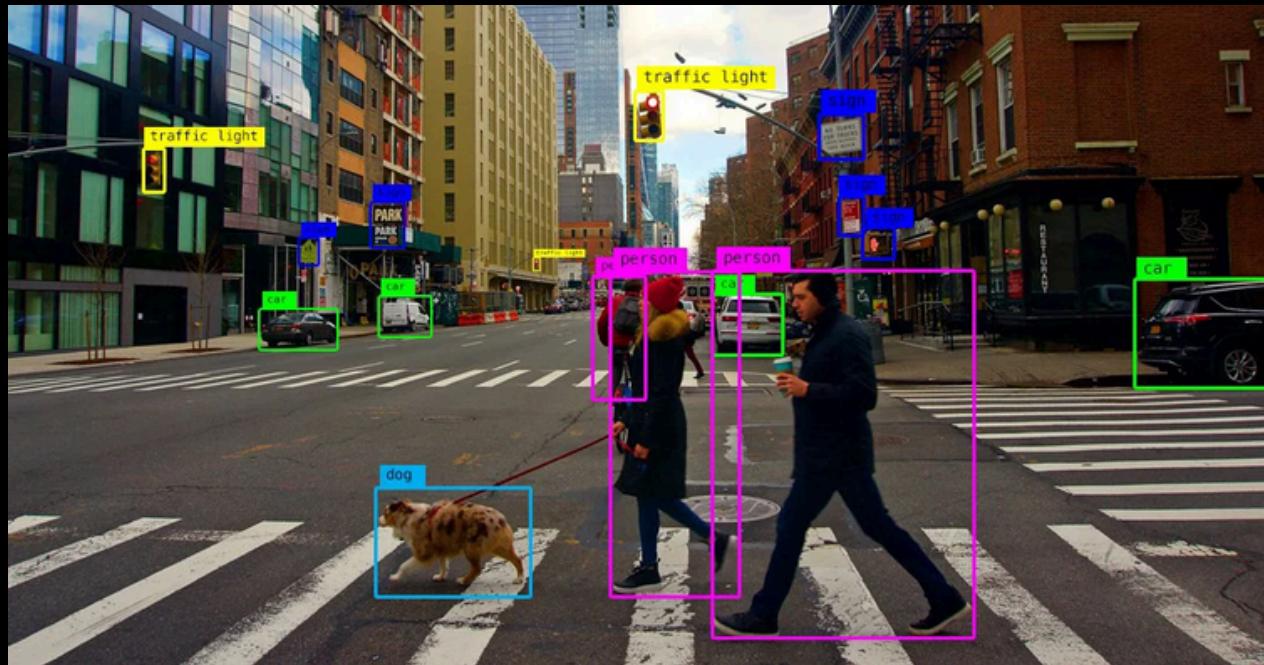




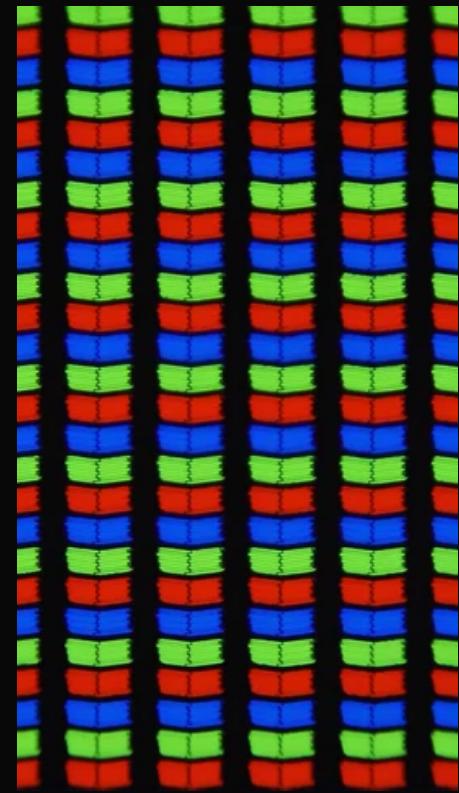
Computer vision



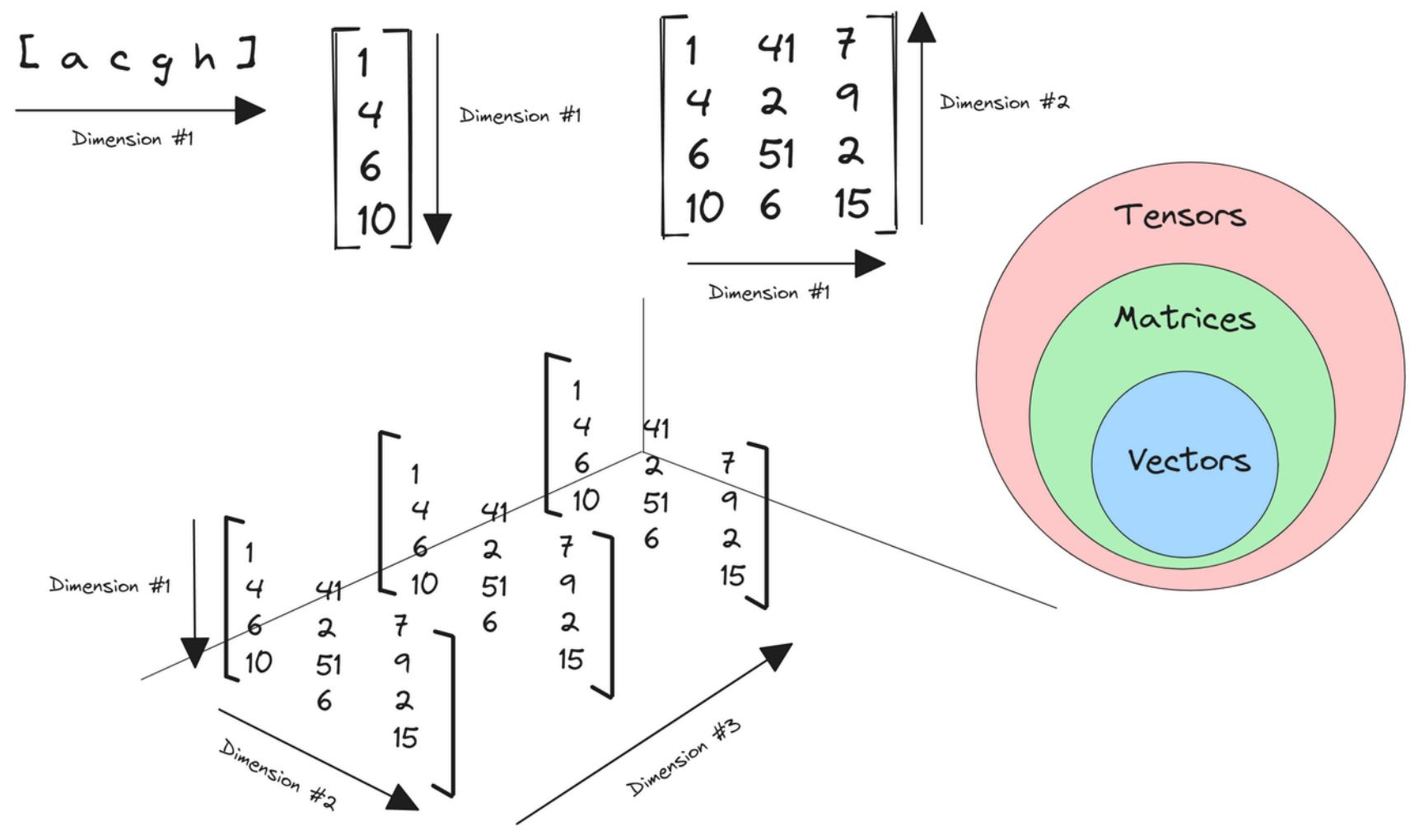
Czym dla komputera jest grafika?



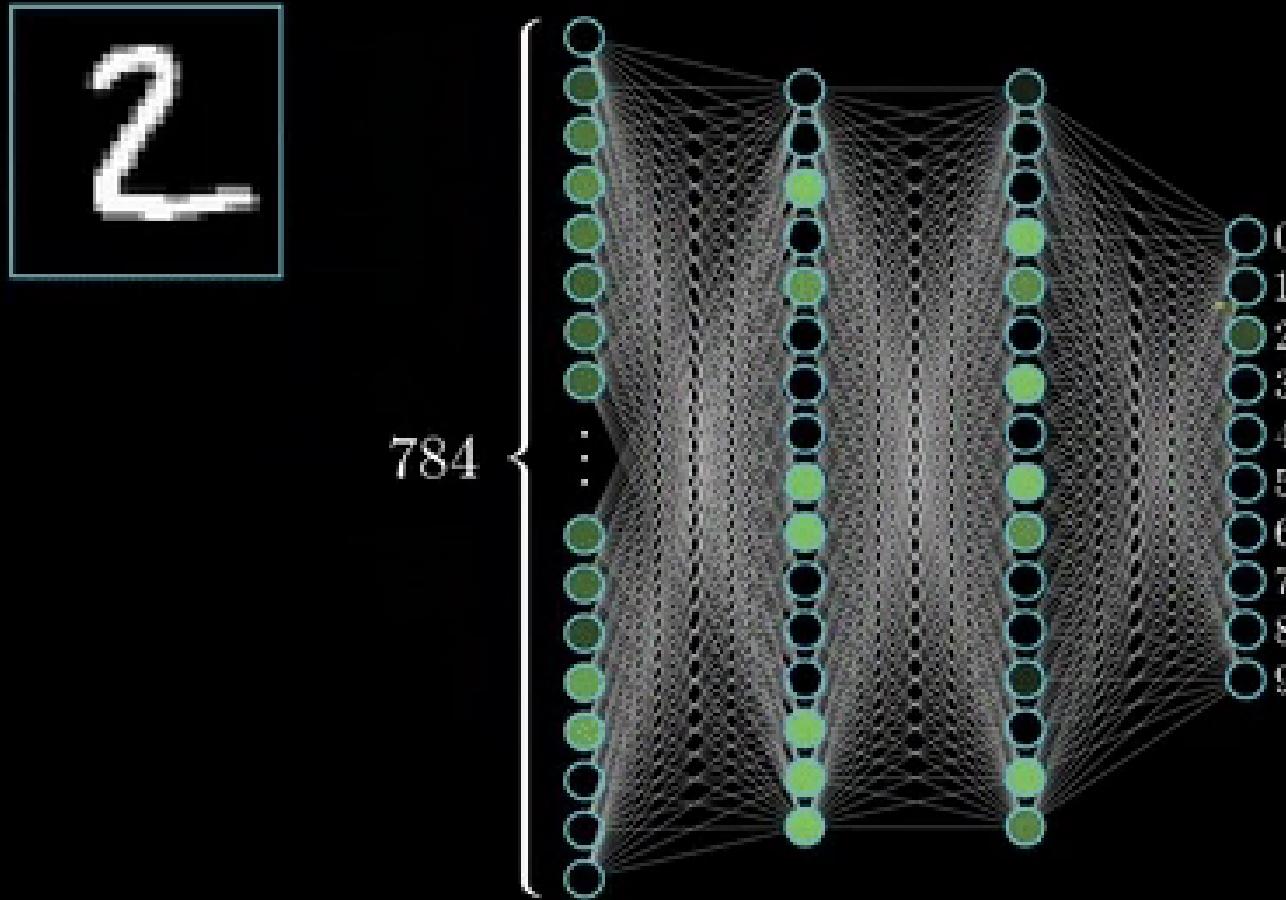
157	153	174	168	150	152	129	151	172	161	155	156
156	182	163	74	75	62	93	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	64	179	209	185	215	211	158	199	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	46	150	79	38	218	241
190	234	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	218
187	196	235	79	1	81	47	0	6	217	255	211
183	292	237	145	0	0	12	108	200	138	243	236
195	296	123	207	177	121	123	200	175	13	96	218



Tensor



MLP



https://miro.medium.com/v2/resize:fit:1000/1*63sGPbvLLpvID16hG1bvmA.gif

Wymiarowość danych

3x32x32

- 480p = $854 * 480 = 409\ 920$
- 720p = $1280 * 720 = 921\ 600$
- 1080p = $1920 * 1080 = 2\ 073\ 600$
- 4k = $3840 * 2160 = 8\ 294\ 400$

(razy 3 jeżeli zdjęcie ma być kolorowe)



Problemy MLP

- nie nadaje się do większych rozdzielczości
- nieodporne na
 - przesunięcia,
 - rotacje
 - skalowanie

nie ma spatial invariance, patrz shuffle_pixels

- “wkuwa” obrazki
- słabo generalizuje

Konwolucja (splot)

Original Image



Sobel Edge Detection



Jak działa?

Input with
padding

0	0	0	0	0	0
0	2	3	1	4	0
0	3	1	3	2	0
0	3	0	1	3	0
0	0	2	0	1	0
0	0	0	0	0	0

conv

kernel		
1	0	-1
1	0	-1
1	0	-1

=

Output of
original size

-4	1	-2	4
-4	3	-5	5
-3	2	-3	4
-2	2	-2	1

Jak działa?

0	0	0	0	0	0	0
0	60	113	56	139	85	0
0	73	121	54	84	128	0
0	131	99	70	129	127	0
0	80	57	115	69	134	0
0	104	126	123	95	130	0
0	0	0	0	0	0	0

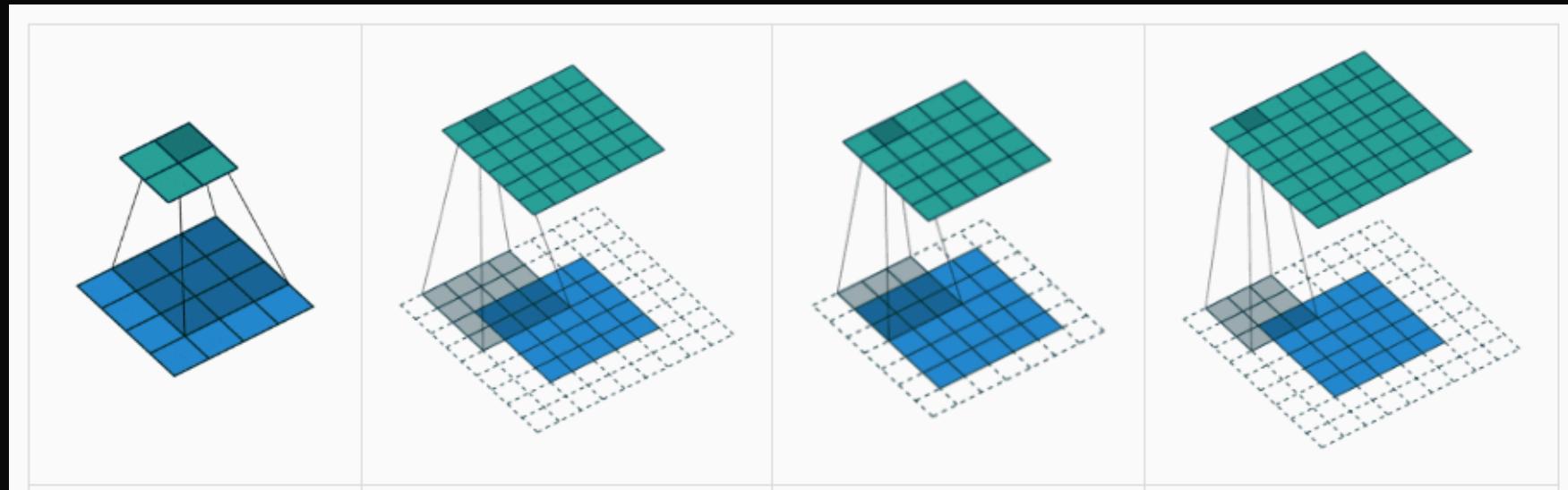
Kernel

0	-1	0
-1	5	-1
0	-1	0

114	328	-26	470	158
53	266	-61	-30	344
403	116	-47	295	244
108	-135	256	-128	344
314	346	279	153	421

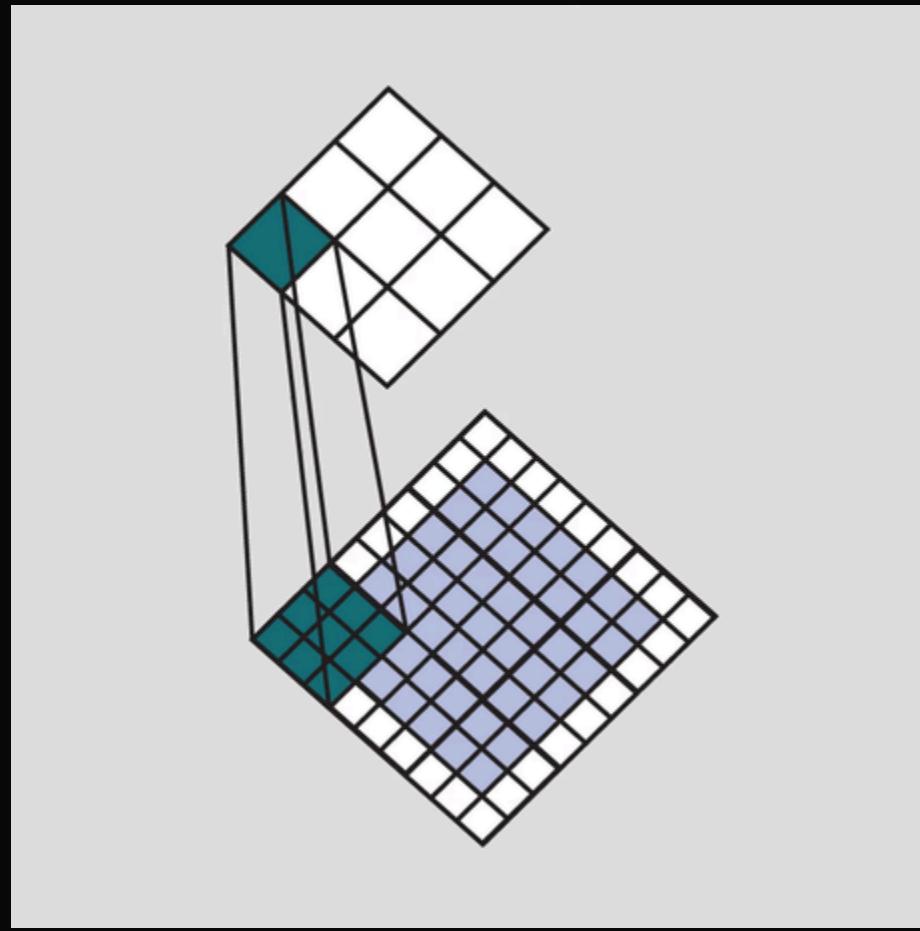
https://svitla.com/uploads/ckeditor/2024/Math%20at%20the%20heart%20of%20CNN/image_930660943761713546482755.gif

Padding



https://miro.medium.com/v2/resize:fit:1400/format:webp/0*jSYji1D43vJWZZ_5

Stride



https://cdn-images-1.medium.com/max/480/1*awHI8cYgPSP4Aj8t929T5g.gif

Konwolucja w sieciach

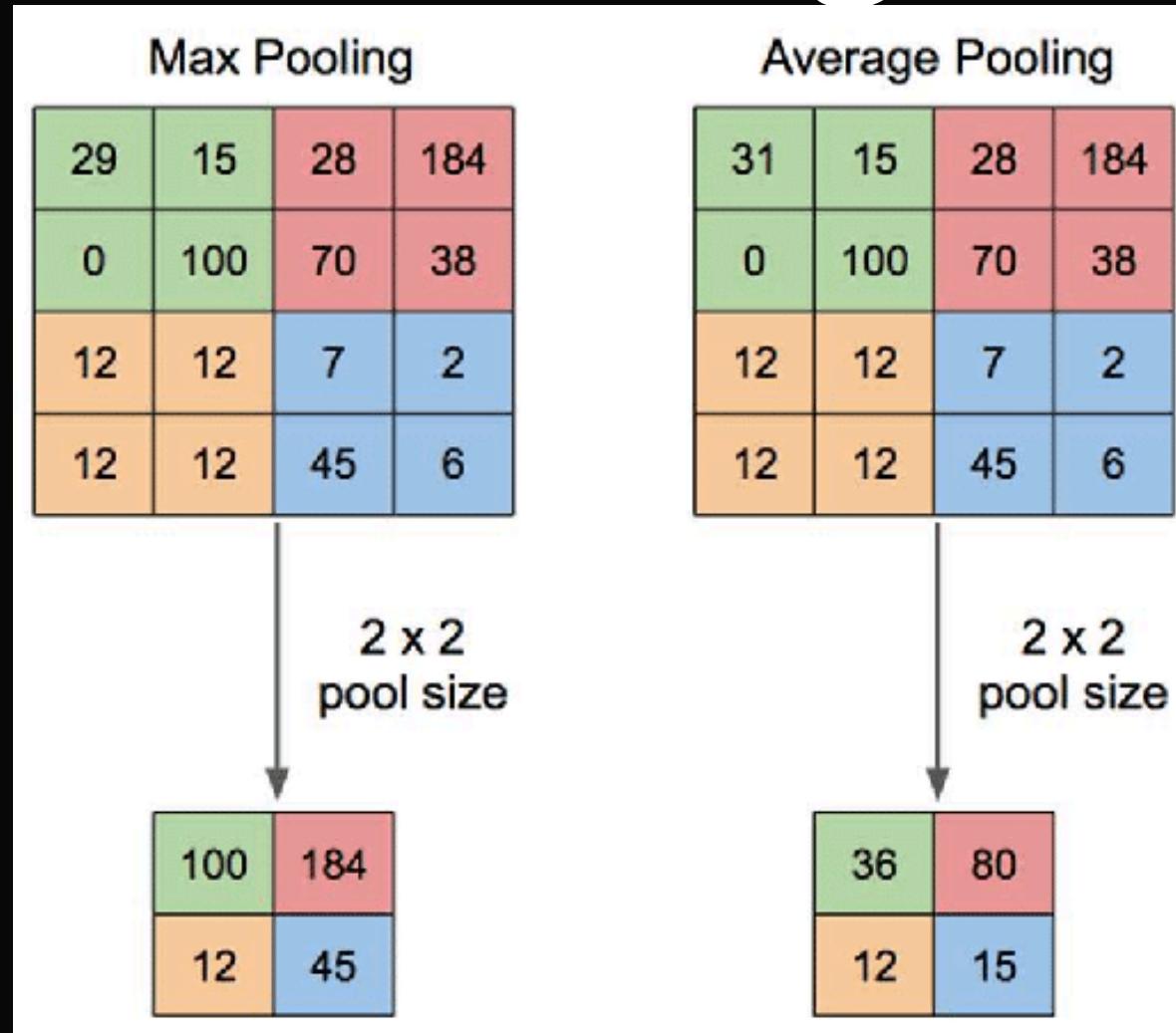
Hiperparametry:

- wymiar jądra
- padding
- stride

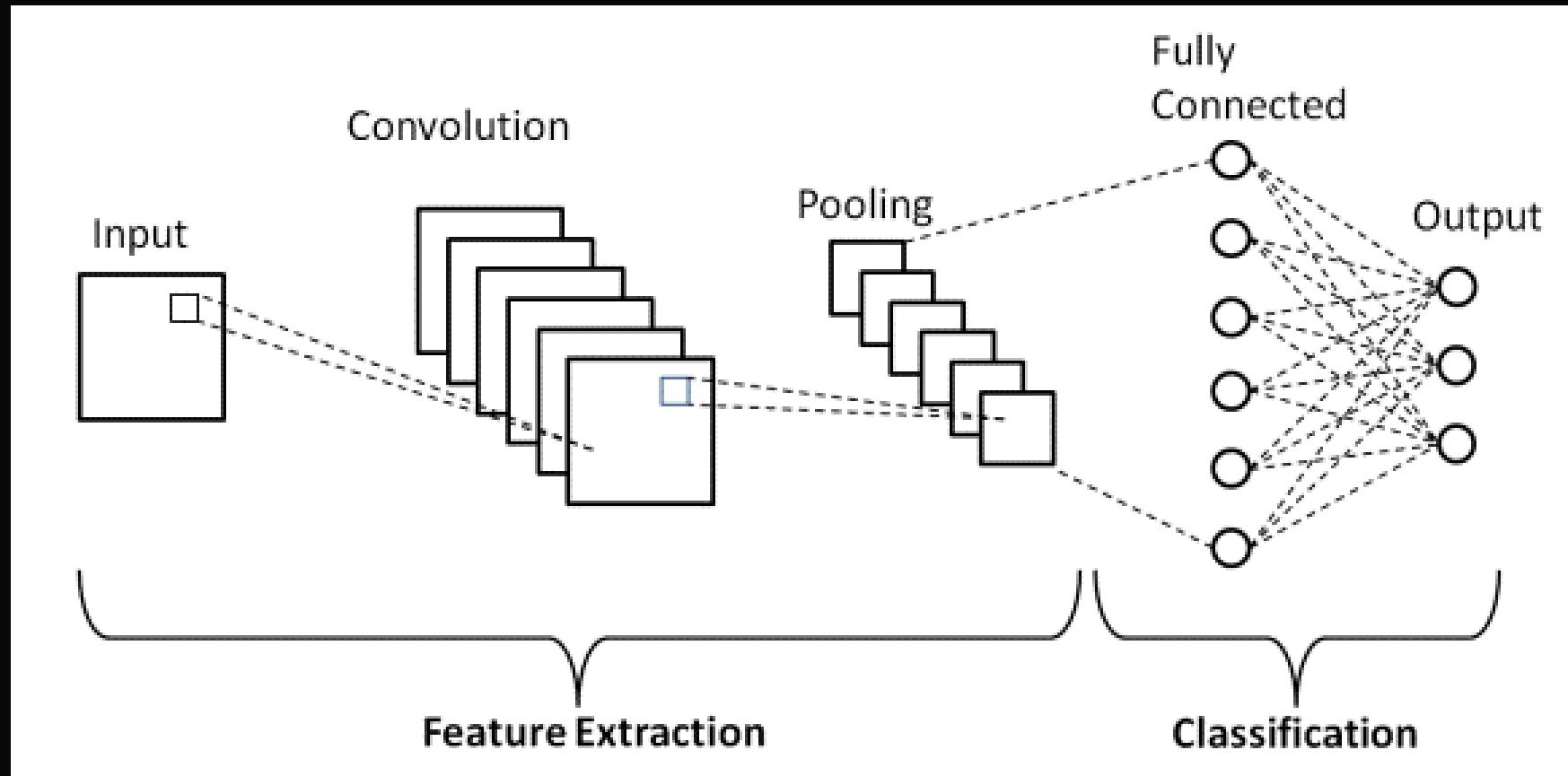
Parametry, których uczy się funkcja:

- wartości konwolucji

Pooling



Architektura CNN

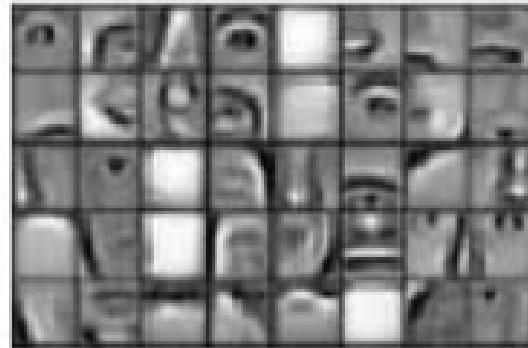
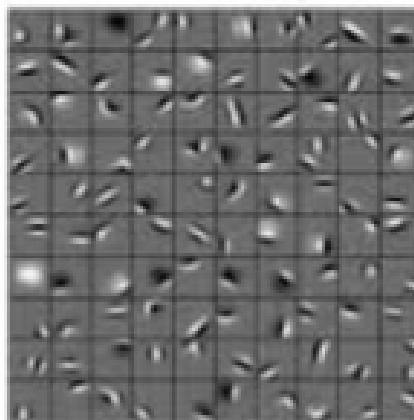


https://adamharley.com/nv_vis/cnn/2d.html

Co się dzieje pod spodem?

Feature Map in Convolutional Neural Networks (CNN)

Low-Level Feature → Overlap → Mid-Level Feature → Overlap → High-Level Feature



Mniej do nauki plusy CNN

Pojedyncze wagi dla jądra wykorzystujemy na całym obrazie!

Spatial invariance plusy CNN

Pooling sprawia, że drobne przesunięcia i rotacje zanikają podczas zmniejszania rozdzielczości zdjęcia.

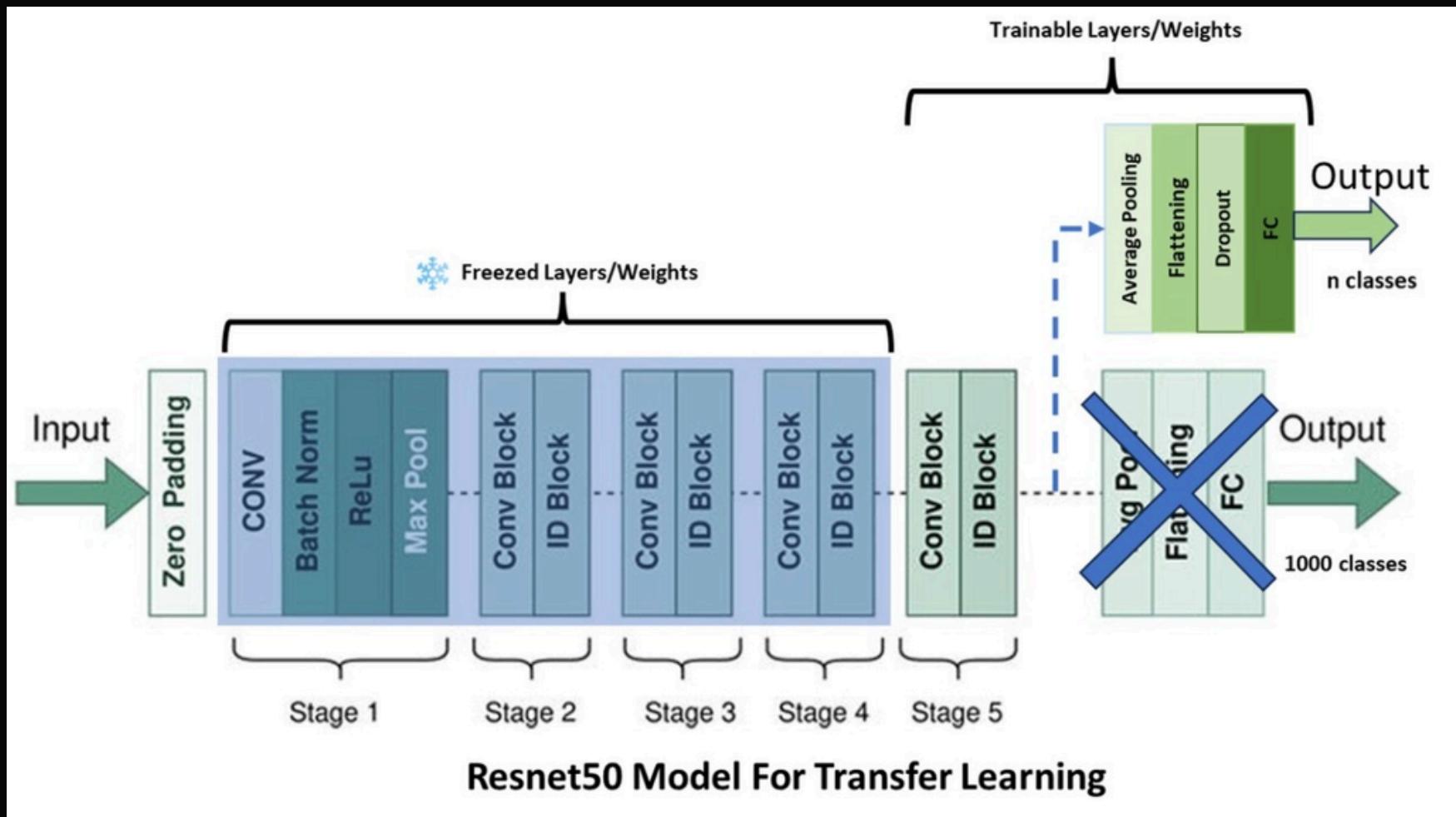
Rozpoznawanie kształtu plusy CNN/Uwagi

CNN opiera się na rozpoznawaniu kształtów.

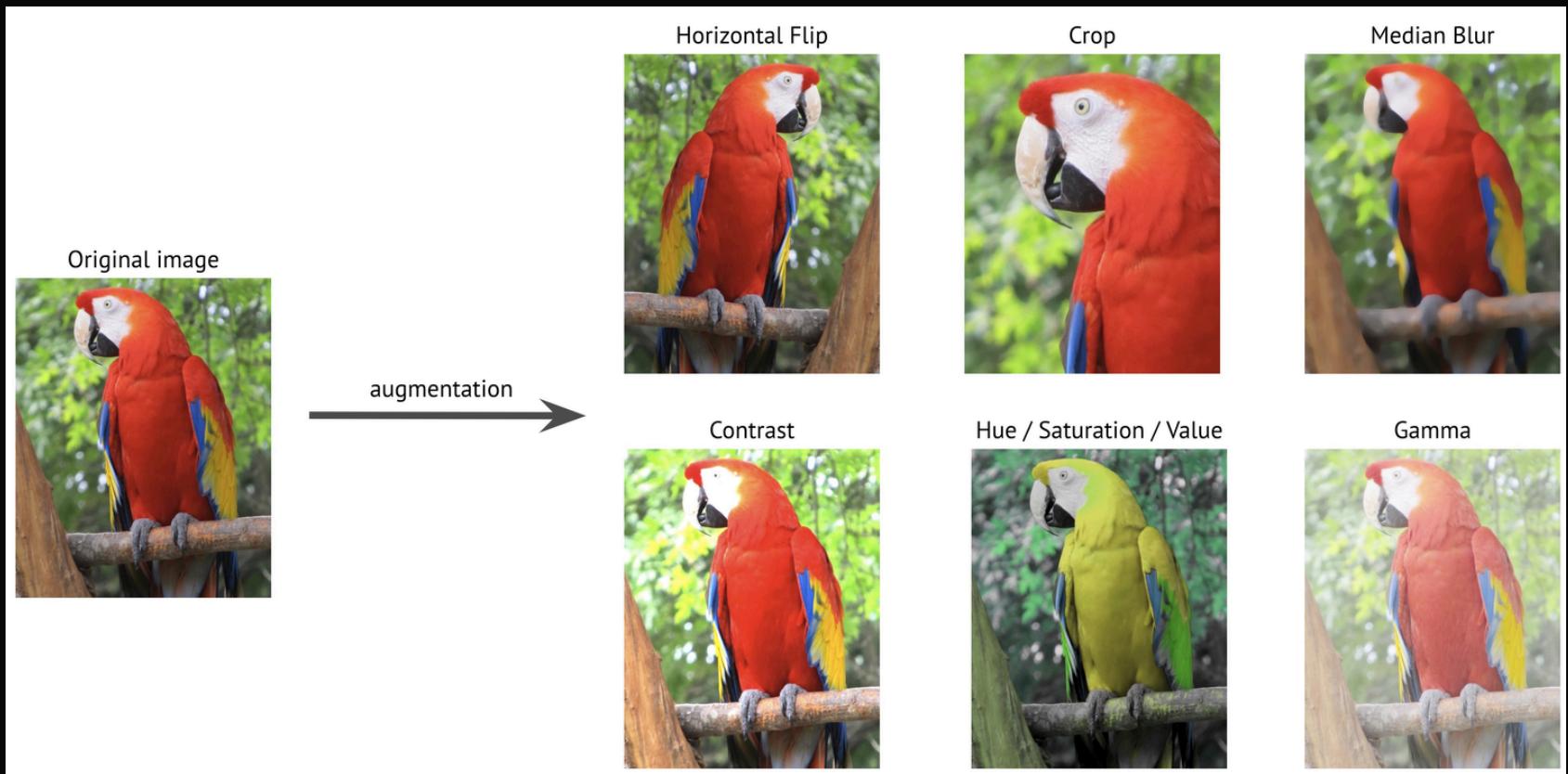
Problem z zanikiem gradientu

Klasyczne CNN nie mogą być zbyt głębokie ponieważ powoduje to, że te najbliższe warstwy nie będą się zbytnio uczyć

Transfer learning!



Augmentacja danych



CIFAR-10

