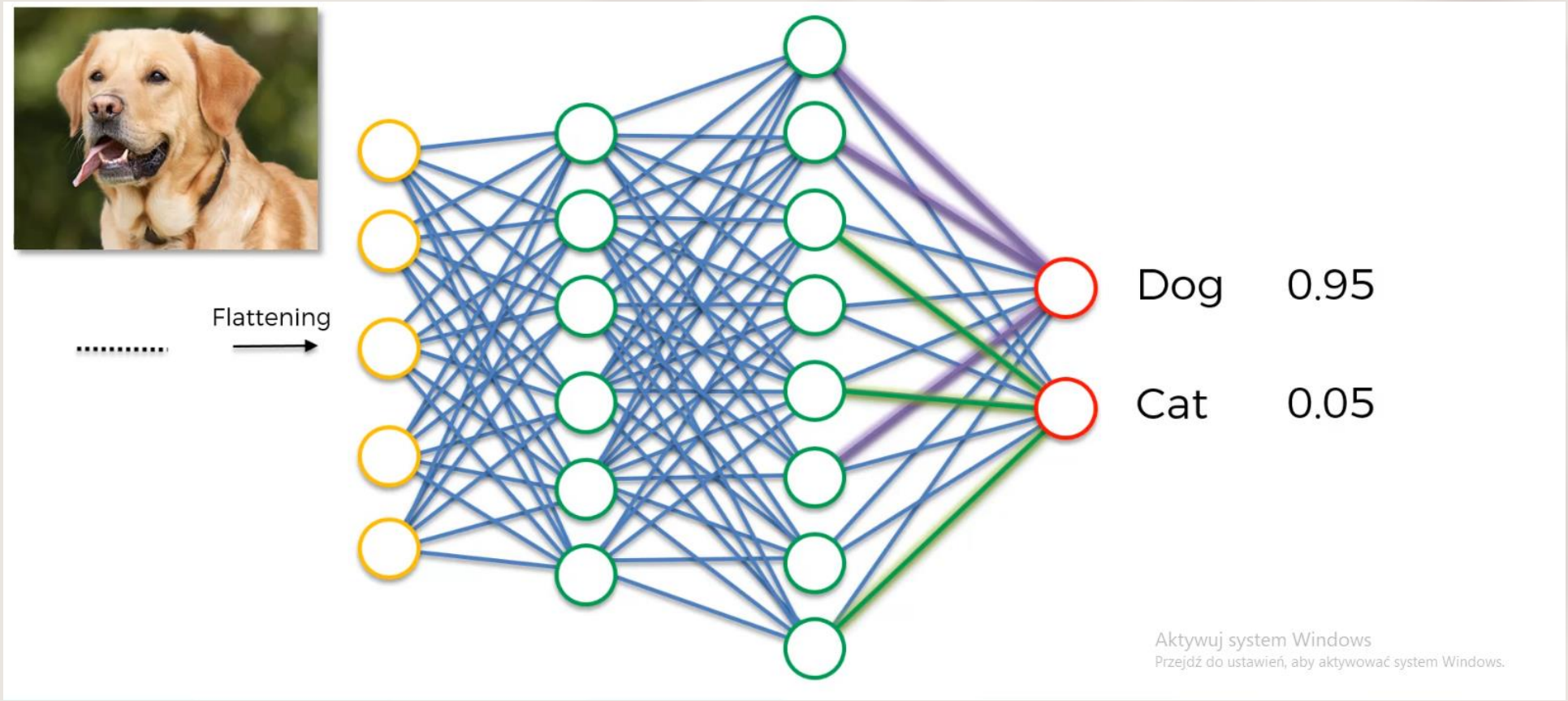
Adit Deshpande (2016)



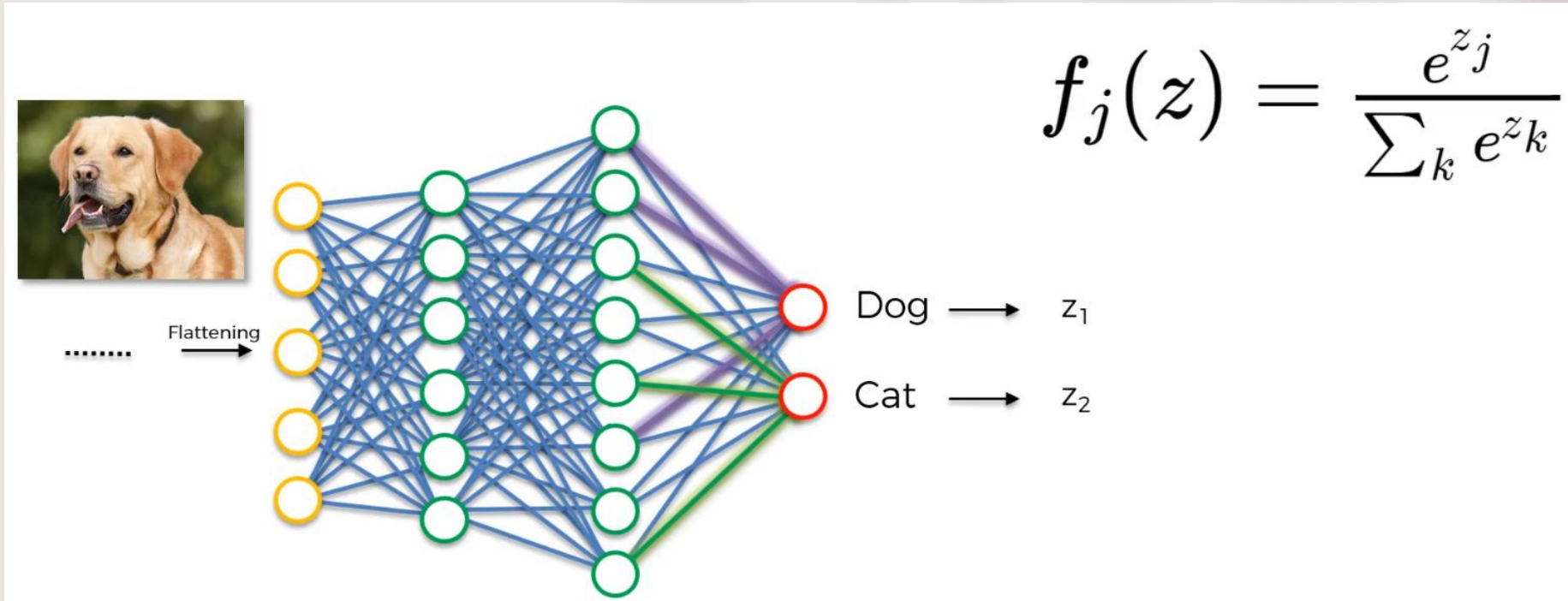
Link:

https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html

Softmax

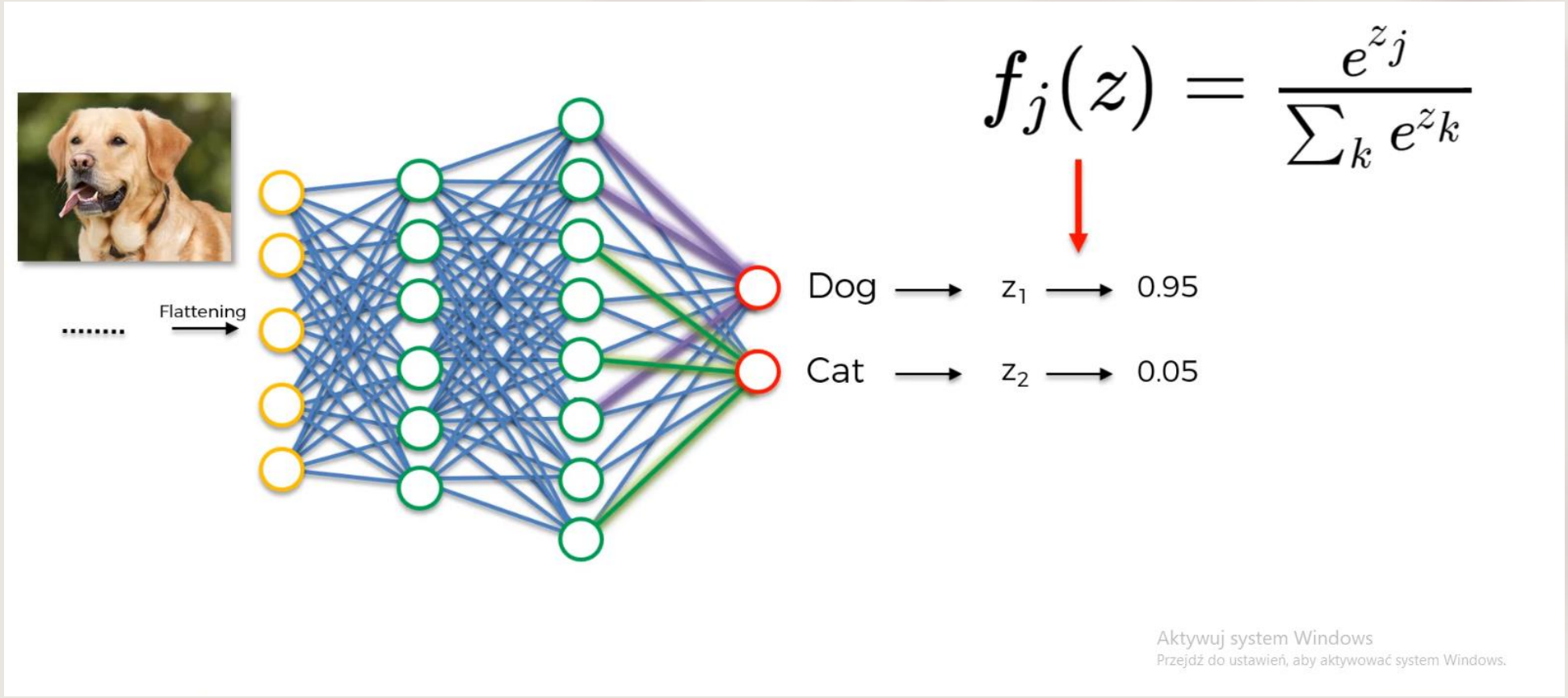


Softmax



$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}}$$

Softmax



Softmax

$$L_i = -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right)$$

Softmax

$$L_i = -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right)$$

$$H(p,q) = -\sum_x p(x)\,\log q(x)$$

Cross-entropy

Dog 0.9

Cat 0.1

Cross-entropy

Dog 0.9

Cat 0.1

1
0

Cross-entropy

Dog 0.9

Cat 0.1

$$H(p,q) = -\sum_{x} p(x) \log q(x)$$

1
0

Cross-entropy

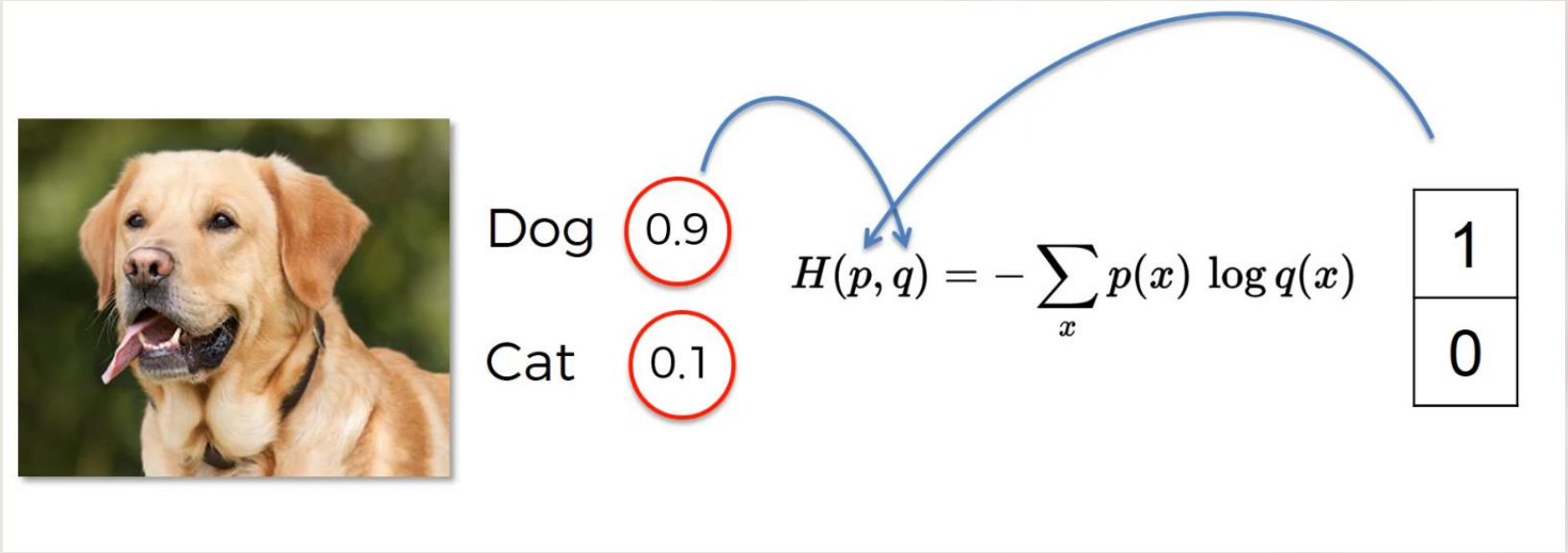
Dog 0.9

Cat 0.1

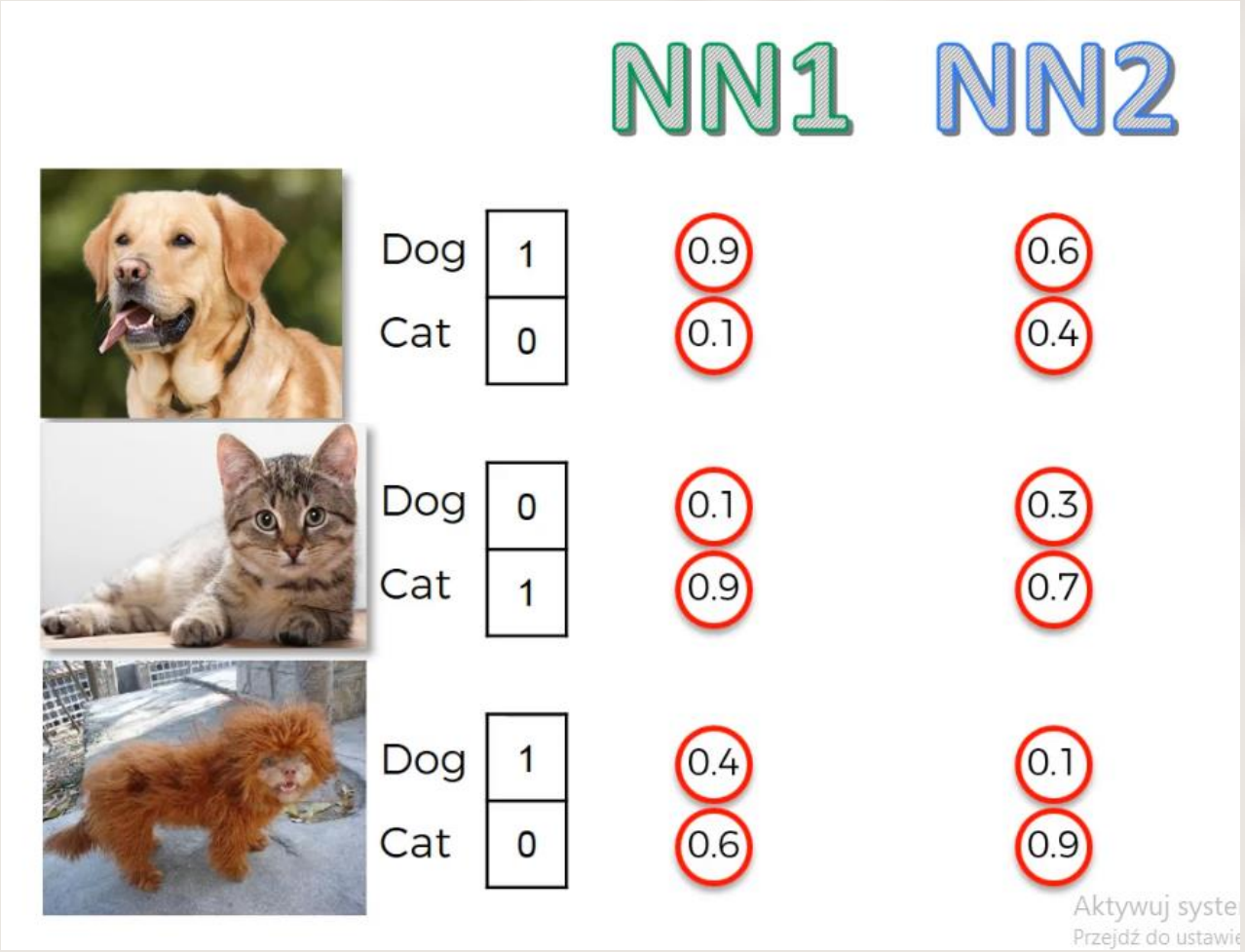
$$H(p,q) = -\sum_{x} p(x) \log q(x)$$

1
0

Cross-entropy



Cross-entropy



Cross-entropy

NN1

Row	Dog^	Cat^	Dog	Cat
#1	0.9	0.1	1	0
#2	0.1	0.9	0	1
#3	0.4	0.6	1	0

NN2

Row	Dog^	Cat^	Dog	Cat
#1	0.6	0.4	1	0
#2	0.3	0.7	0	1
#3	0.1	0.9	1	0

Cross-entropy

NN1

Row	Dog^	Cat^	Dog	Cat
#1	0.9	0.1	1	0
#2	0.1	0.9	0	1
#3	0.4	0.6	1	0

NN2

Row	Dog^	Cat^	Dog	Cat
#1	0.6	0.4	1	0
#2	0.3	0.7	0	1
#3	0.1	0.9	1	0

Classification Error

1/3 = 0.33 | 1/3 = 0.33

Mean Squared Error

0.25 | 0.71

Cross-Entropy

0.38 | 1.06

Cross-entropy

Additional Reading:

How to implement a neural network Intermezzo 2

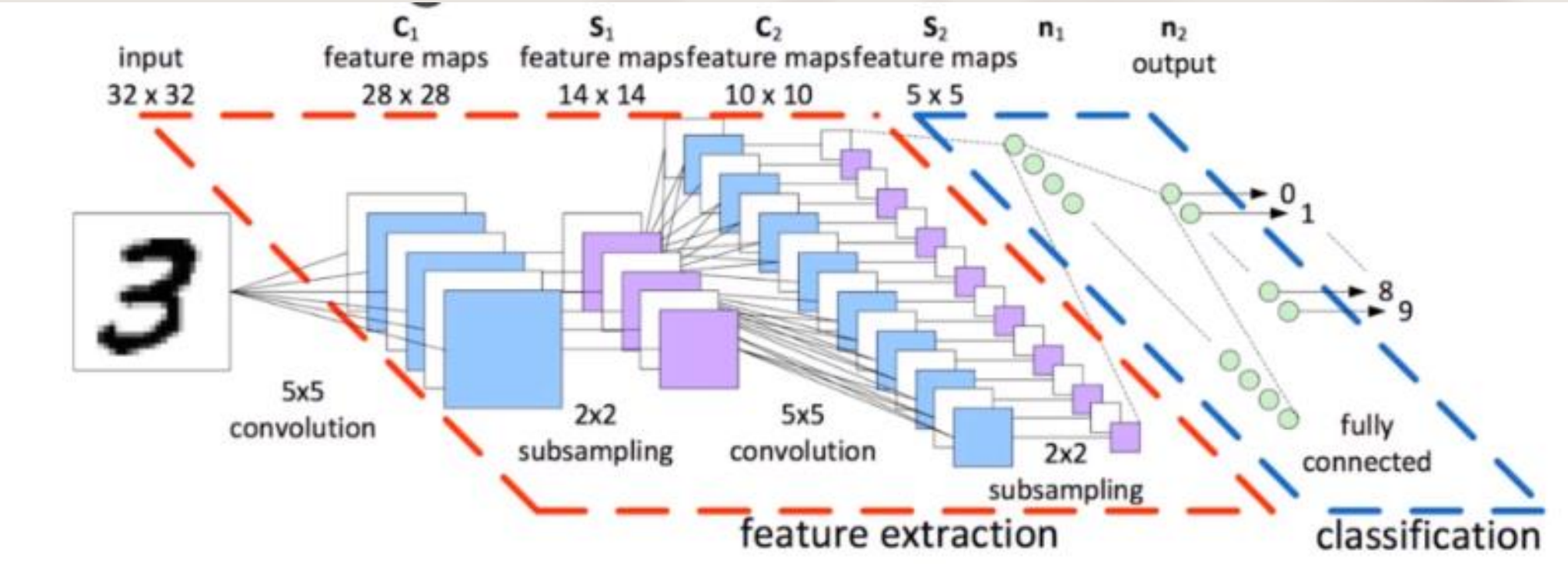
By Peter Roelants (2016)

$$
\begin{aligned}
\frac{\partial \xi}{\partial z_i} &= -\sum_{j=1}^{C} \frac{\partial t_j log(y_j)}{\partial z_i} = -\sum_{j=1}^{C} t_j \frac{\partial log(y_j)}{\partial z_i} = -\sum_{j=1}^{C} t_j \frac{1}{y_j} \frac{\partial y_j}{\partial z_i} \\
&= -\frac{t_i}{y_i} \frac{\partial y_i}{\partial z_i} - \sum_{j \neq i}^{C} \frac{t_j}{y_j} \frac{\partial y_j}{\partial z_i} = -\frac{t_i}{y_i} y_i (1 - y_i) - \sum_{j \neq i}^{C} \frac{t_j}{y_j} (-y_j y_i) \\
&= -t_i + t_i y_i + \sum_{j \neq i}^{C} t_j y_i = -t_i + \sum_{j=1}^{C} t_j y_i = -t_i + y_i \sum_{j=1}^{C} t_j \\
&= y_i - t_i
\end{aligned}
$$

Link:

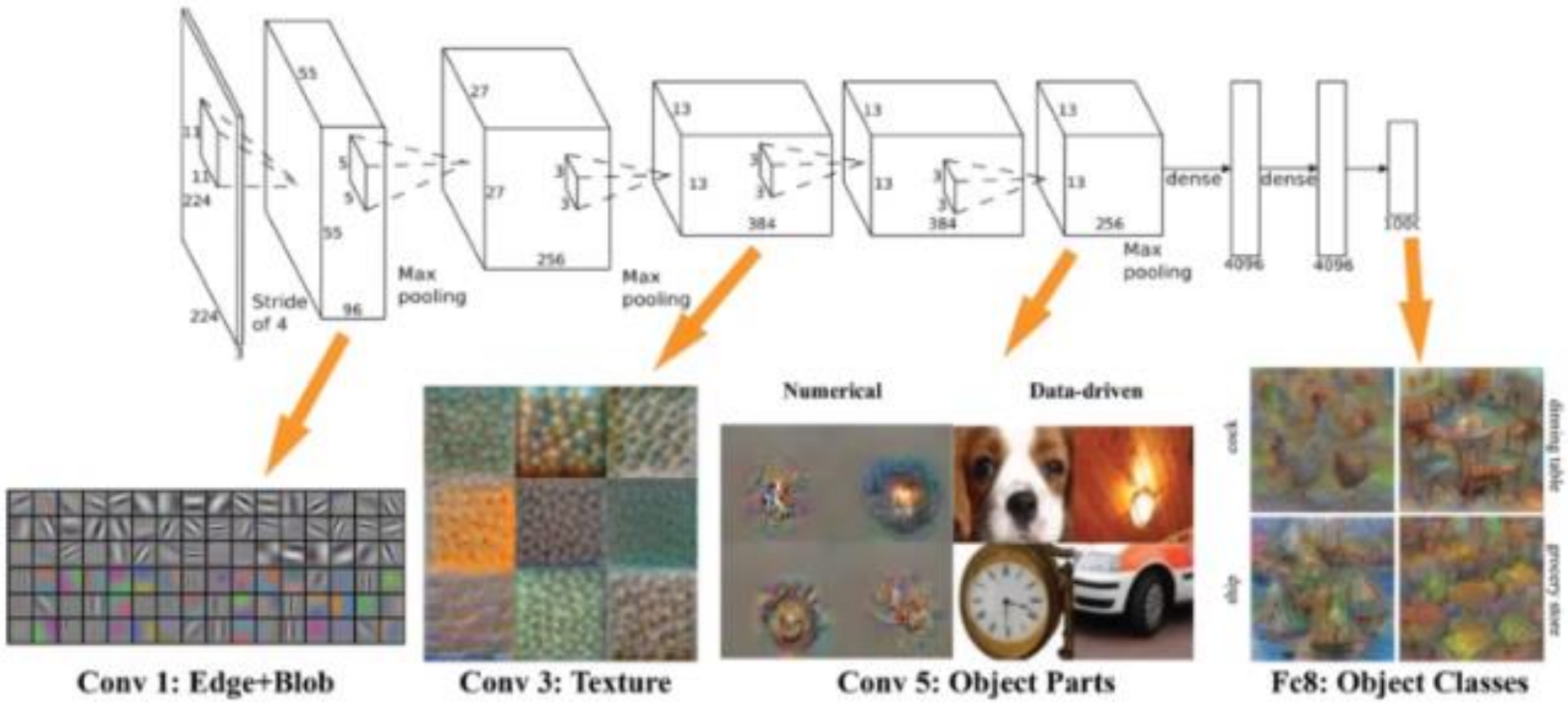
http://peterroelants.github.io/posts/neural_network_implementation_intermezzo02/

Sieci splotowe



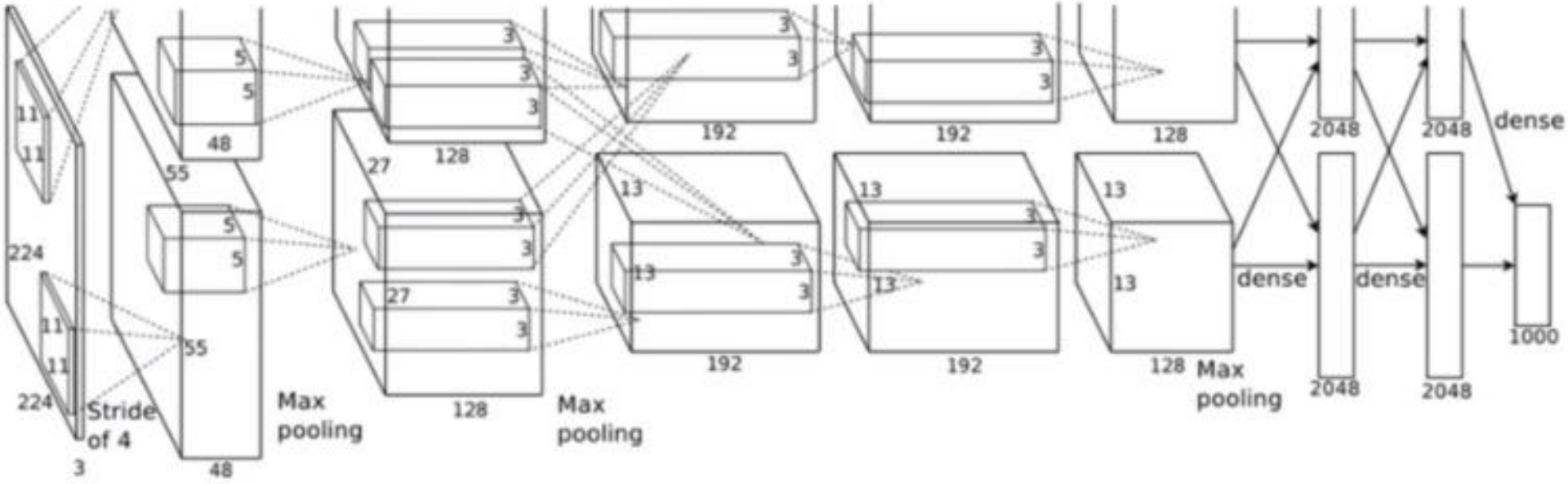
Sieci splotowe

- AlexNet



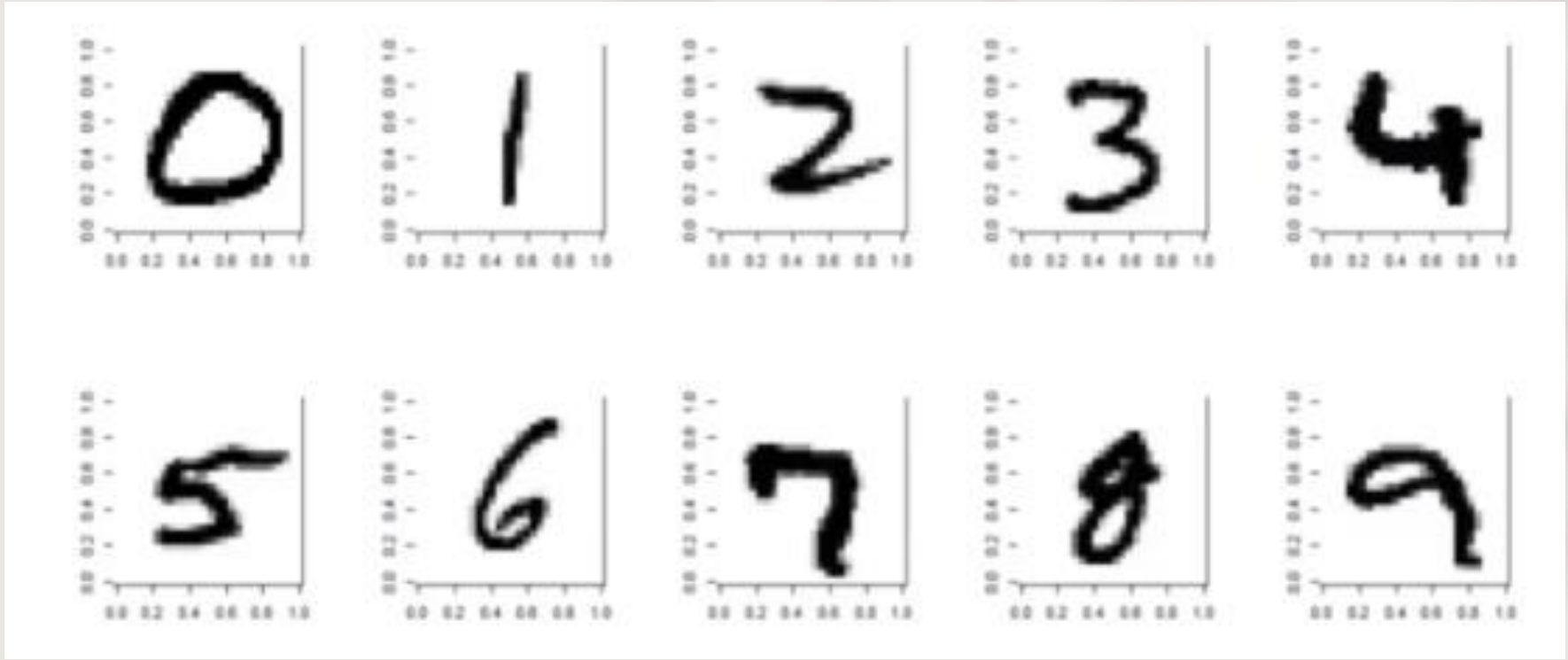
Sieci splotowe

- AlexNet

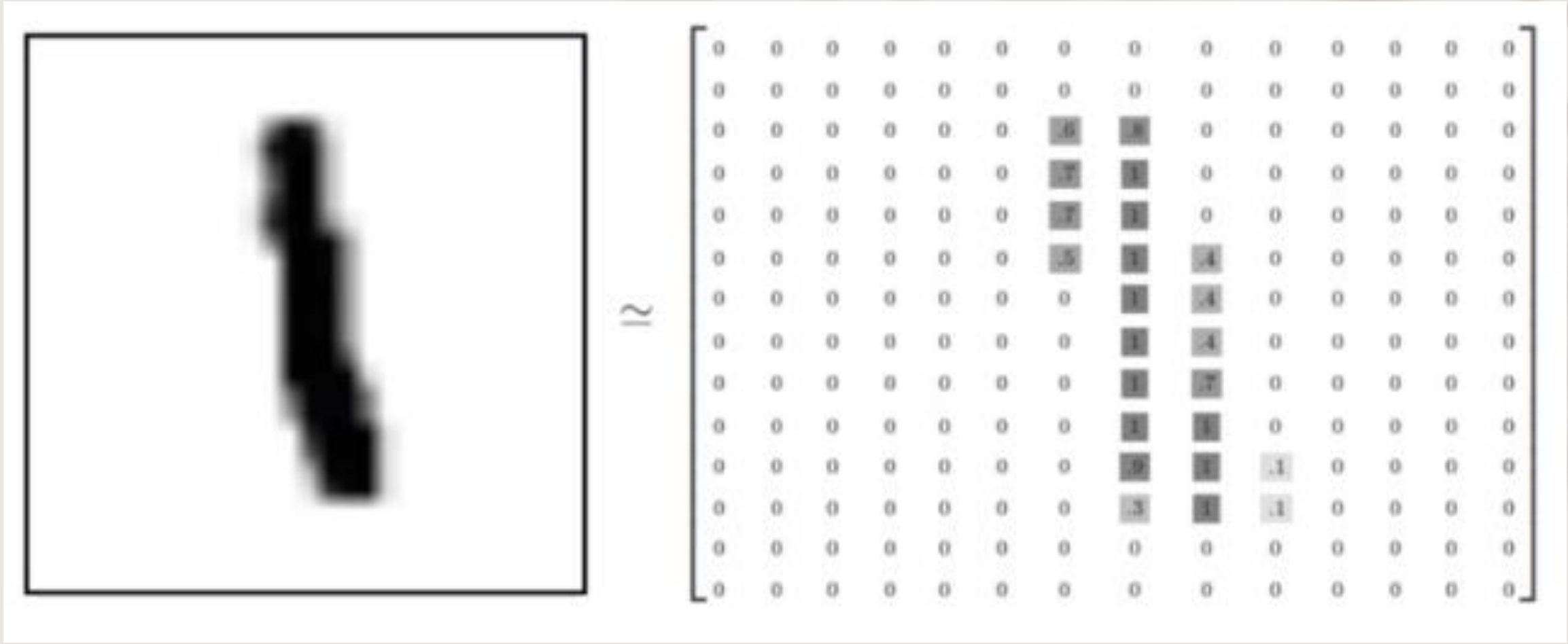


MNIST

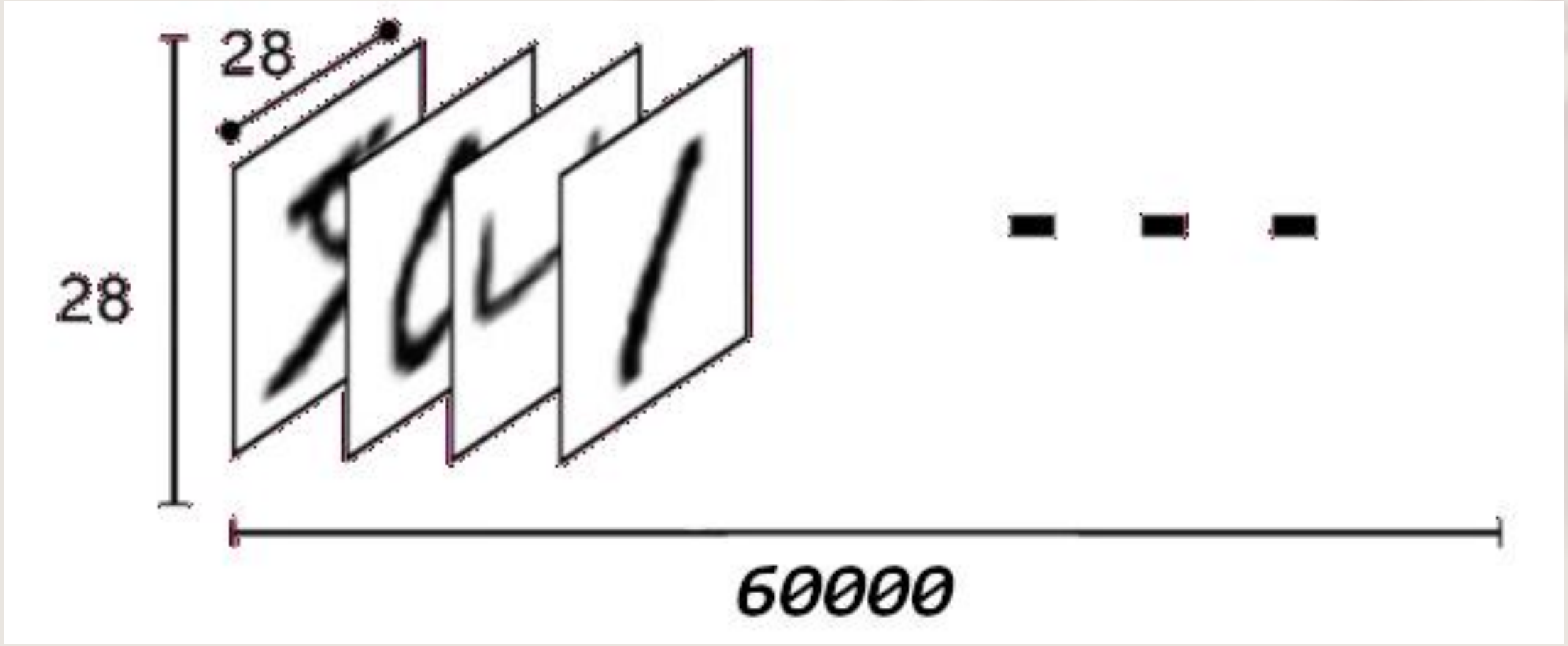
MNIST



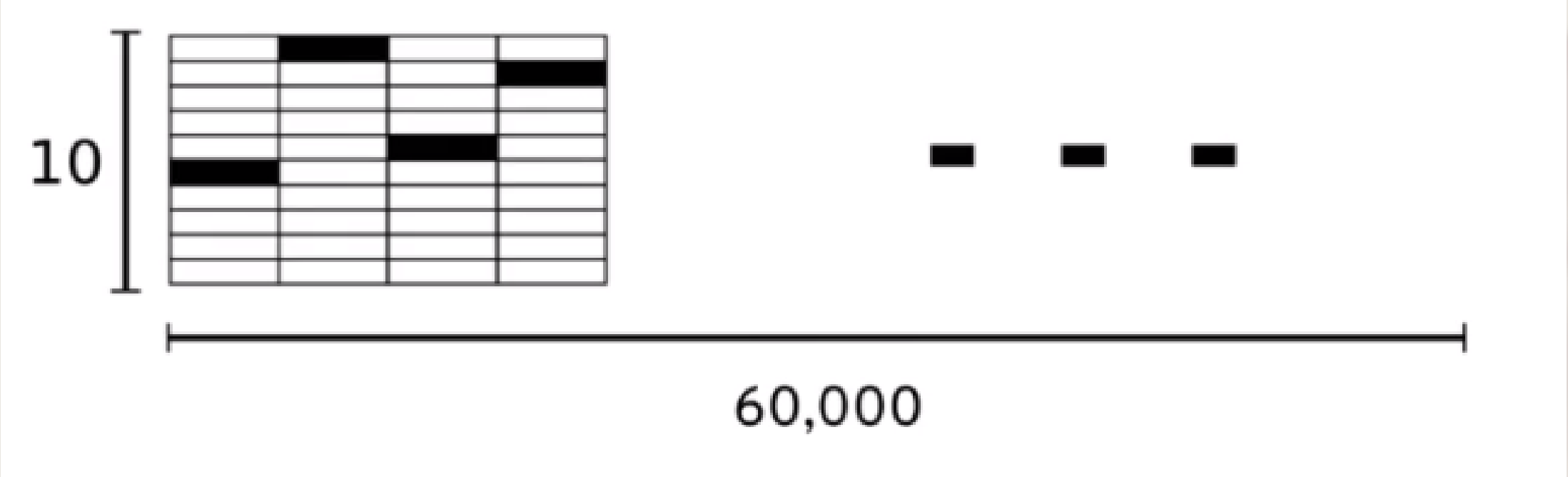
MNIST



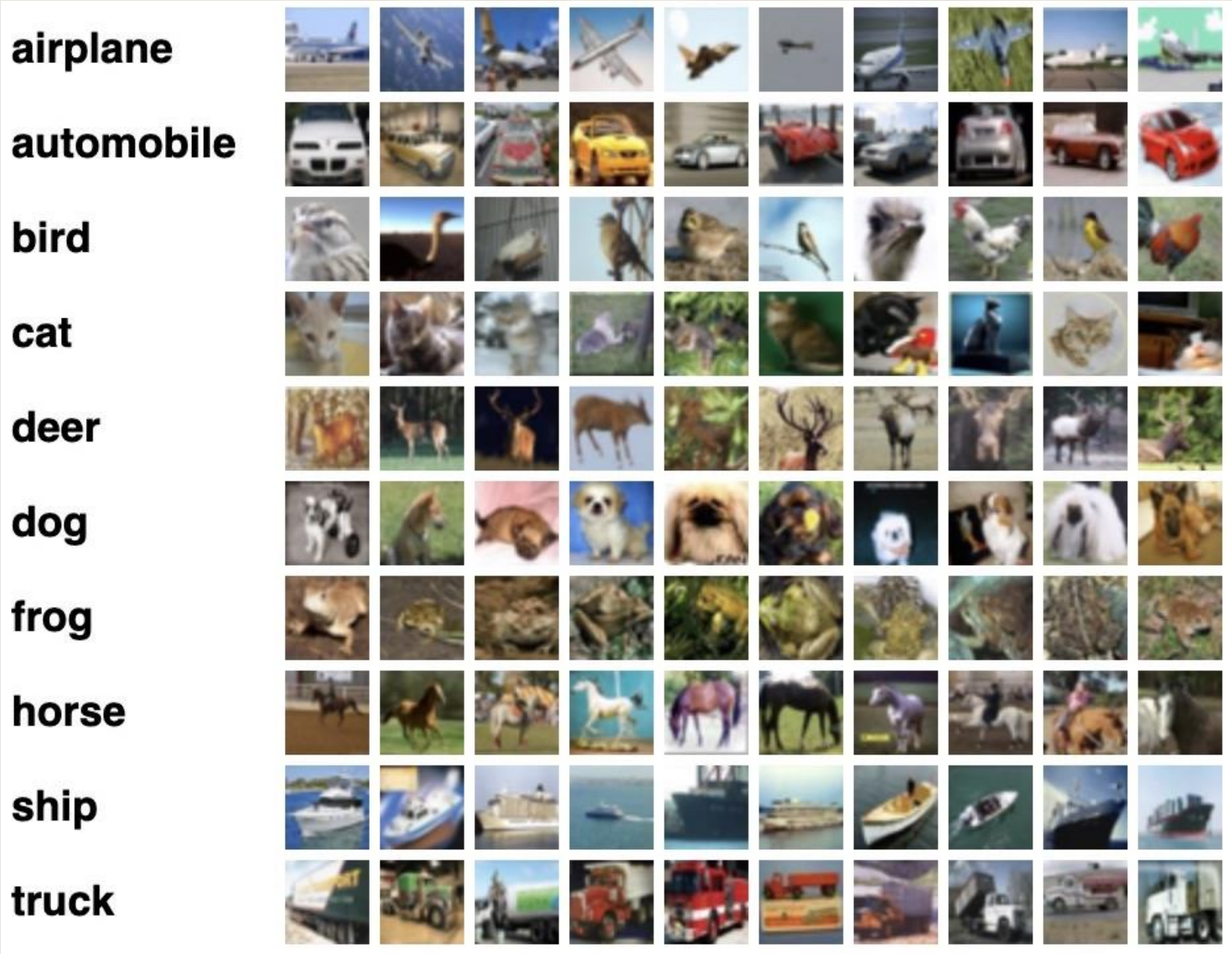
MNIST - feature



MNIST - targets



CIFAR-10


airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

LeNet

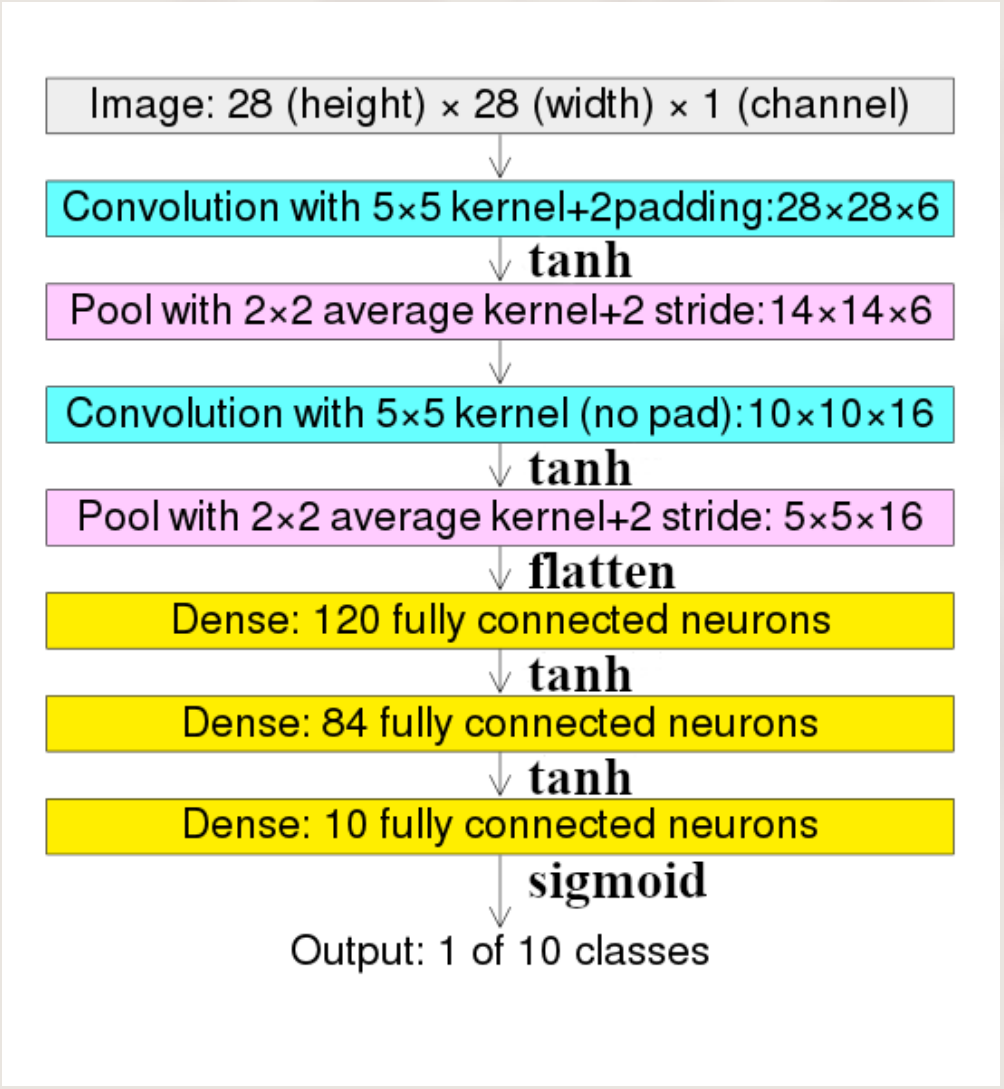
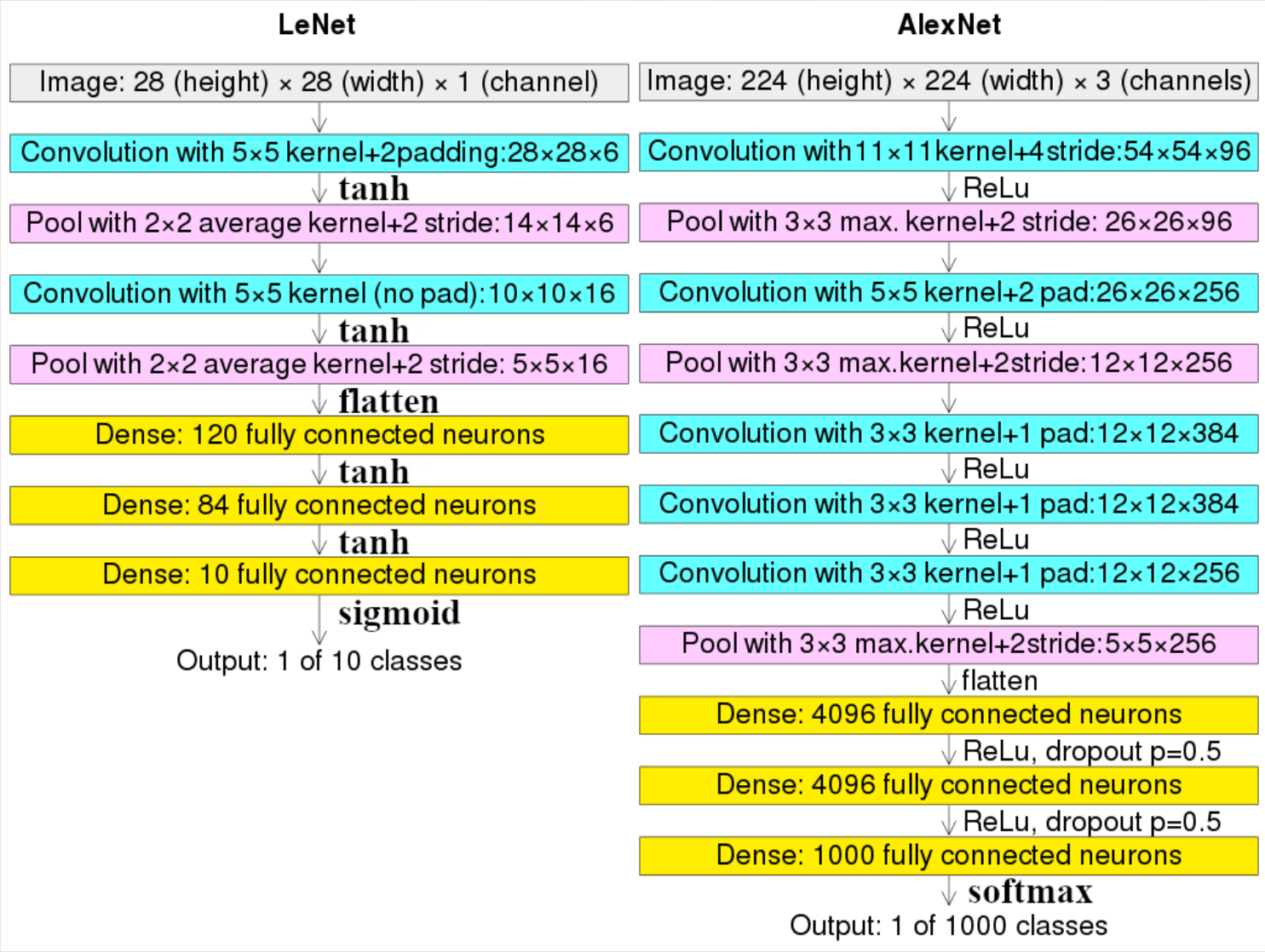


Image: 28 (height) × 28 (width) × 1 (channel)
Convolution with 5×5 kernel+2padding:28×28×6
tanh
Pool with 2×2 average kernel+2 stride:14×14×6
Convolution with 5×5 kernel (no pad):10×10×16
tanh
Pool with 2×2 average kernel+2 stride: 5×5×16
flatten
Dense: 120 fully connected neurons
tanh
Dense: 84 fully connected neurons
tanh
Dense: 10 fully connected neurons
sigmoid
Output: 1 of 10 classes

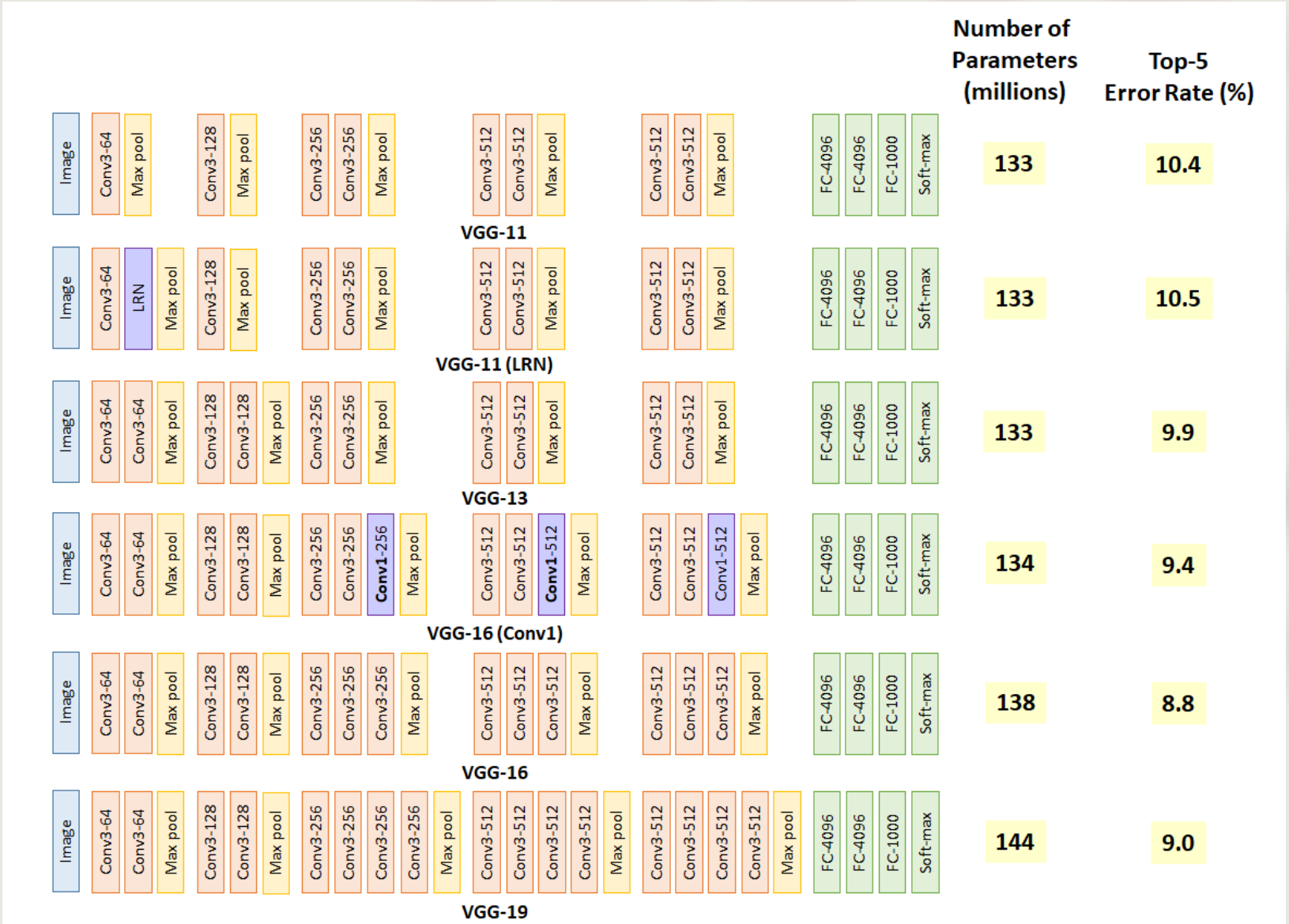
AlexNet


LeNet
Image: 28 (height) × 28 (width) × 1 (channel)
Convolution with 5×5 kernel+2padding:28×28×6
tanh
Pool with 2×2 average kernel+2 stride:14×14×6
Convolution with 5×5 kernel (no pad):10×10×16
tanh
Pool with 2×2 average kernel+2 stride: 5×5×16
flatten
Dense: 120 fully connected neurons
tanh
Dense: 84 fully connected neurons
tanh
Dense: 10 fully connected neurons
sigmoid
Output: 1 of 10 classes
AlexNet
Image: 224 (height) × 224 (width) × 3 (channels)
Convolution with 11×11 kernel+4 stride:54×54×96
ReLu
Pool with 3×3 max. kernel+2 stride: 26×26×96
Convolution with 5×5 kernel+2 pad:26×26×256
ReLu
Pool with 3×3 max.kernel+2stride:12×12×256
Convolution with 3×3 kernel+1 pad:12×12×384
ReLu
Convolution with 3×3 kernel+1 pad:12×12×384
ReLu
Convolution with 3×3 kernel+1 pad:12×12×256
ReLu
Pool with 3×3 max.kernel+2stride:5×5×256
flatten
Dense: 4096 fully connected neurons
ReLu, dropout p=0.5
Dense: 4096 fully connected neurons
ReLu, dropout p=0.5
Dense: 1000 fully connected neurons
softmax
Output: 1 of 1000 classes

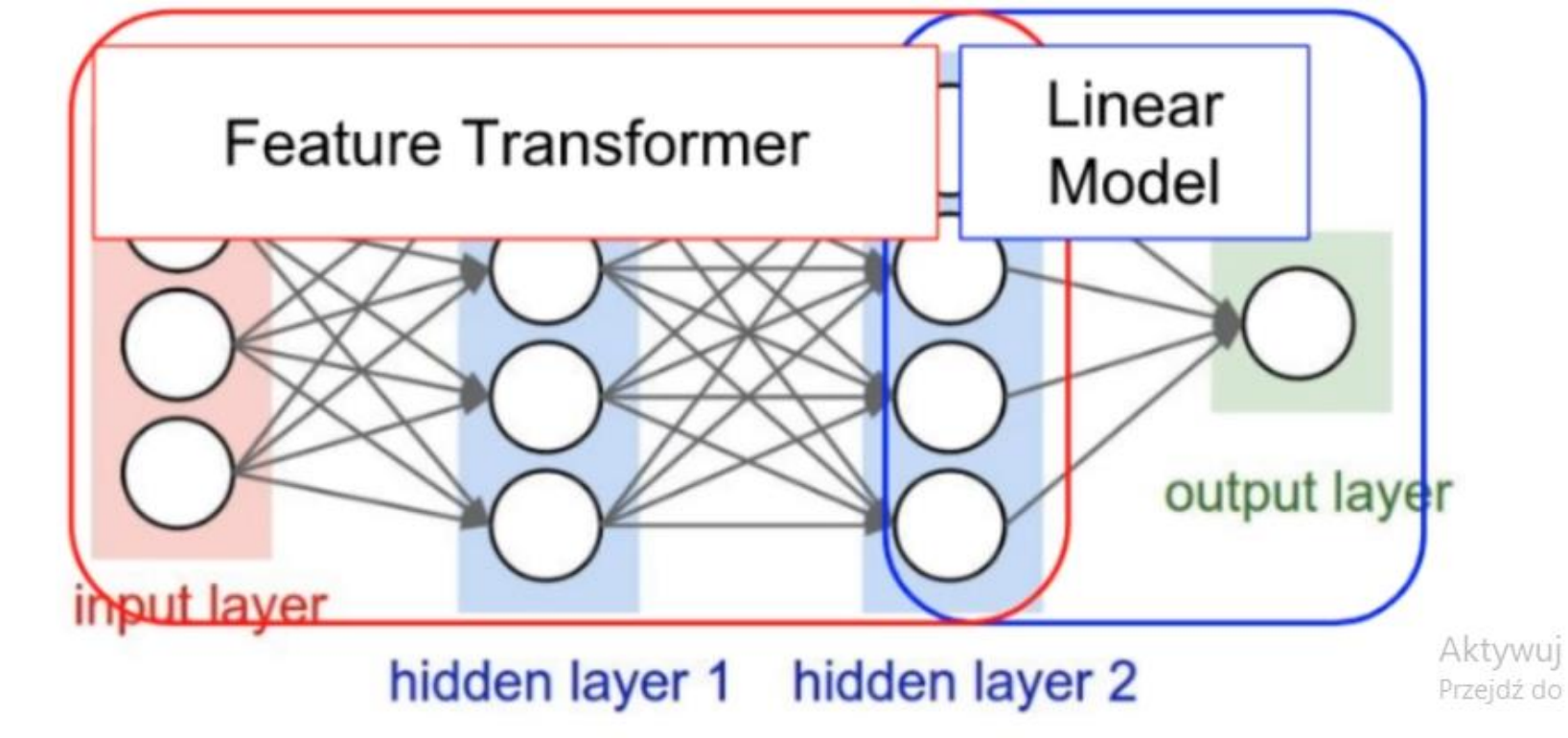
VGG


3x3 conv, 64
3x3 conv, 64
pool/2
3x3 conv, 128
3x3 conv, 128
pool/2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
pool/2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
pool/2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
pool/2
fc 4096
fc 4096
fc 4096
Size:224
Size:112
Size:56
Size:28
Size:14
Size:7

VGG

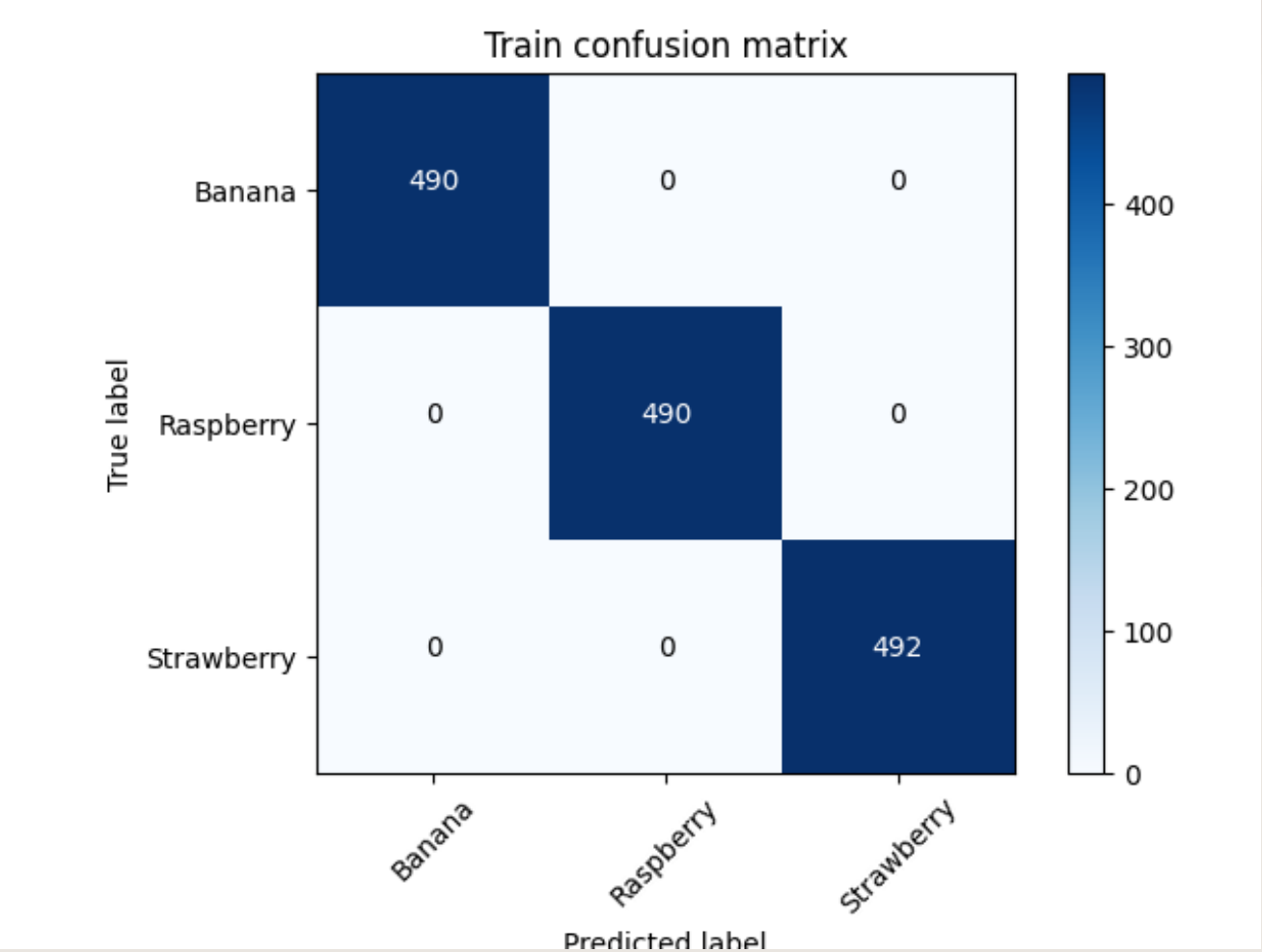

Number of Parameters (millions)
Top-5 Error Rate (%)
Image
Conv3-64
Max pool
Conv3-128
Max pool
Conv3-256
Conv3-256
Max pool
Conv3-512
Conv3-512
Max pool
Conv3-512
Conv3-512
Max pool
FC-4096
FC-4096
FC-1000
Soft-max
133
10.4
VGG-11
Image
Conv3-64
LRN
Max pool
Conv3-128
Max pool
Conv3-256
Conv3-256
Max pool
Conv3-512
Conv3-512
Max pool
Conv3-512
Conv3-512
Max pool
FC-4096
FC-4096
FC-1000
Soft-max
133
10.5
VGG-11 (LRN)
Image
Conv3-64
Conv3-64
Max pool
Conv3-128
Conv3-128
Max pool
Conv3-256
Conv3-256
Max pool
Conv3-512
Conv3-512
Max pool
Conv3-512
Conv3-512
Max pool
FC-4096
FC-4096
FC-1000
Soft-max
133
9.9
VGG-13
Image
Conv3-64
Conv3-64
Max pool
Conv3-128
Conv3-128
Max pool
Conv3-256
Conv3-256
Conv1-256
Max pool
Conv3-512
Conv3-512
Conv1-512
Max pool
Conv3-512
Conv3-512
Conv1-512
Max pool
FC-4096
FC-4096
FC-1000
Soft-max
134
9.4
VGG-16 (Conv1)
Image
Conv3-64
Conv3-64
Max pool
Conv3-128
Conv3-128
Max pool
Conv3-256
Conv3-256
Conv3-256
Max pool
Conv3-512
Conv3-512
Conv3-512
Max pool
Conv3-512
Conv3-512
Conv3-512
Max pool
FC-4096
FC-4096
FC-1000
Soft-max
138
8.8
VGG-16
Image
Conv3-64
Conv3-64
Max pool
Conv3-128
Conv3-128
Max pool
Conv3-256
Conv3-256
Conv3-256
Conv3-256
Max pool
Conv3-512
Conv3-512
Conv3-512
Conv3-512
Max pool
Conv3-512
Conv3-512
Conv3-512
Conv3-512
Max pool
FC-4096
FC-4096
FC-1000
Soft-max
144
9.0
VGG-19

Transfer learning

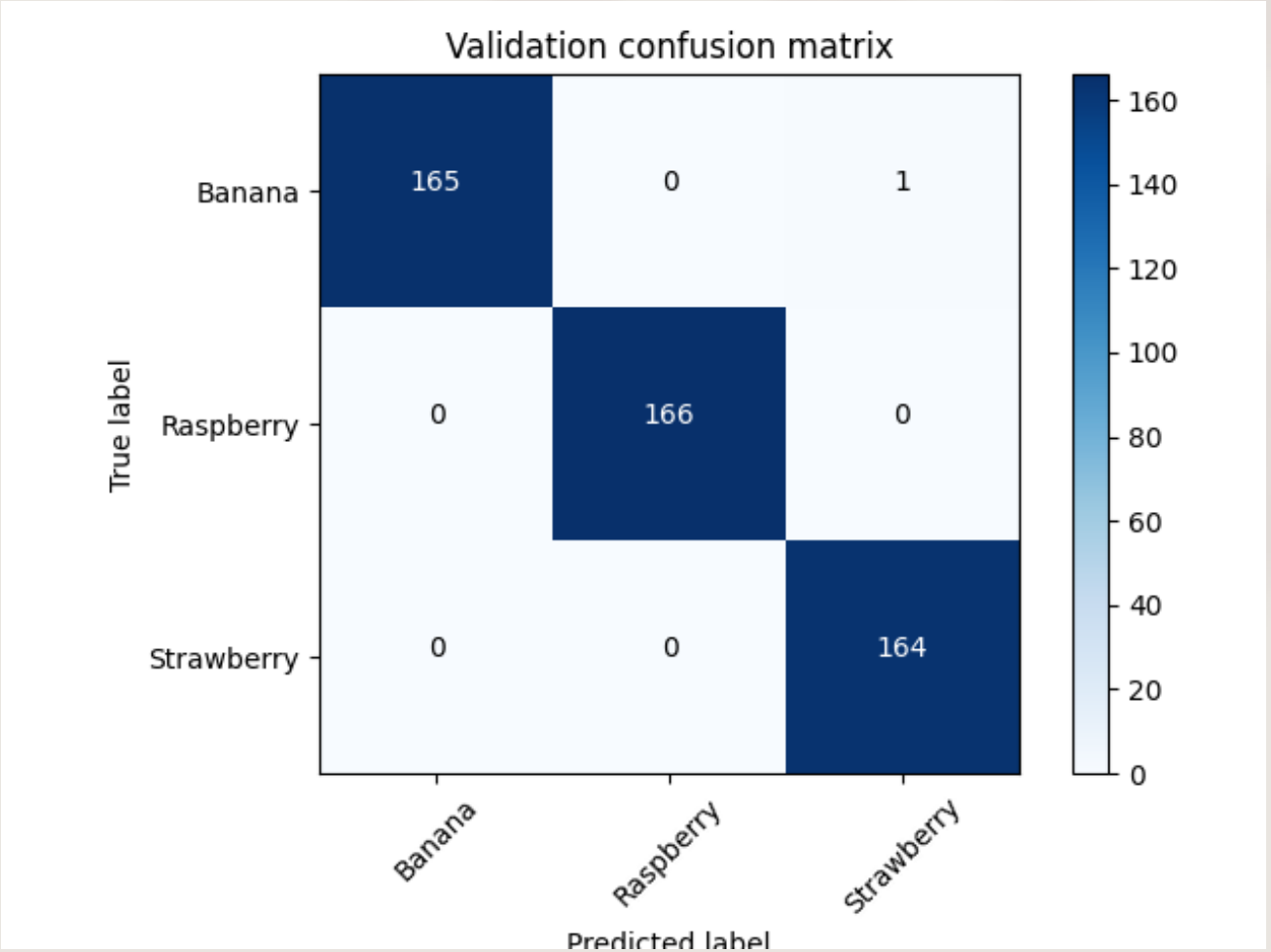


Transfer learning

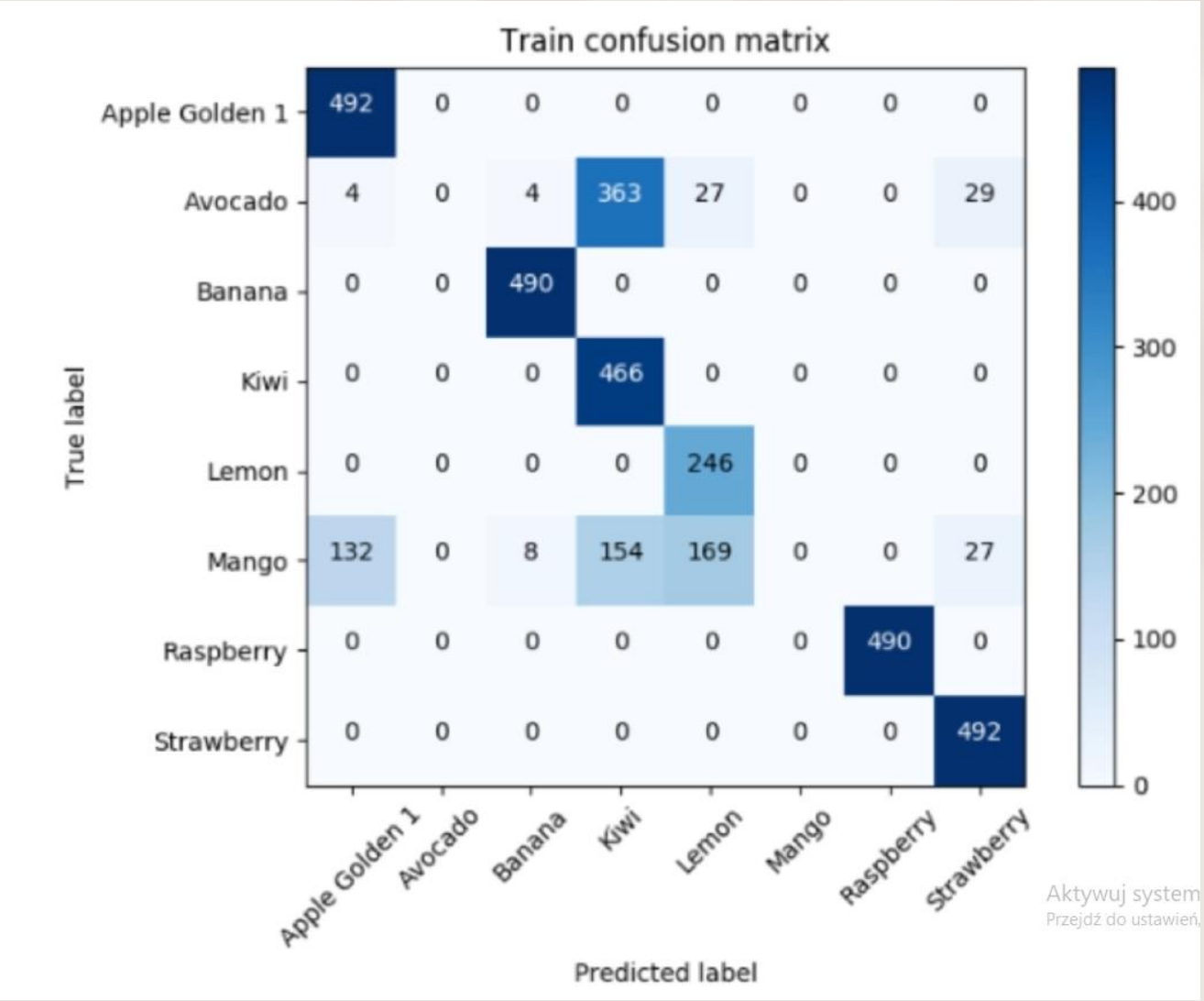
VGG


Train confusion matrix
Banana
Raspberry
Strawberry
True label
490
0
0
0
490
0
0
0
492
Banana
Raspberry
Strawberry
Predicted label
400
300
200
100
0

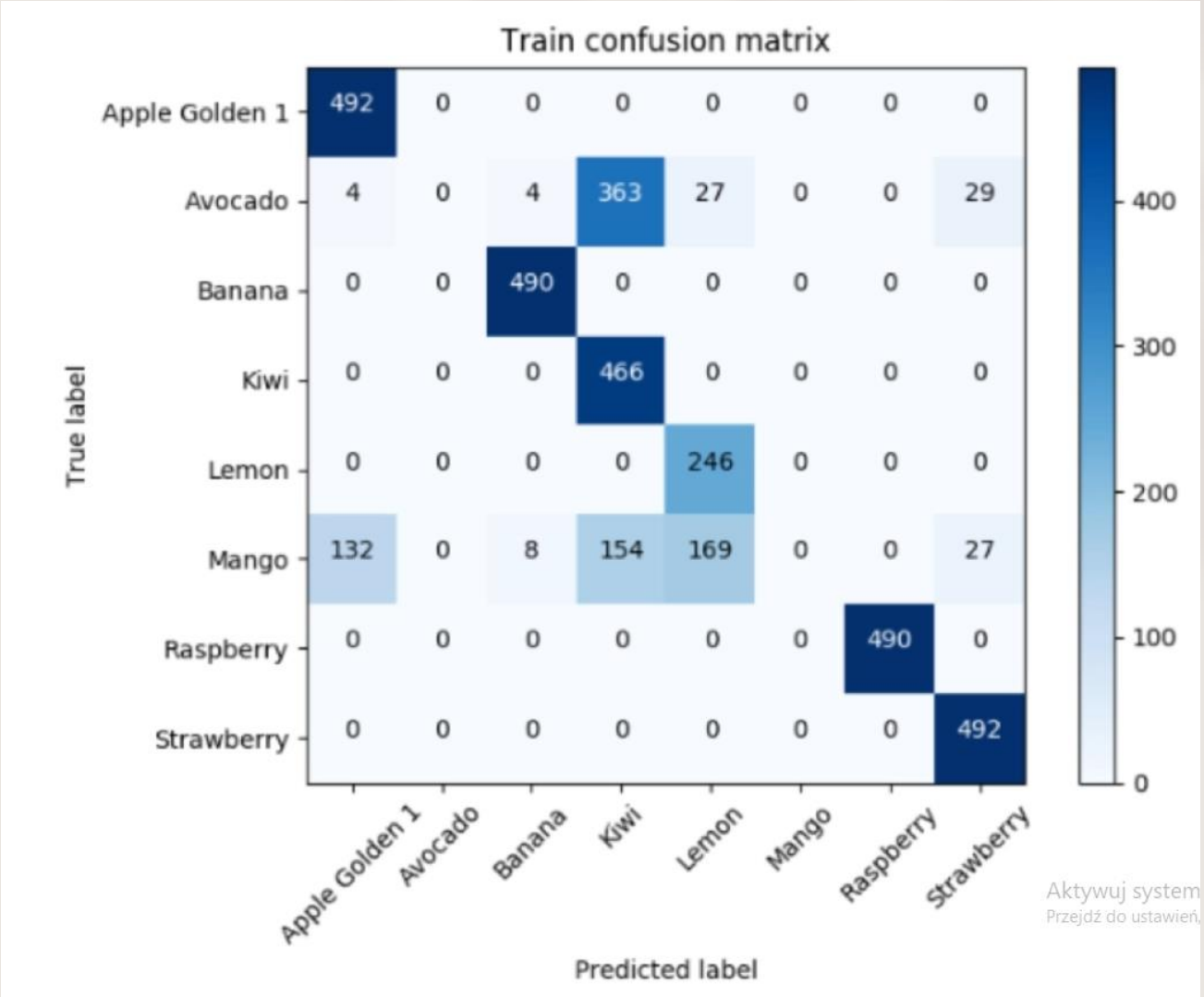
VGG


Validation confusion matrix
Banana
Raspberry
Strawberry
True label
165
0
1
0
166
0
0
0
164
Banana
Raspberry
Strawberry
Predicted label
160
140
120
100
80
60
40
20
0

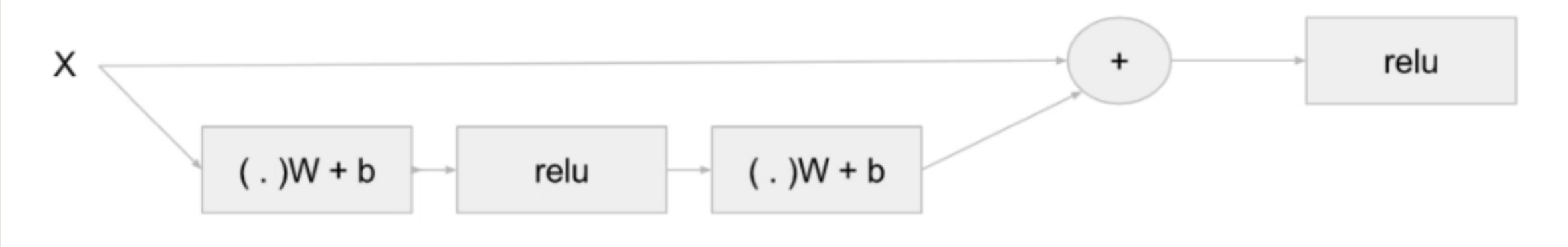
VGG


Train confusion matrix
Apple Golden 1
Avocado
Banana
Kiwi
Lemon
Mango
Raspberry
Strawberry
True label
Predicted label
492 0 0 0 0 0 0 0
4 0 4 363 27 0 0 29
0 0 490 0 0 0 0 0
0 0 0 466 0 0 0 0
0 0 0 0 246 0 0 0
132 0 8 154 169 0 0 27
0 0 0 0 0 0 490 0
0 0 0 0 0 0 0 492
400
300
200
100
0
Aktywuj system
Przejdź do ustawień.

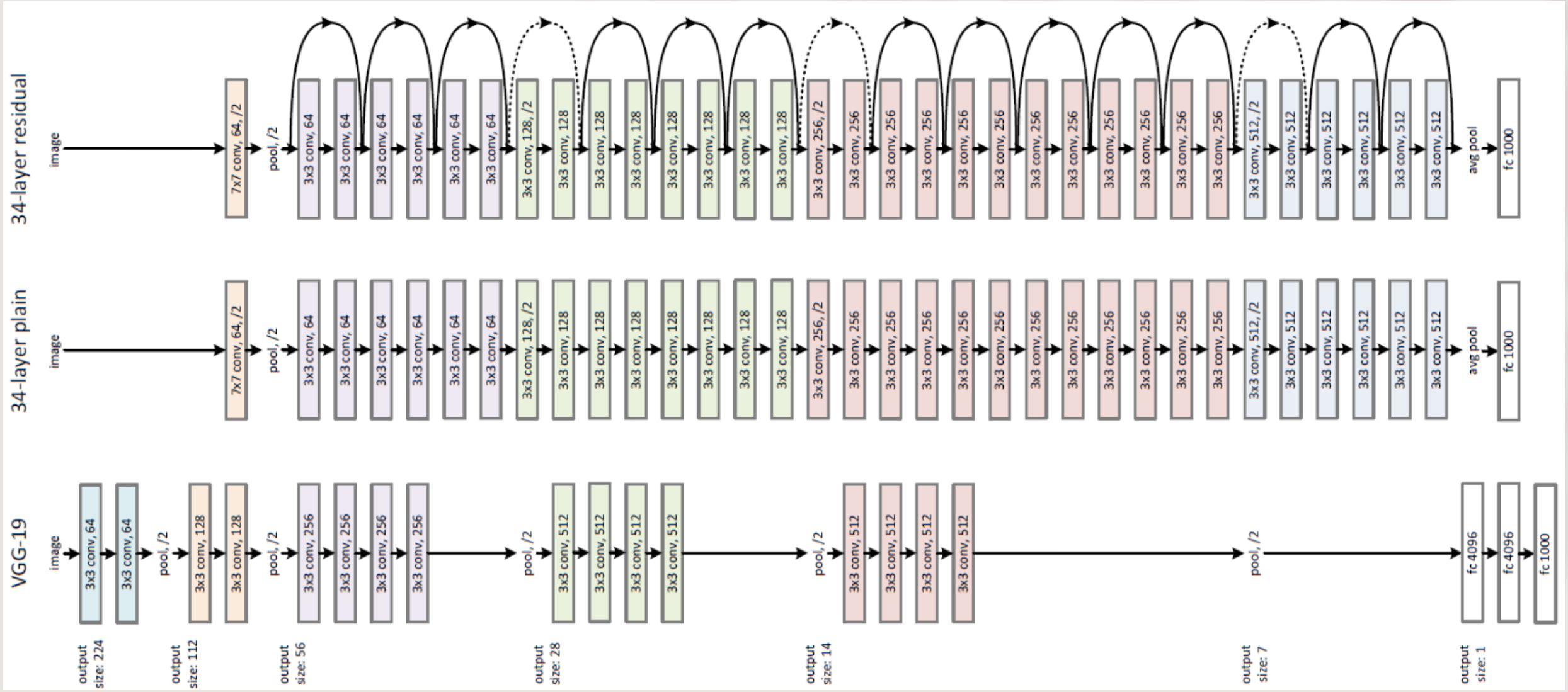
VGG


Train confusion matrix
Apple Golden 1
Avocado
Banana
Kiwi
Lemon
Mango
Raspberry
Strawberry
True label
Predicted label
492 0 0 0 0 0 0 0
4 0 4 363 27 0 0 29
0 0 490 0 0 0 0 0
0 0 0 466 0 0 0 0
0 0 0 0 246 0 0 0
132 0 8 154 169 0 0 27
0 0 0 0 0 0 490 0
0 0 0 0 0 0 0 492
400
300
200
100
0
Aktywuj system
Przejdź do ustawień.

ResNet



ResNet


VGG-19
image
3x3 conv, 64
3x3 conv, 64
pool, /2
3x3 conv, 128
3x3 conv, 128
pool, /2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
pool, /2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
pool, /2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
pool, /2
fc 4096
fc 4096
fc 1000
output size: 224
output size: 112
output size: 56
output size: 28
output size: 14
output size: 7
output size: 1
34-layer plain
image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 128
3x3 conv, 128, /2
3x3 conv, 256
3x3 conv, 256, /2
3x3 conv, 512
3x3 conv, 512, /2
avg pool
fc 1000
34-layer residual
image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 128
3x3 conv, 128, /2
3x3 conv, 256
3x3 conv, 256, /2
3x3 conv, 512
3x3 conv, 512, /2
avg pool
fc 1000