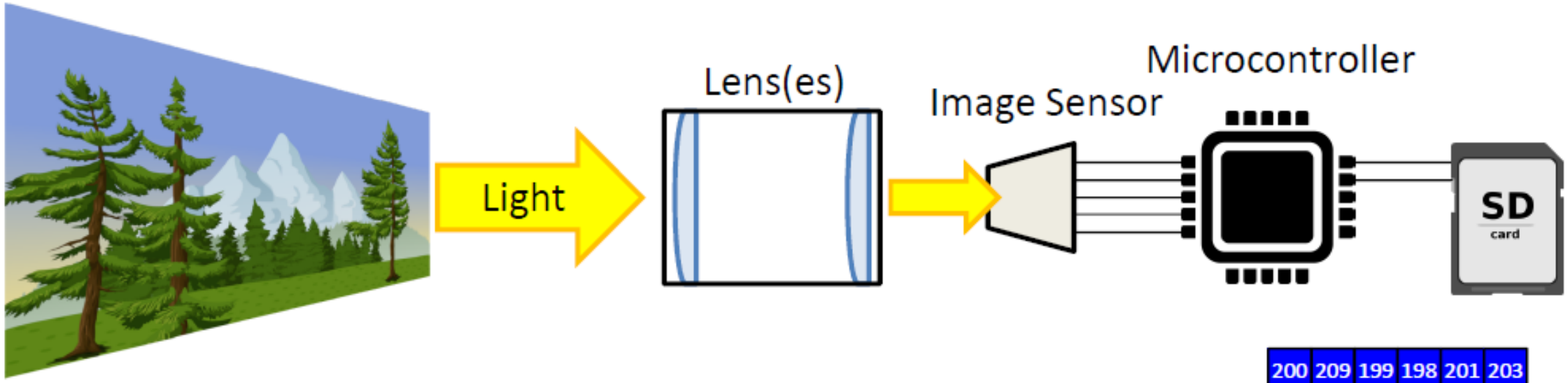
Institución Universitaria
Reacreditada en Alta Calidad

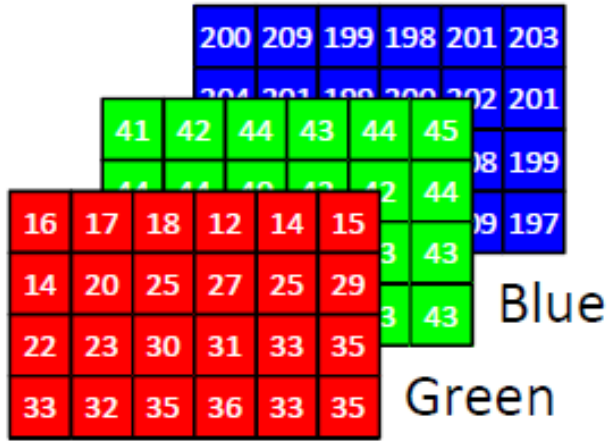
Computer Vision - Convolution

Aprendizaje Automático Embebido

Outline

1. Computer vision – digital images.
2. Image Classification
3. Deep Learning

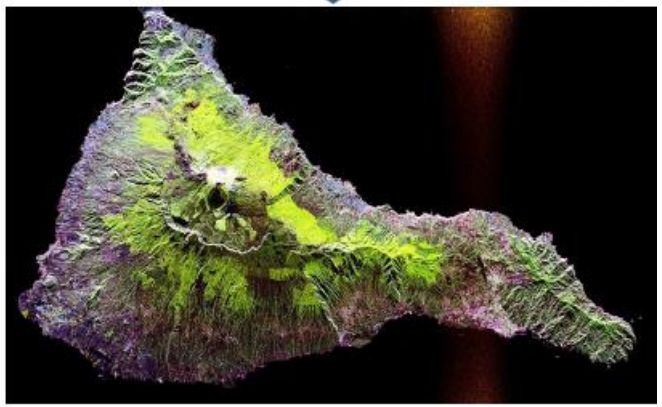




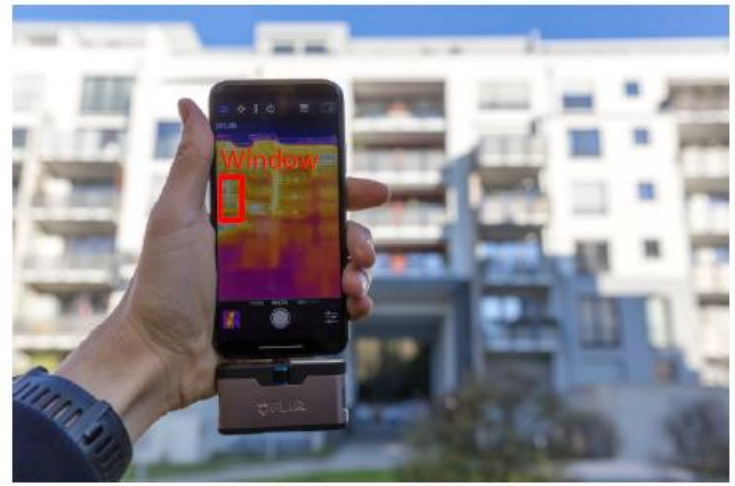
"Thermal image of a building", by Marco Verch (CC BY 2.0)

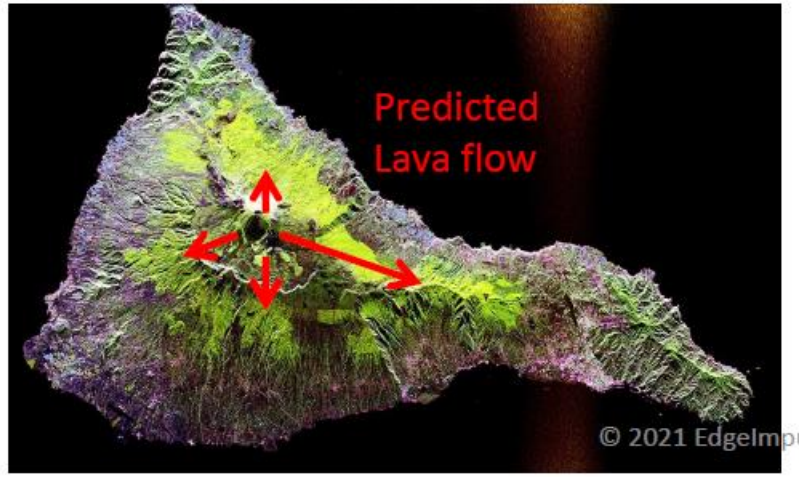
SIR-C/X-SAR imaging radar
on Space Shuttle Endeavour

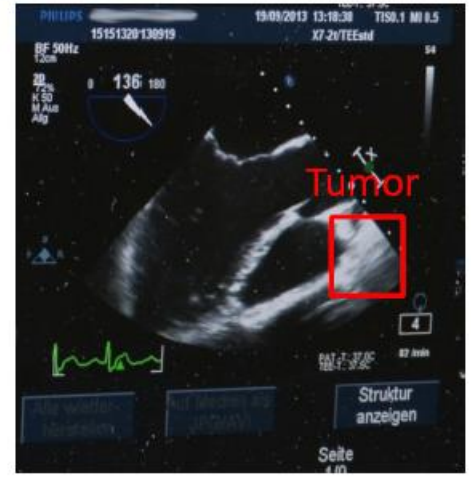
Teide volcano

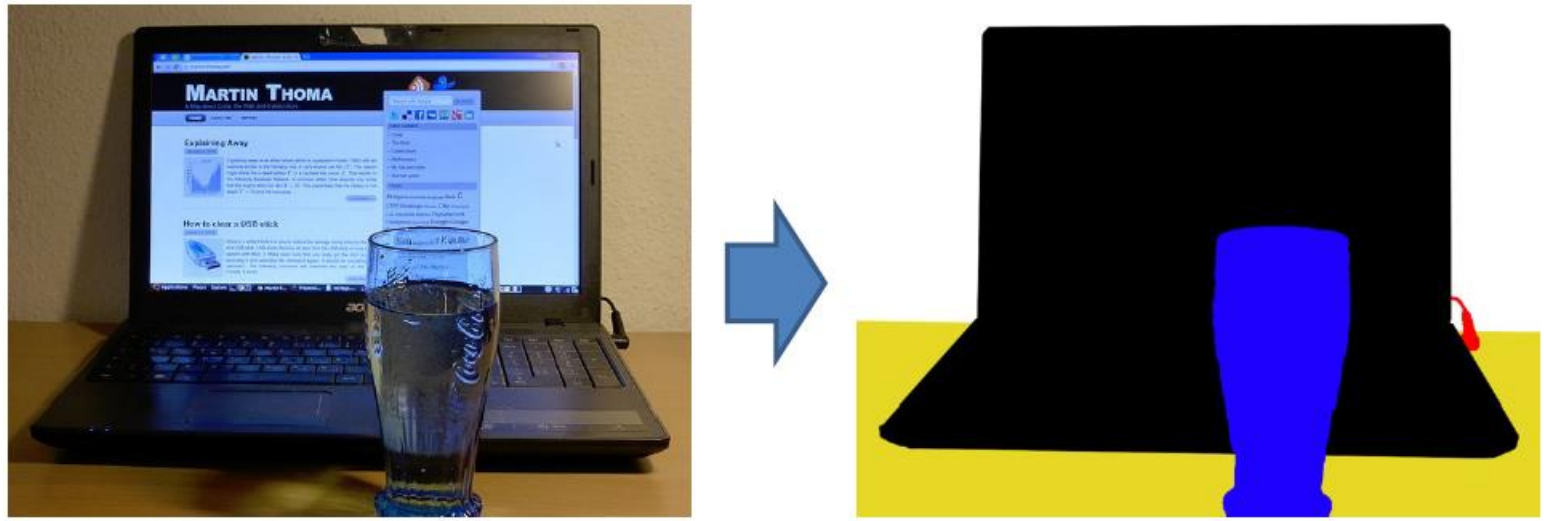


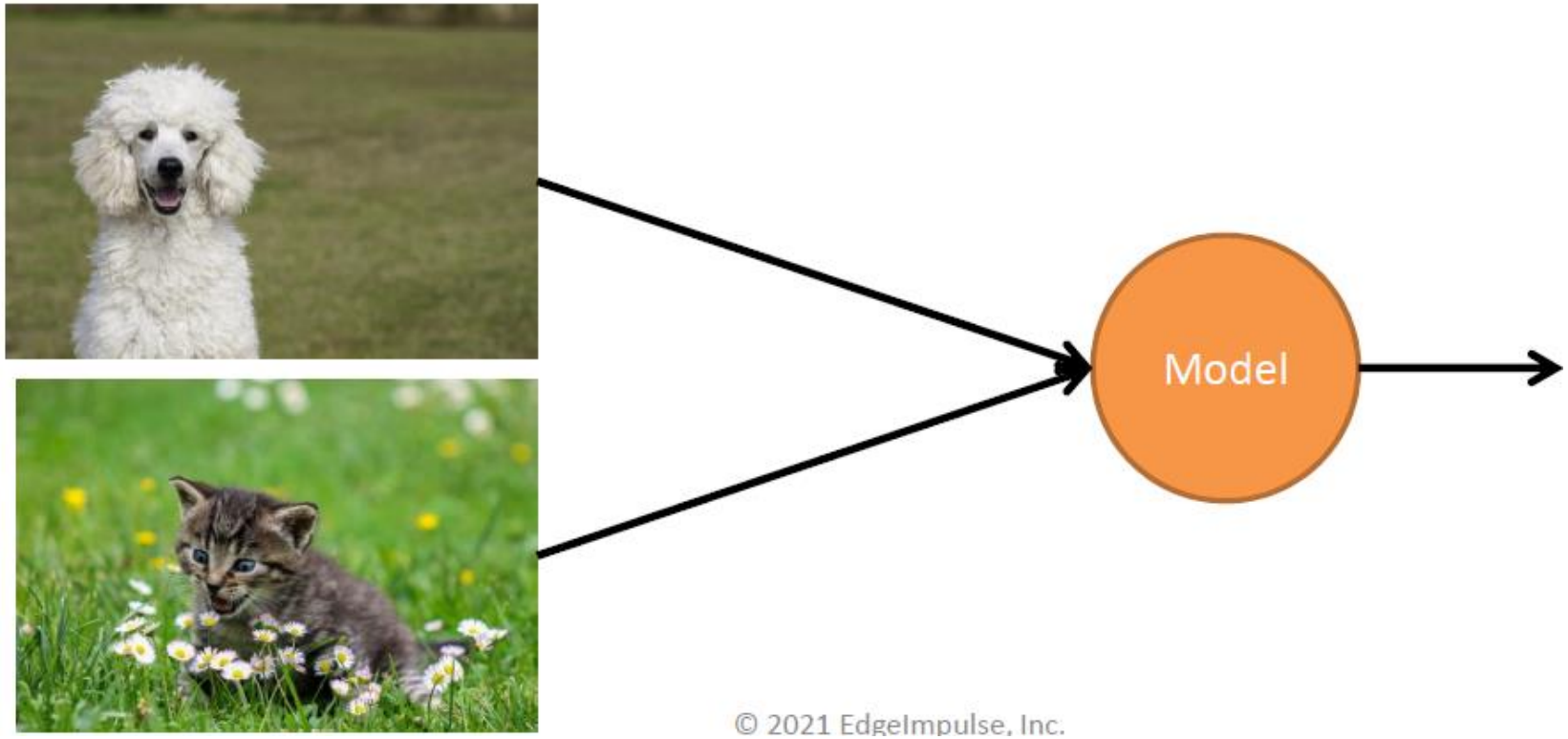
npulse, Inc.

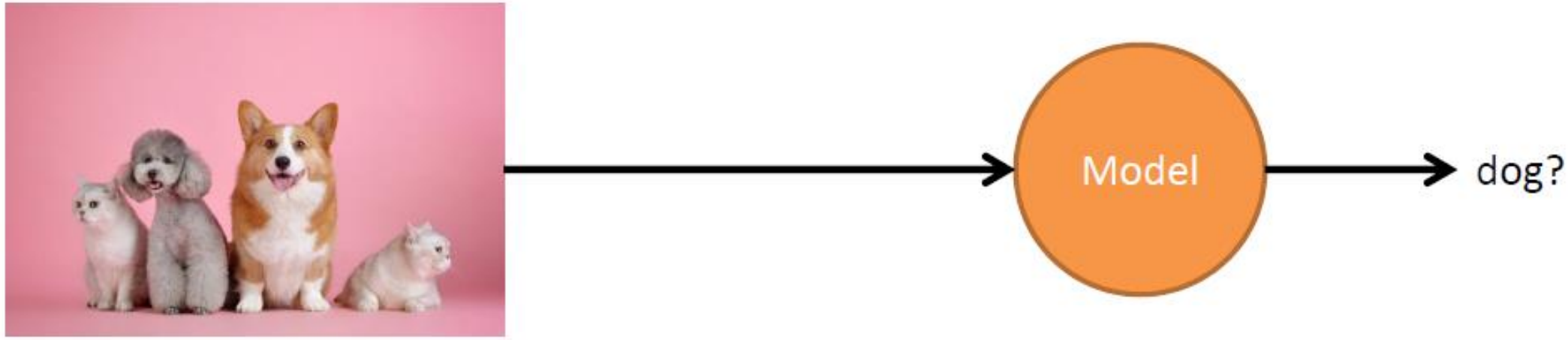


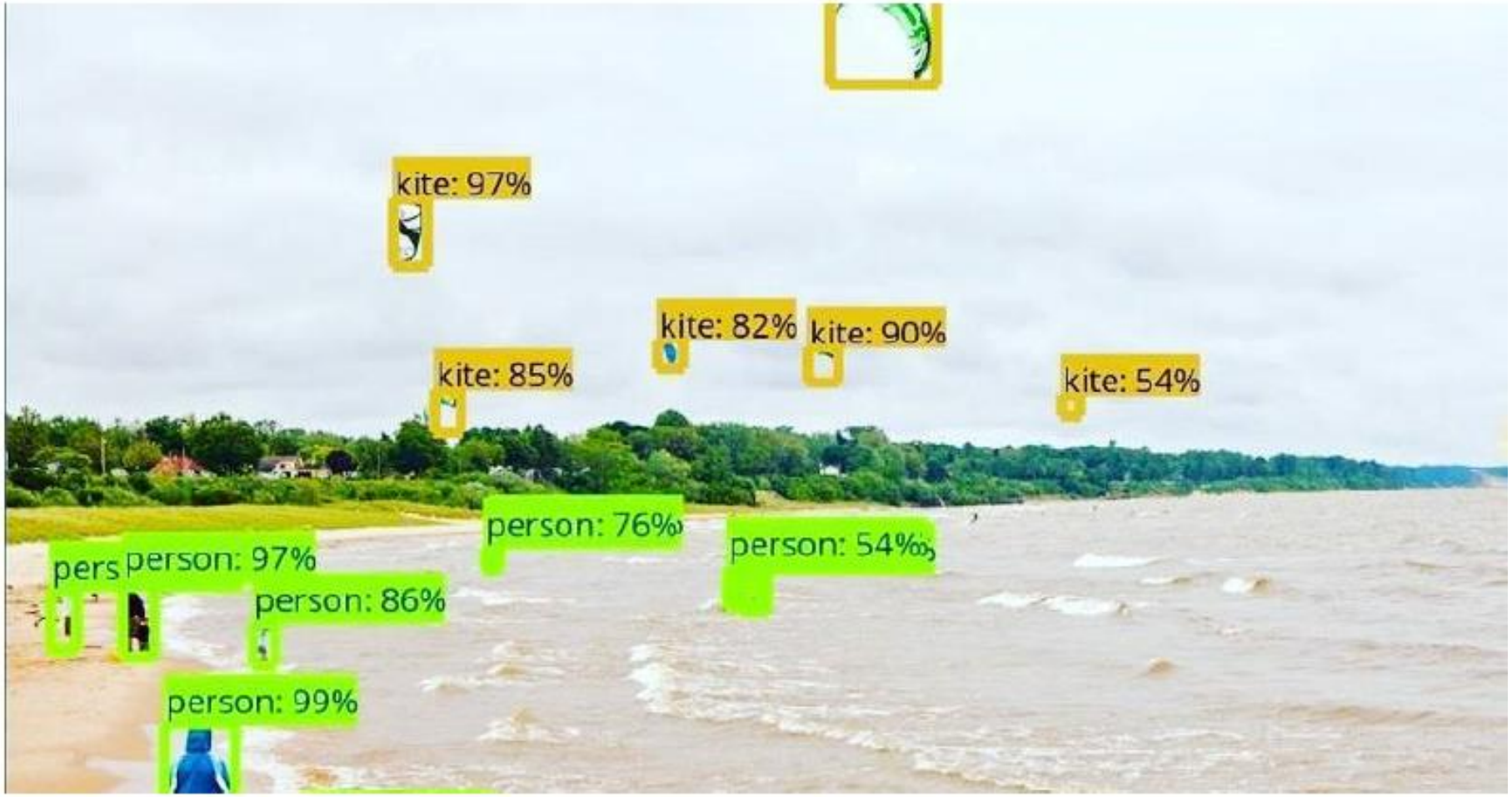




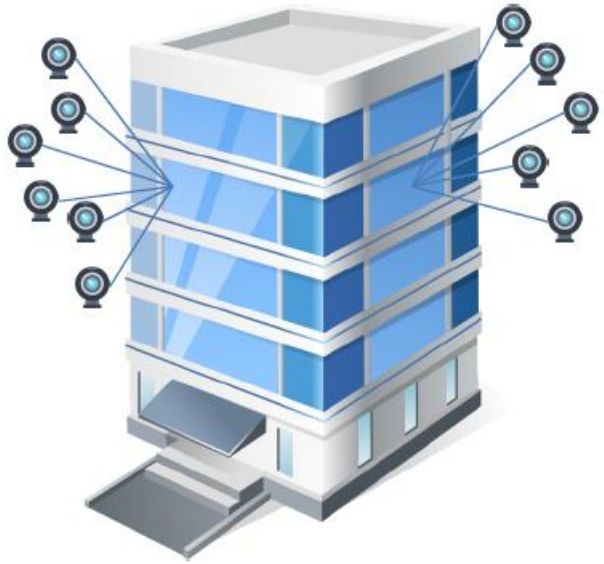








"The future of computer vision," by ShashiBellamkonda (CC BY 2.0)



240 x 240 x 8 x 30 x 1 = 13.824 Mbps

Number of pixels / Num. cameras

Bits per pixel Frames per second

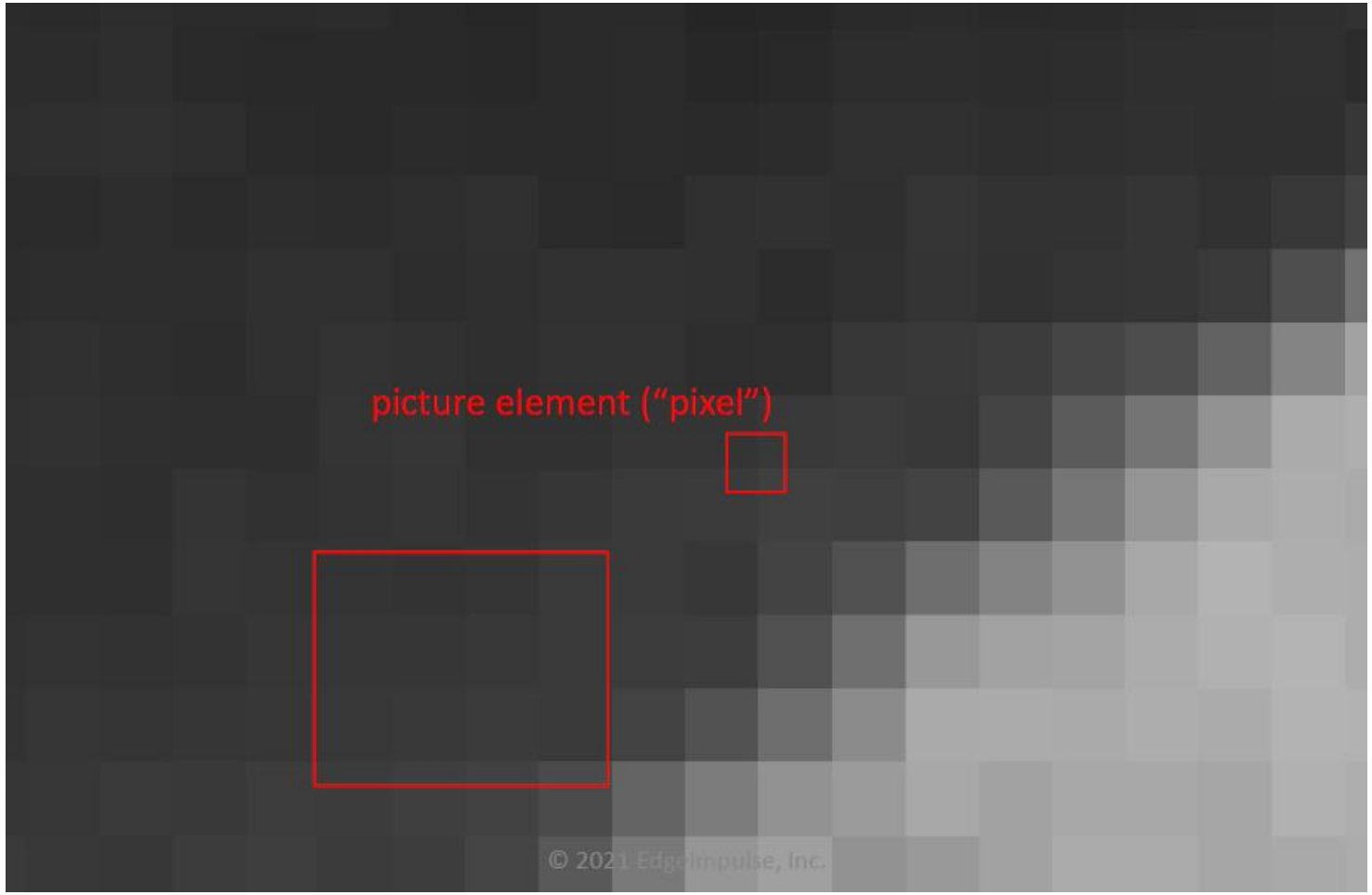
"OpenMV H7 Camera", by SparkFun Electronics (CC BY 2.0)

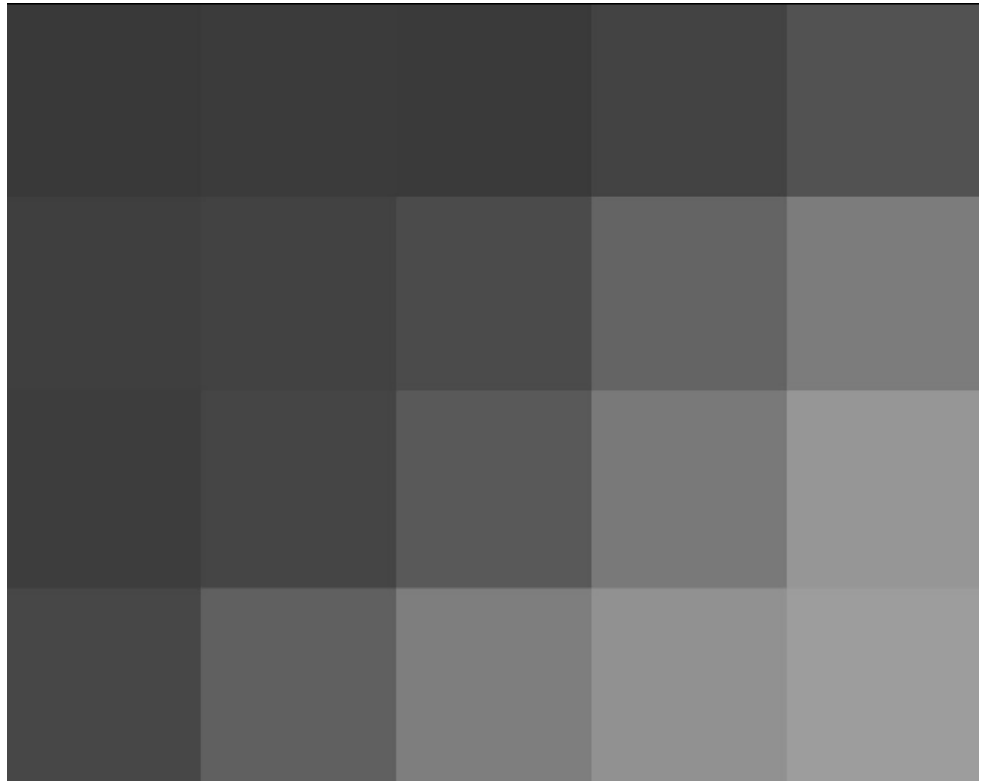
1 x 30 x 1 = 30 bps

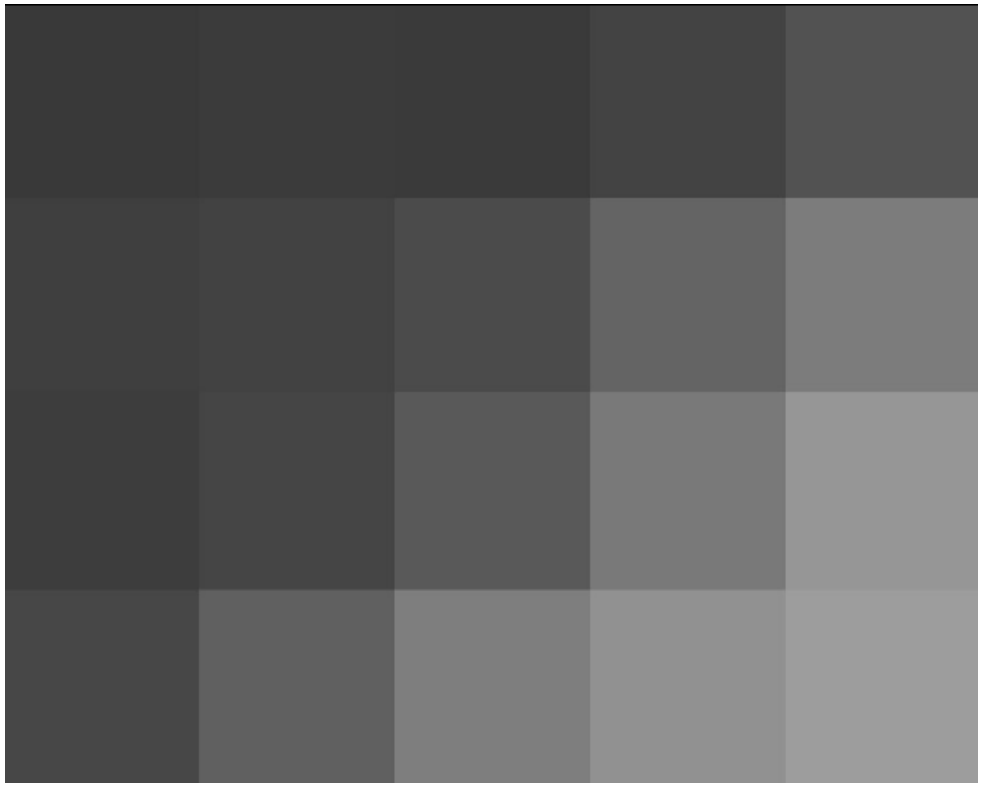
Num. cameras

Person or not person Frames per second

"Hands-free Driving" by Steve Jurvetson (CC BY 2.0)


picture element (“pixel”)
© 2021 EdgeImpulse, Inc.



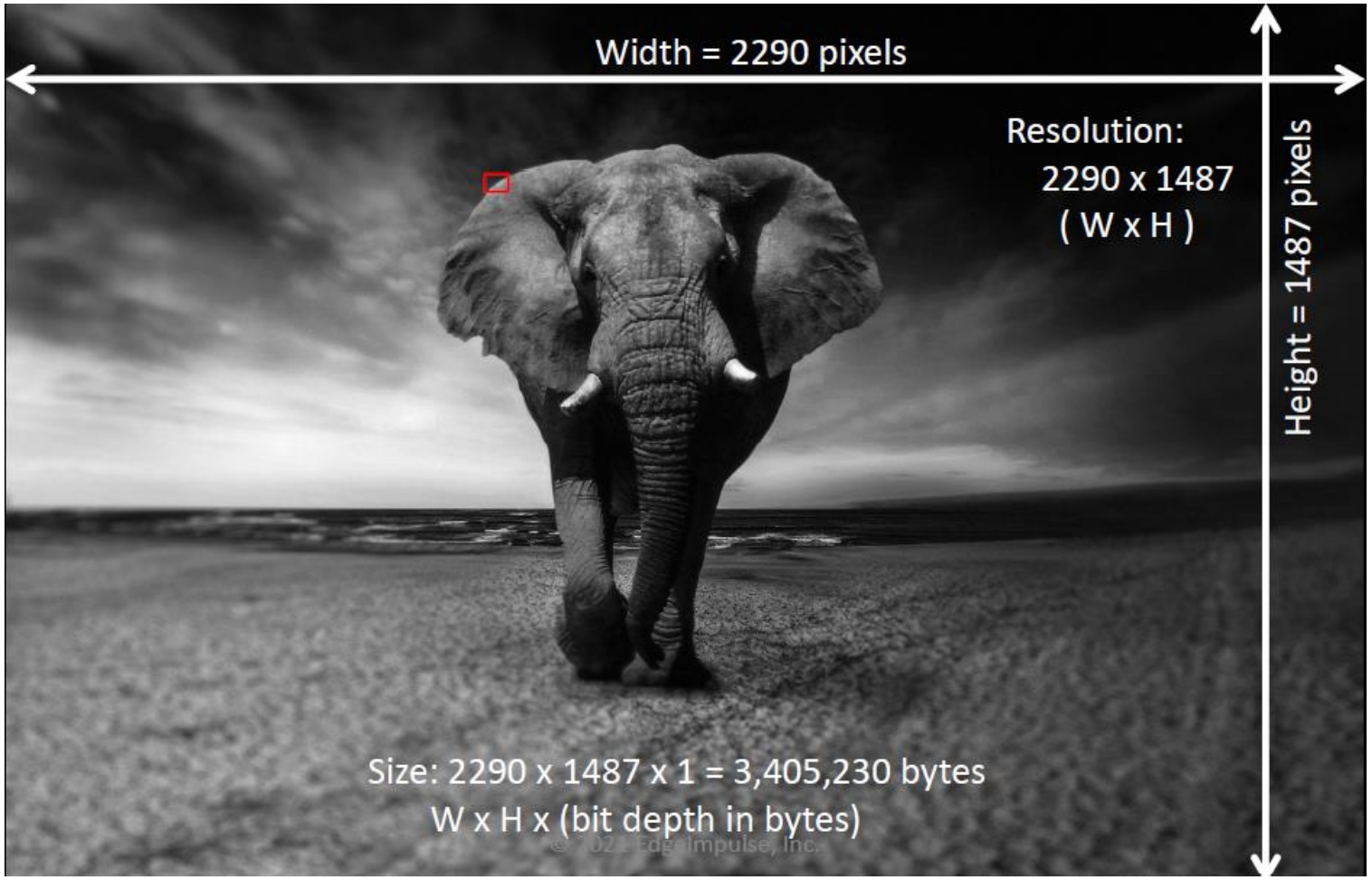


0.22	0.23	0.23	0.26	0.32
0.25	0.26	0.29	0.39	0.49
0.24	0.27	0.35	0.47	0.59
0.28	0.38	0.49	0.57	0.62

0.22	0.23	0.23	0.26	0.32
0.25	0.26	0.29	0.39	0.49
0.24	0.27	0.35	0.47	0.59
0.28	0.38	0.49	0.57	0.62

57	59	58	67	82
63	66	75	100	124
61	69	89	121	150
71	96	126	145	157

Bit depth: 8 bits
- 0 = black
- 255 = white




© 2021 EdgeImpulse, Inc.


© 2021 EdgeImpulse, Inc.

R: 16 G: 159 B: 165	R: 34 G: 179 B: 176	R: 30 G: 161 B: 147	R: 55 G: 163 B: 131	R: 131 G: 204 B: 148
R: 19 G: 160 B: 152	R: 34 G: 166 B: 145	R: 55 G: 161 B: 125	R: 119 G: 187 B: 136	R: 184 G: 200 B: 135
R: 44 G: 166 B: 143	R: 73 G: 173 B: 135	R: 140 G: 204 B: 152	R: 186 G: 208 B: 144	R: 208 G: 181 B: 112
R: 101 G: 189 B: 149	R: 162 G: 215 B: 159	R: 203 G: 212 B: 145	R: 216 G: 190 B: 116	R: 208 G: 151 B: 80

Bit depth: 8x3=24 bits

- 1 byte: Red
- 1 byte: Green
- 1 byte: Blue

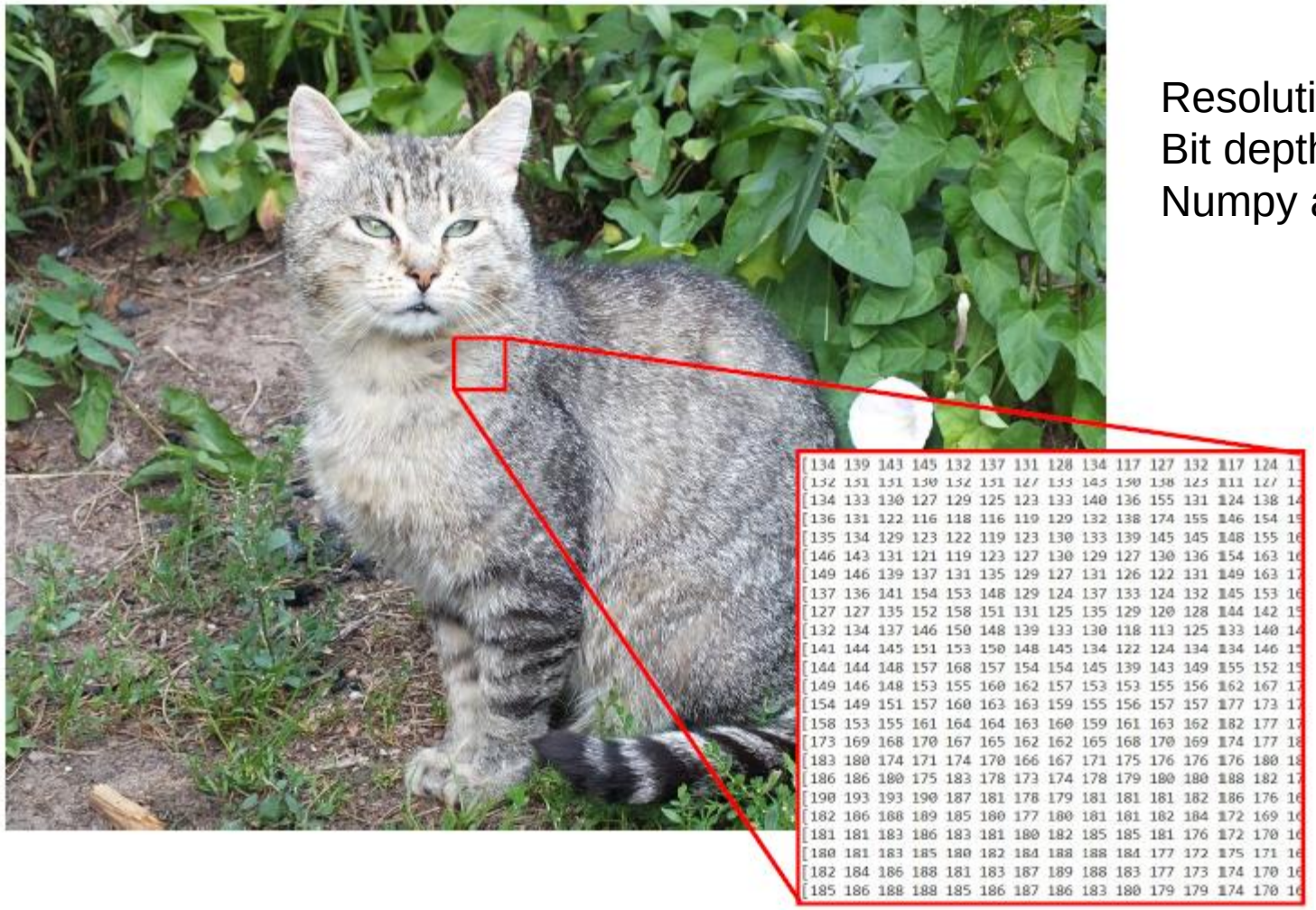
Optional

- 1 byte: Alpha

Outline

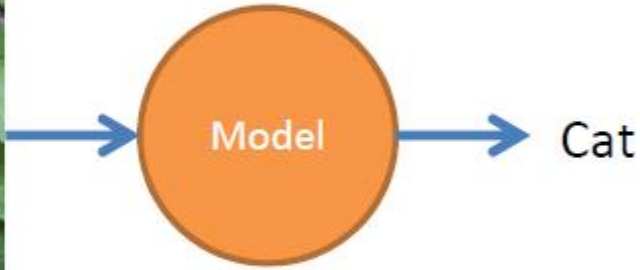
1. Computer vision – digital images.
2. Image Classification
3. Deep Learning

Image Classification



Resolution: 2048 x 1536 pixels
Bit depth: 24 bits
Numpy array shape: (1536, 2048, 3)

Image Classification



Binary Classification

One-vs-One

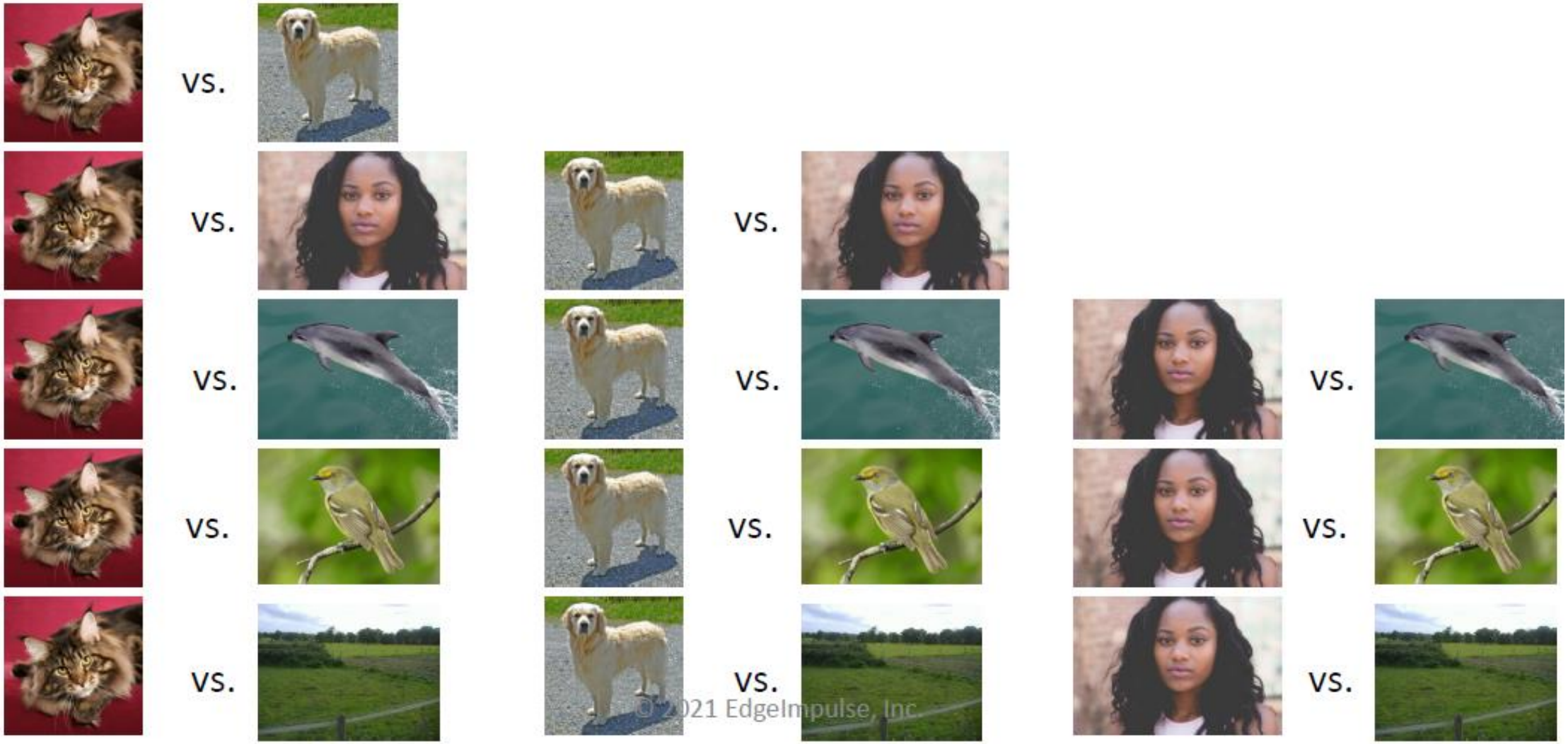
vs.

One-vs-Rest

vs.

Multiclass Classification

Multiclass: One-vs-One

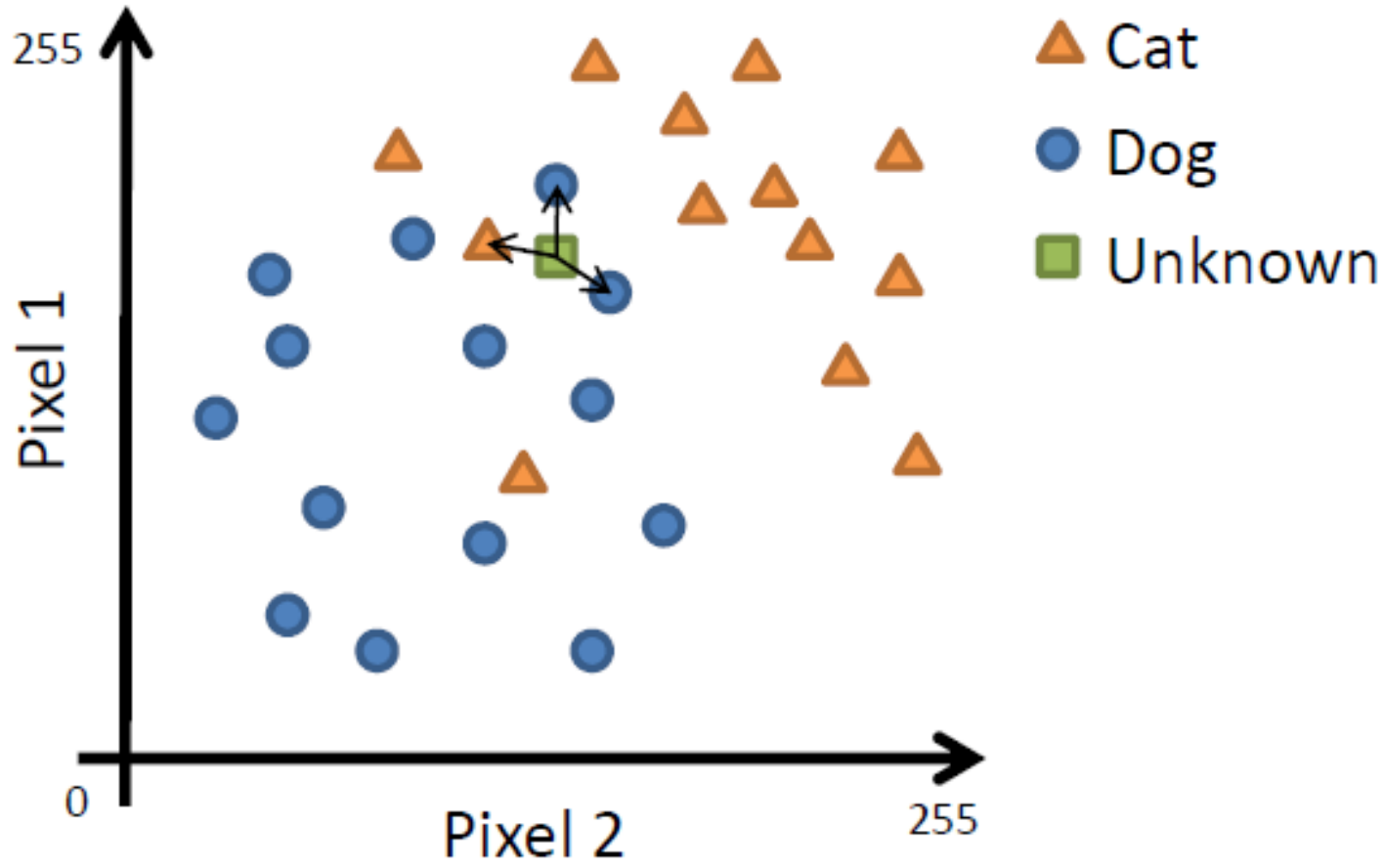


Multiclass: One-vs-Rest

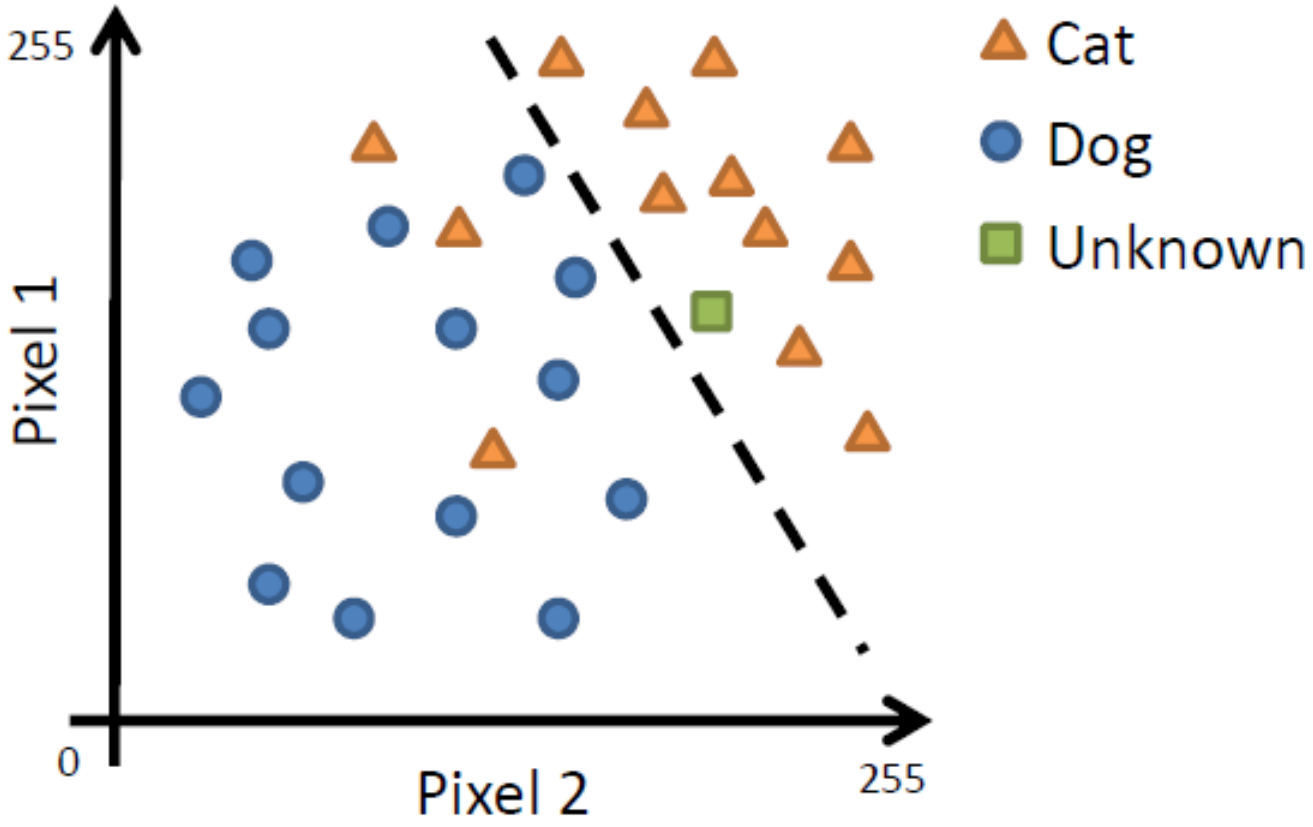
vs.

vs.

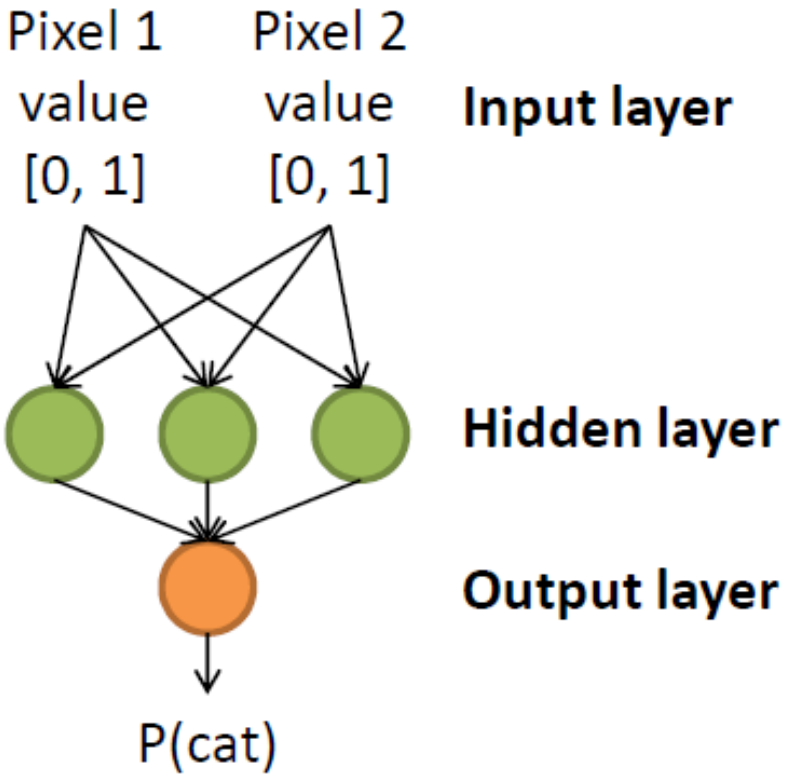
k-Nearest Neighbors (k-NN)



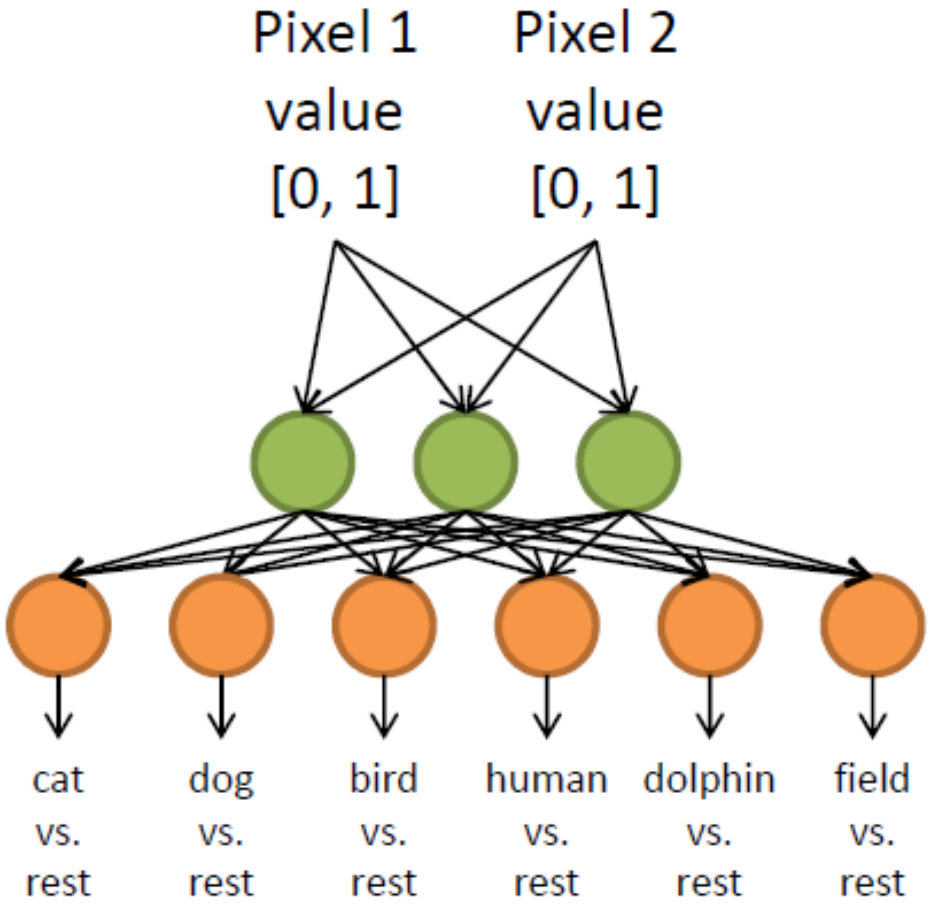
Support vector machine (SVM)



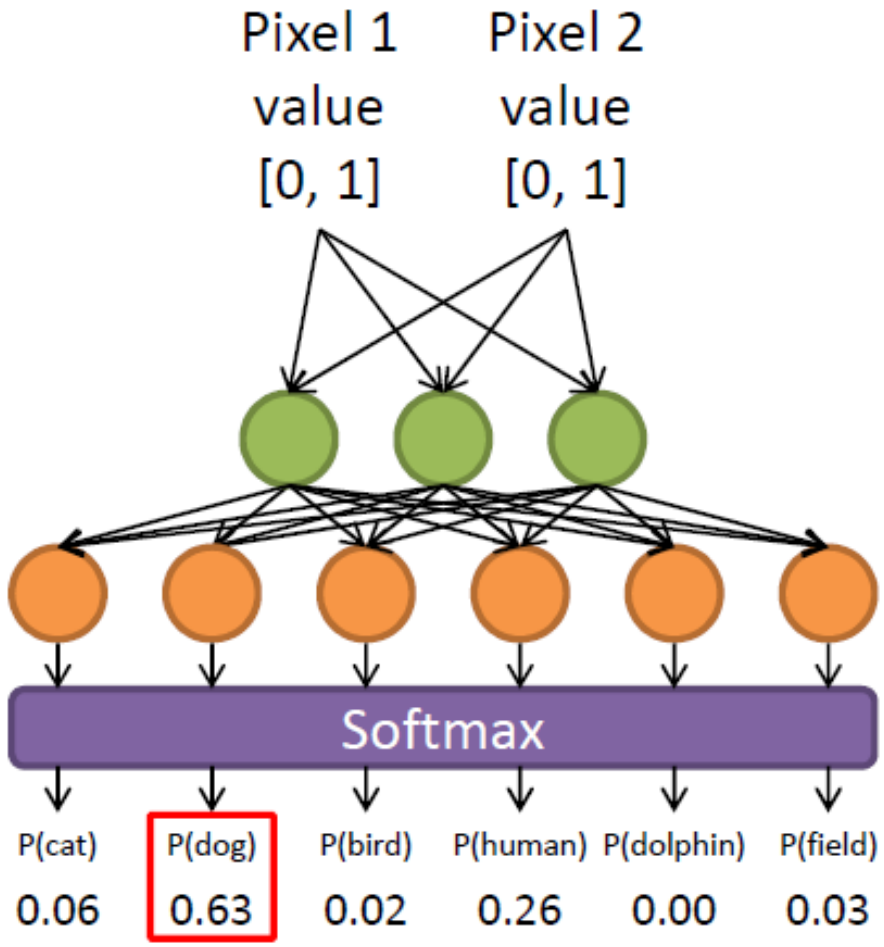
Neural Network



Neural Network

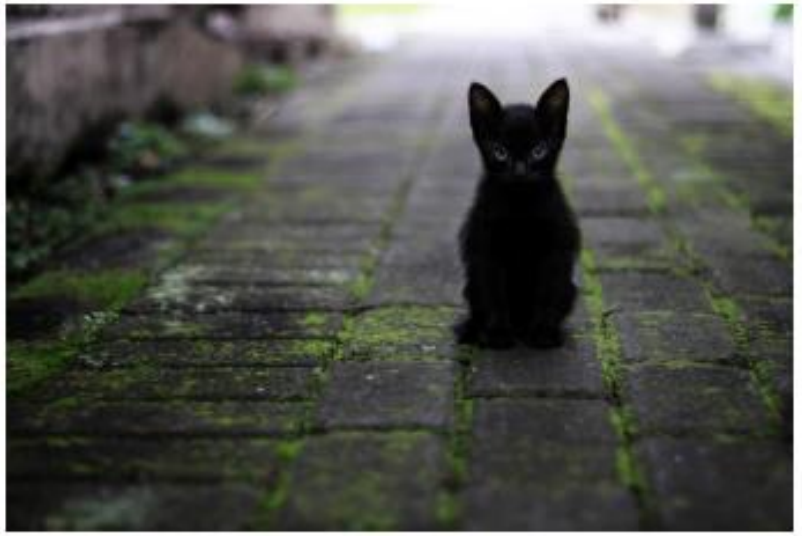


Neural Network

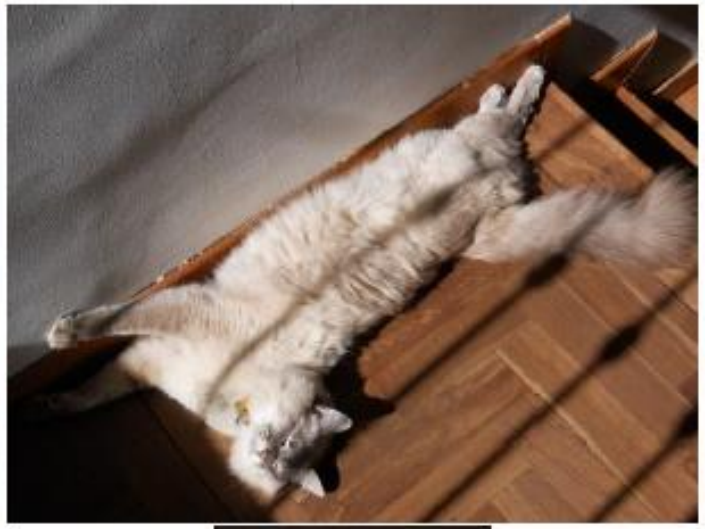


P(cat) + P(dog) + P(bird) + P(human) + P(dolphin) + P(field) = 1.0

Challenge


Challenge: deformation



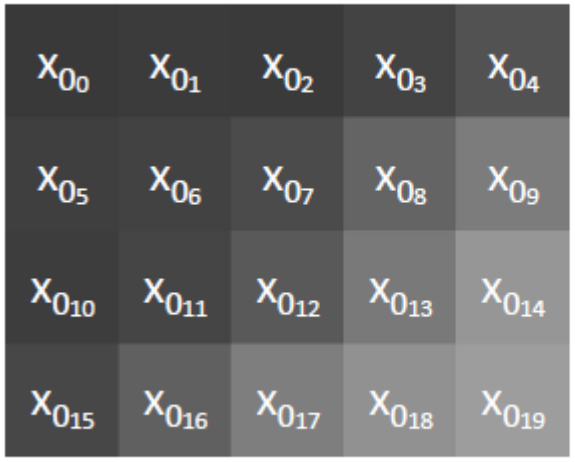
Challenge: Occlusion

Outline

1. Computer vision – digital images.
2. Image Classification
3. Deep Learning

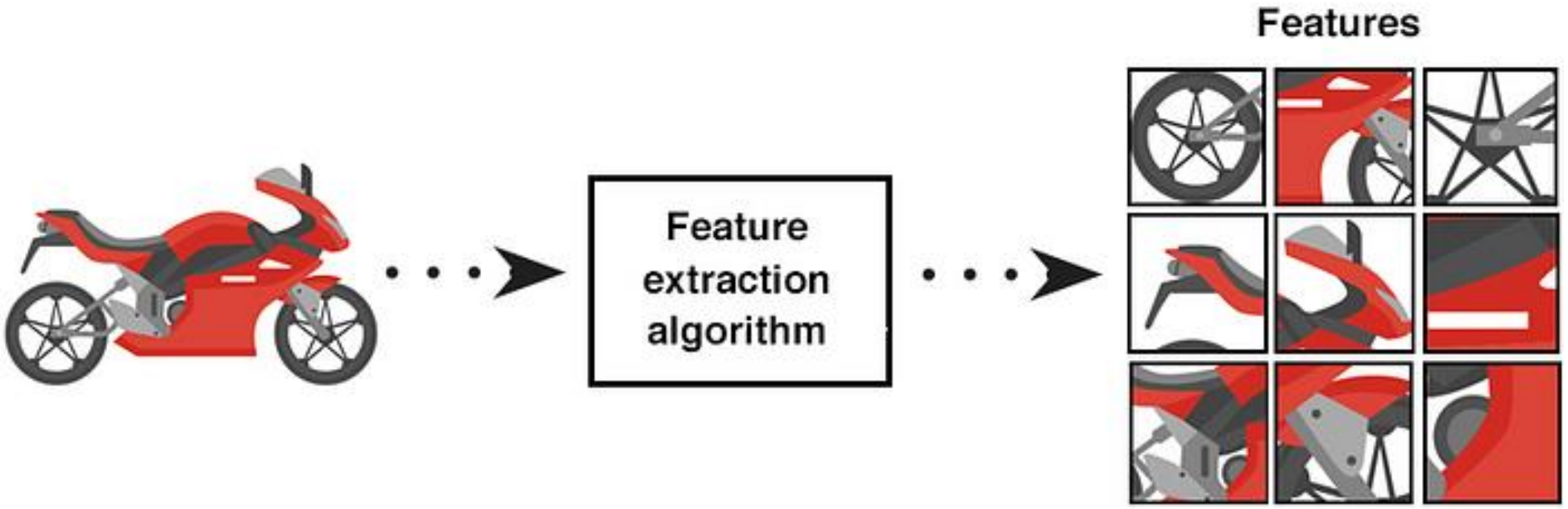
Raw Data

Example of sample x_0



$x_{0_0} = 0.22$
$x_{0_1} = 0.23$
$x_{0_2} = 0.23$
$x_{0_3} = 0.26$
$X_{0_4} = 0.32$
$x_{0_5} = 0.25$
$x_{0_6} = 0.26$
⋮

Feature extraction

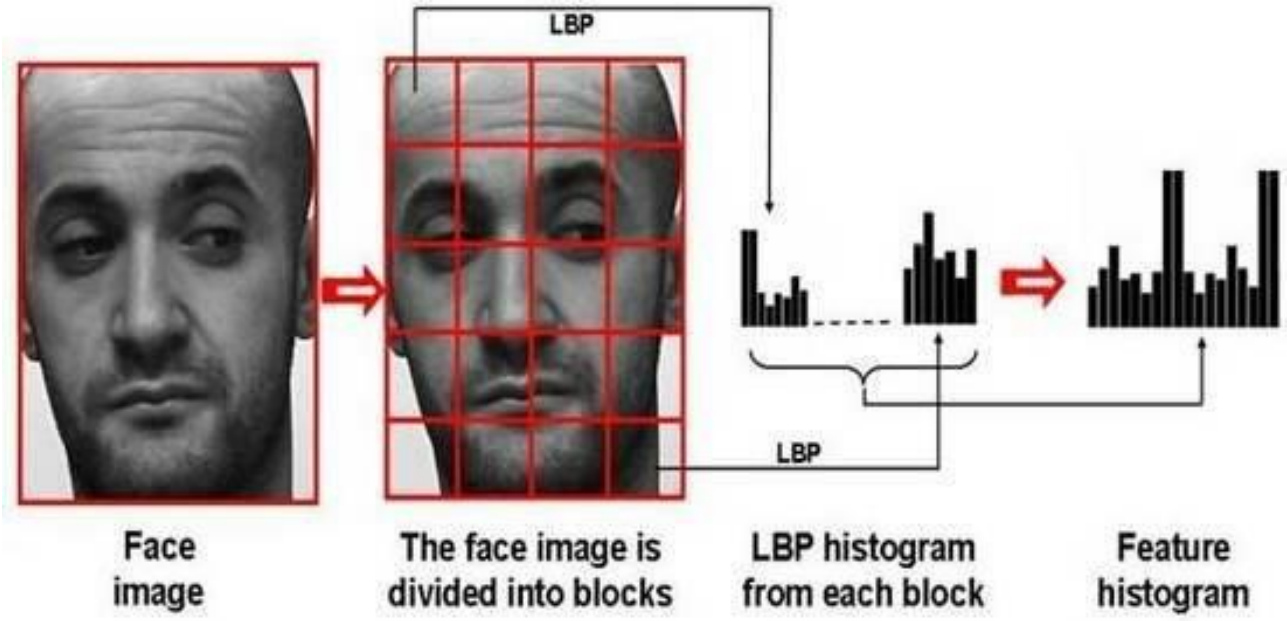


Feature Extraction Methods

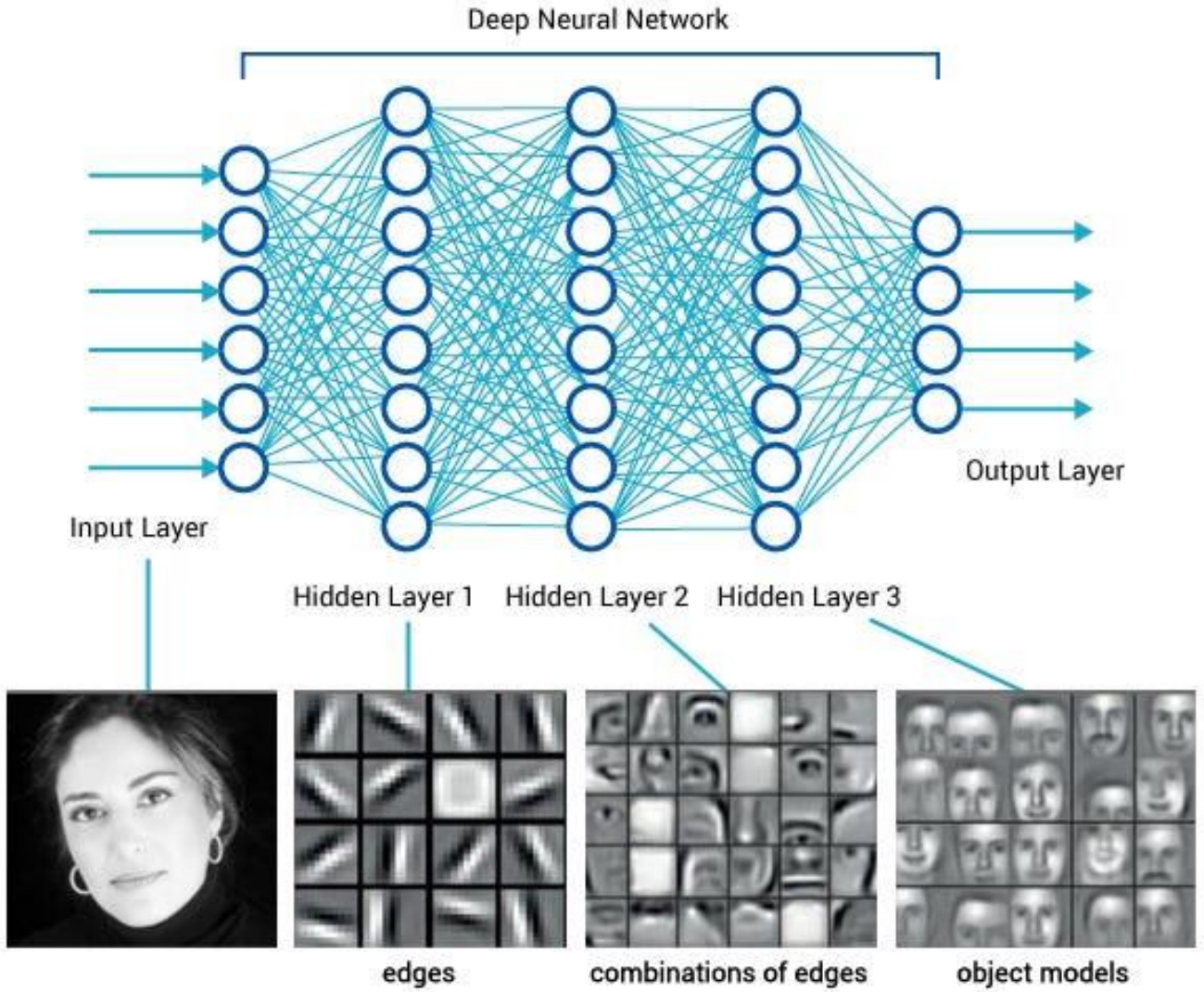
1. Convolutional neural networks.
2. Grey scale features: Image similarity, Text Recognition, Edge Detection, Medical Imaging, Facial Recognition.
3. Mean Pixel value of channels: Mean intensity value of each cannel.
4. Edge Features.
5. AutoEnconders.
6. Histogram of Oriented Gradients.
7. Scale-Invariant Feature Transform (SIFT).
8. Local Binary Patterns (LBP).
9. Frecuency-based features.
10. Color based features.

https://medium.com/@mohitjavali/10-ways-to-extract-features-from-an-image-f44c8e9b0fbf

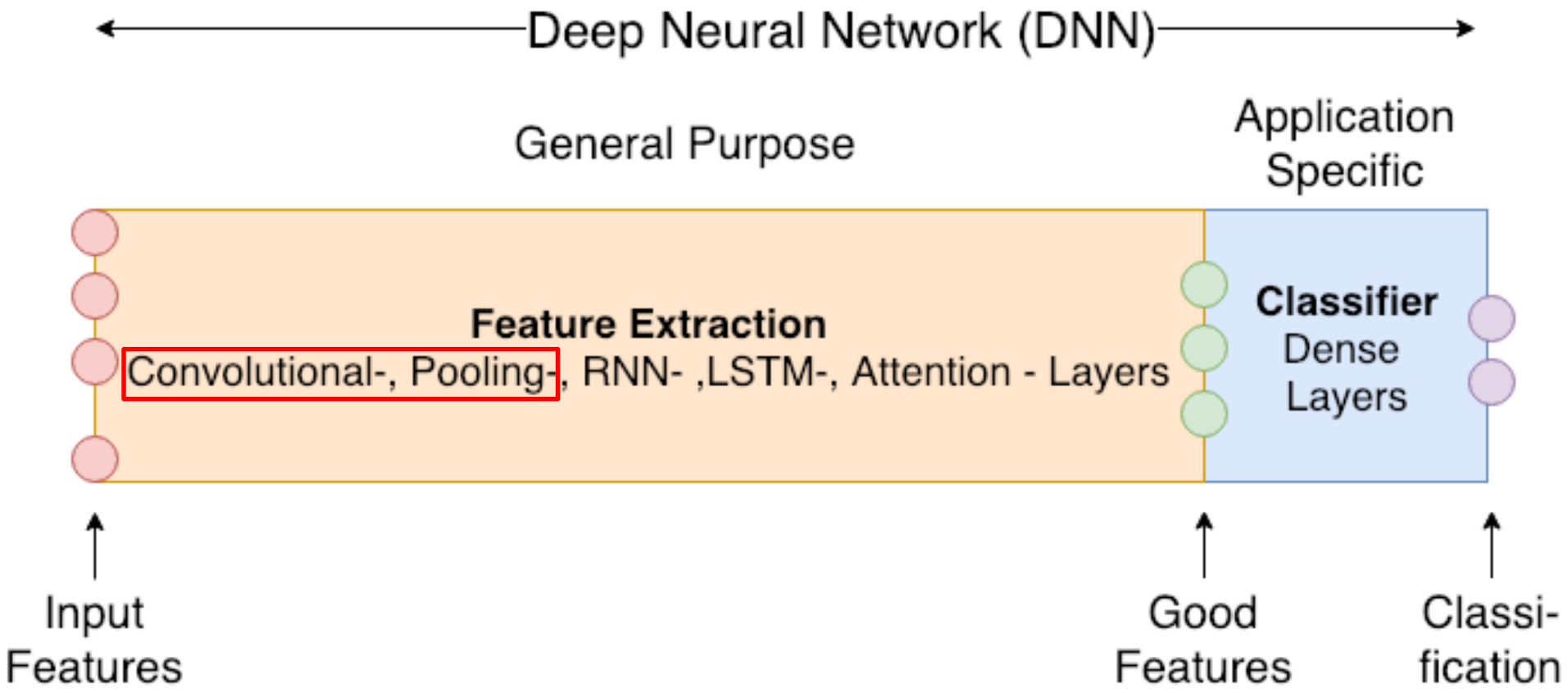
Features - LBP



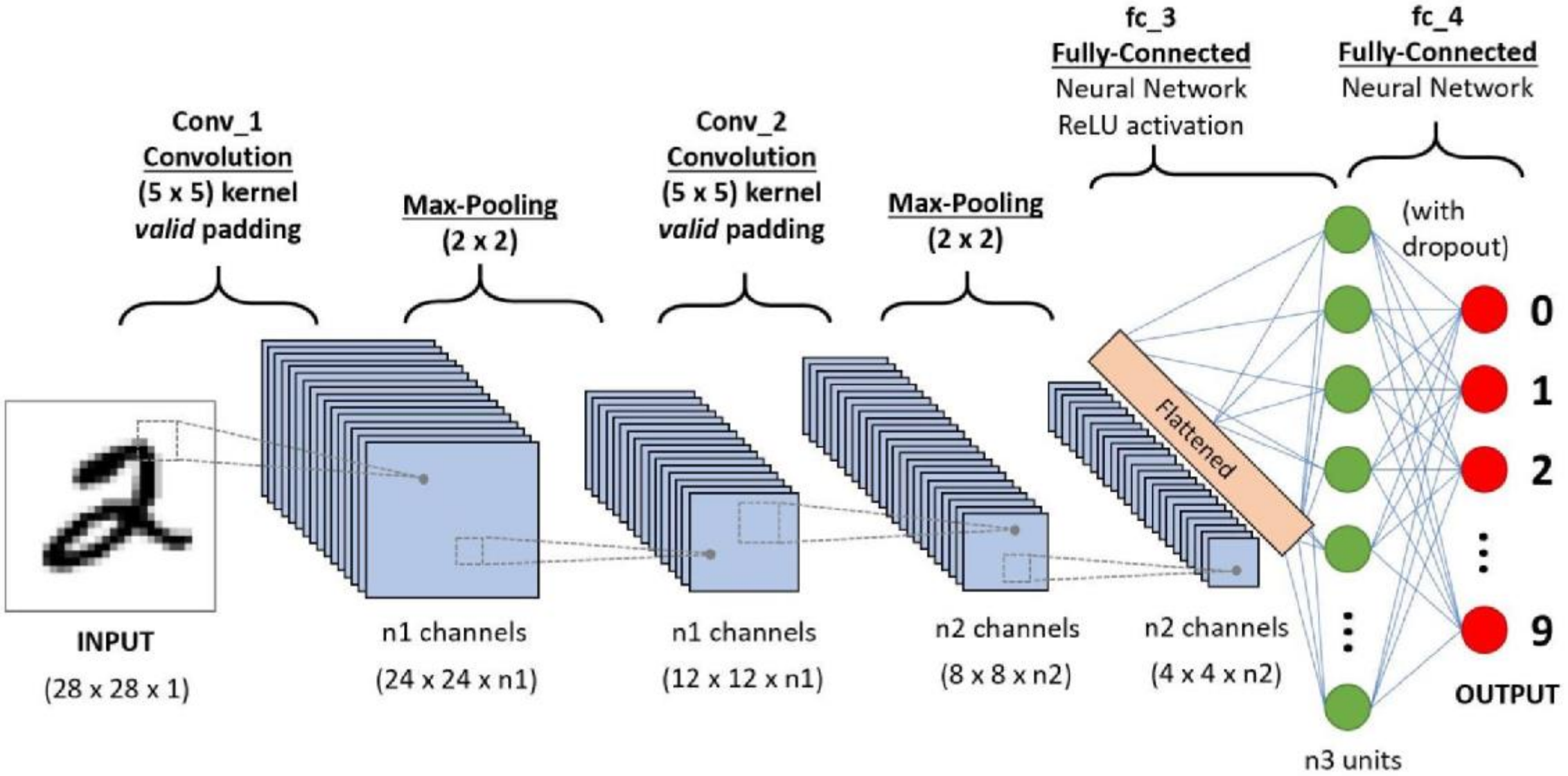
Deep learning



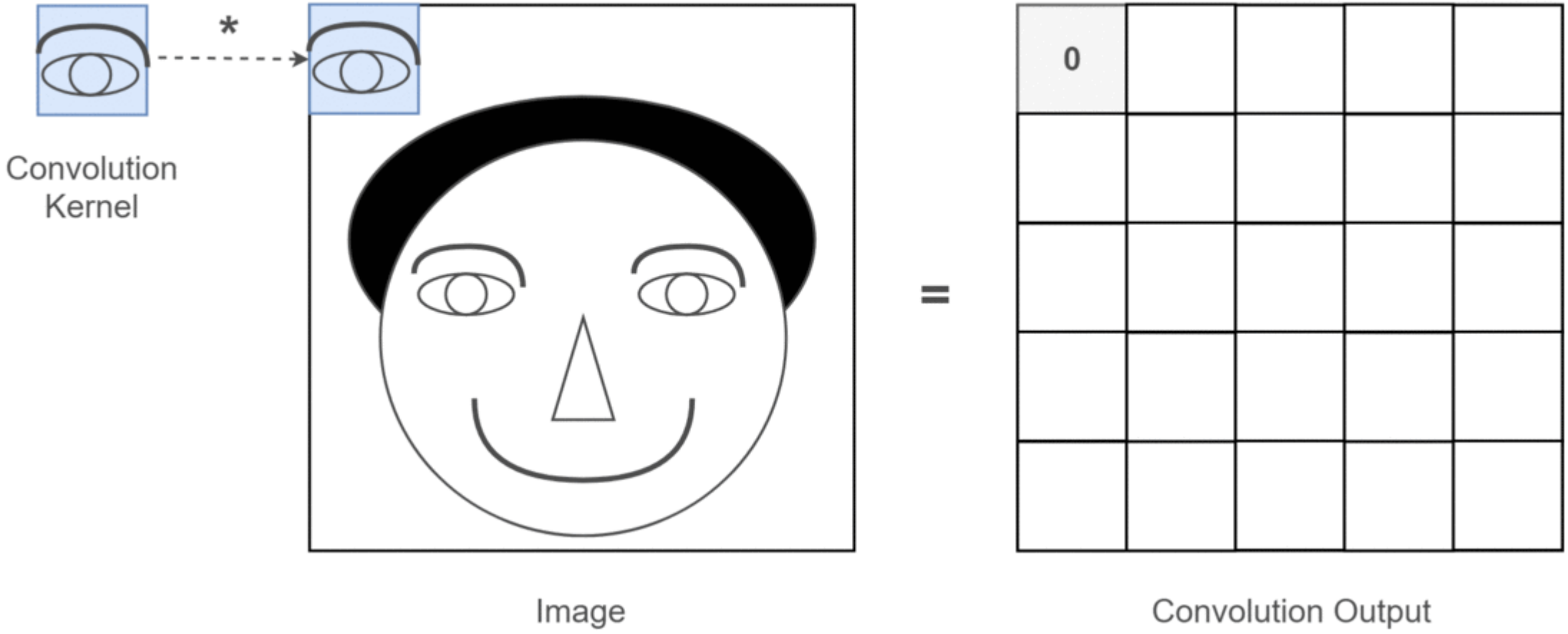
Deep Learning



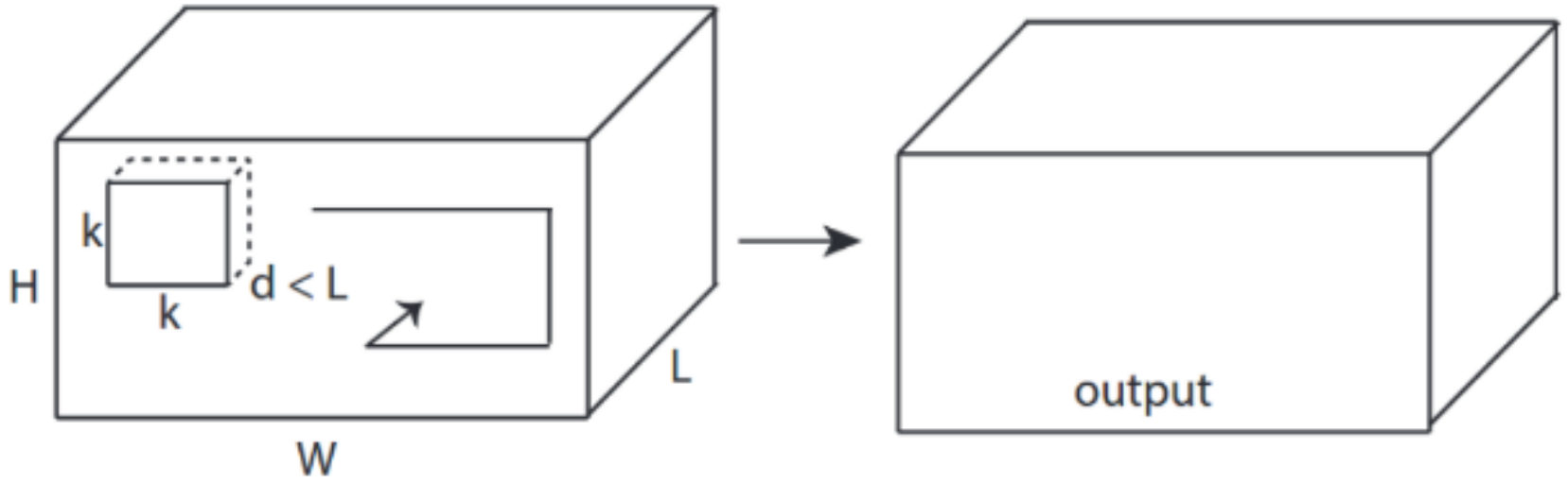
Convolutional Neural Network

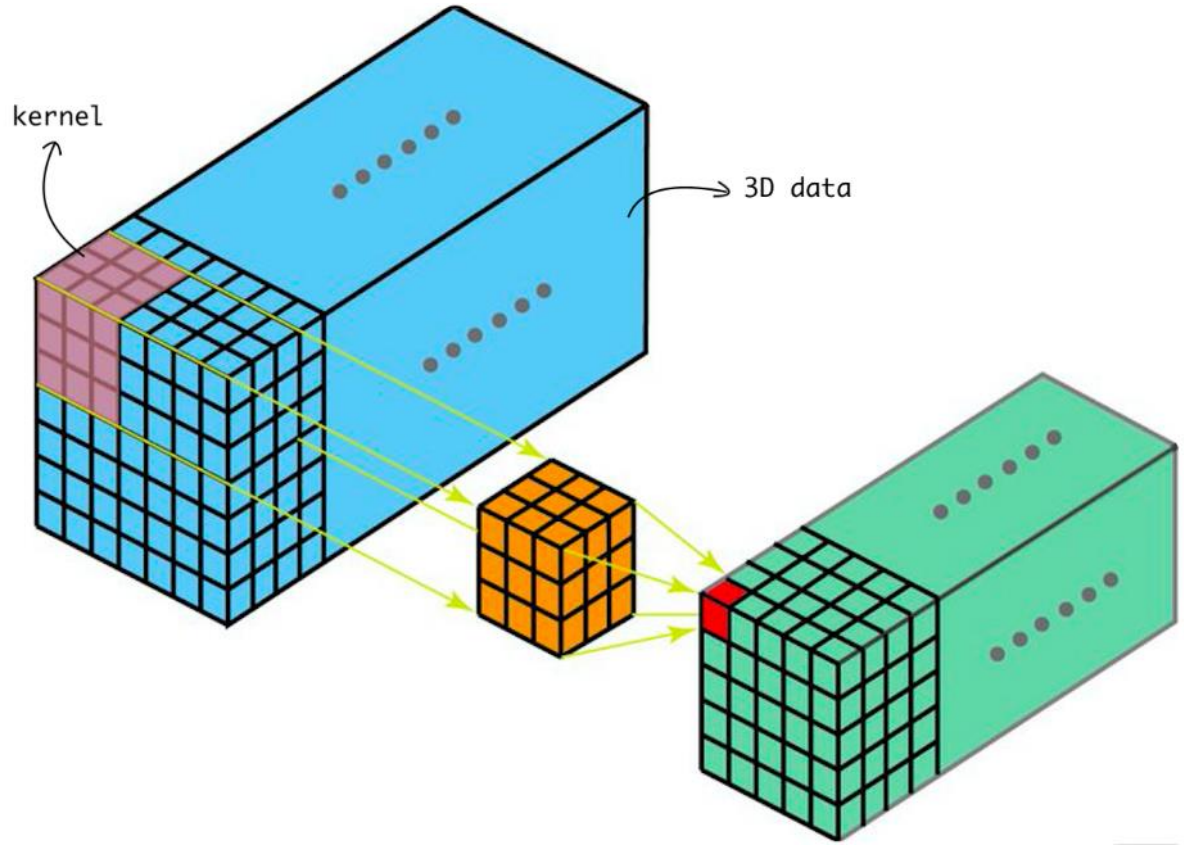


Convolution 2D

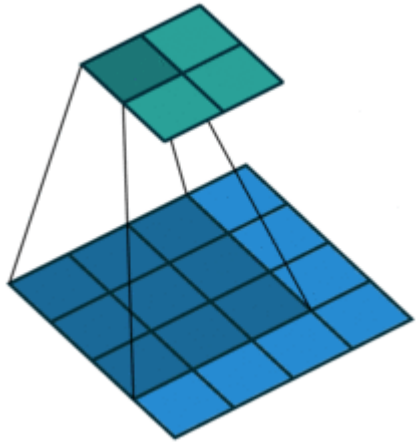


3D CNN

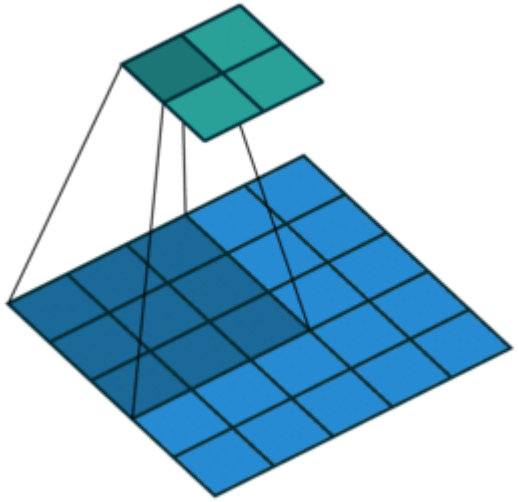




Convolution 2D

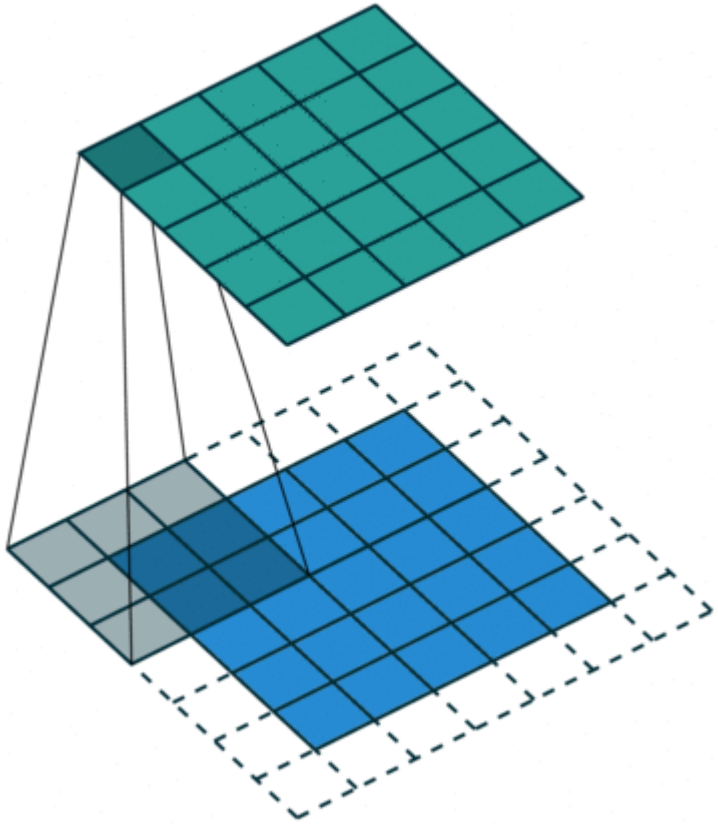


padding = 0, stride = 1

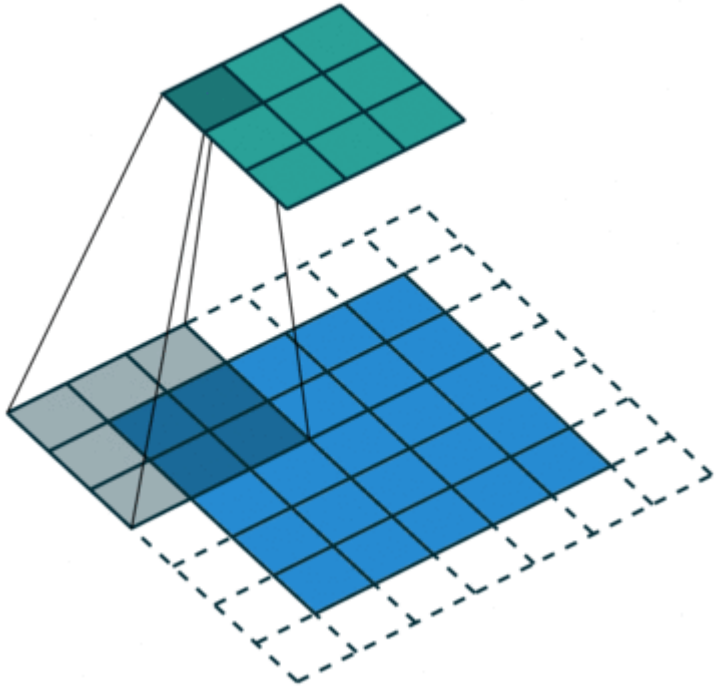


padding = 0, stride = 2

Convolution 2D

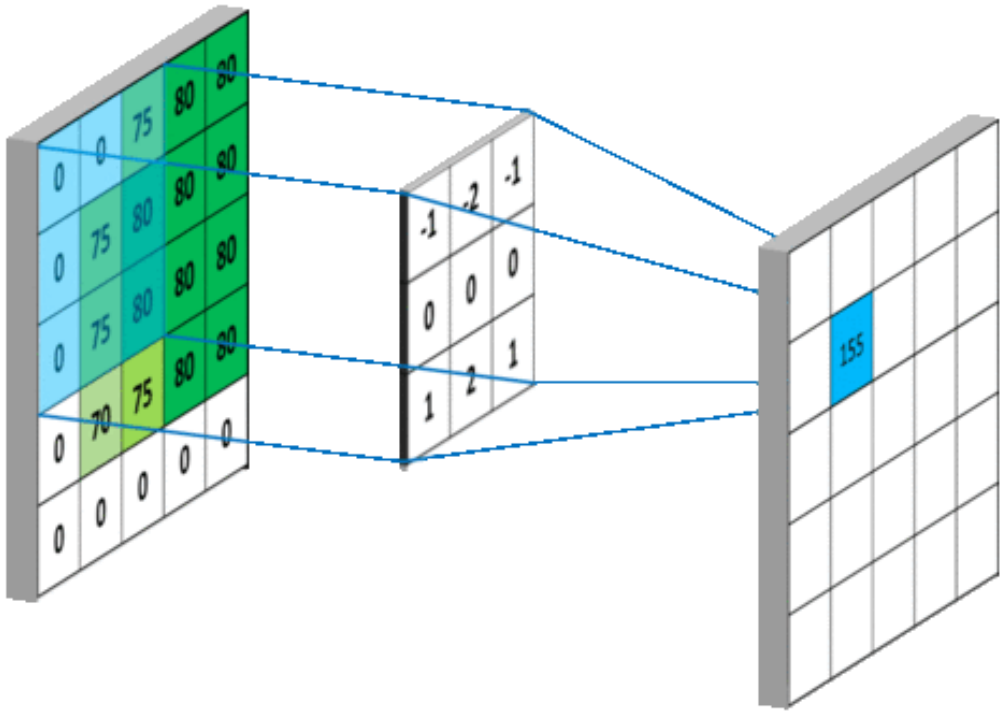


padding = 1, stride = 1



padding = 1, stride = 2

Convolutional layer - example



Excercise

Image

59	58	67	82
66	75	100	124
69	89	121	150

Kernel

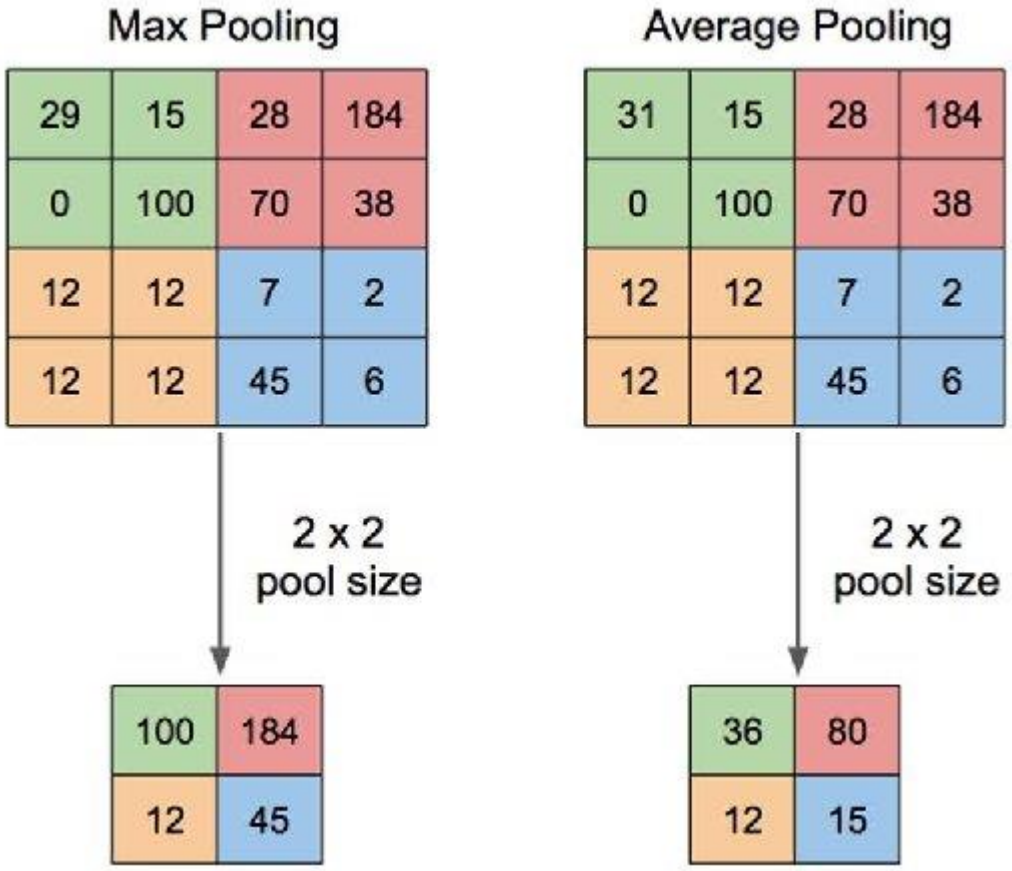
-1	-1
-1	3

Kernel: 2x2
Stride: 1
Padding: valid

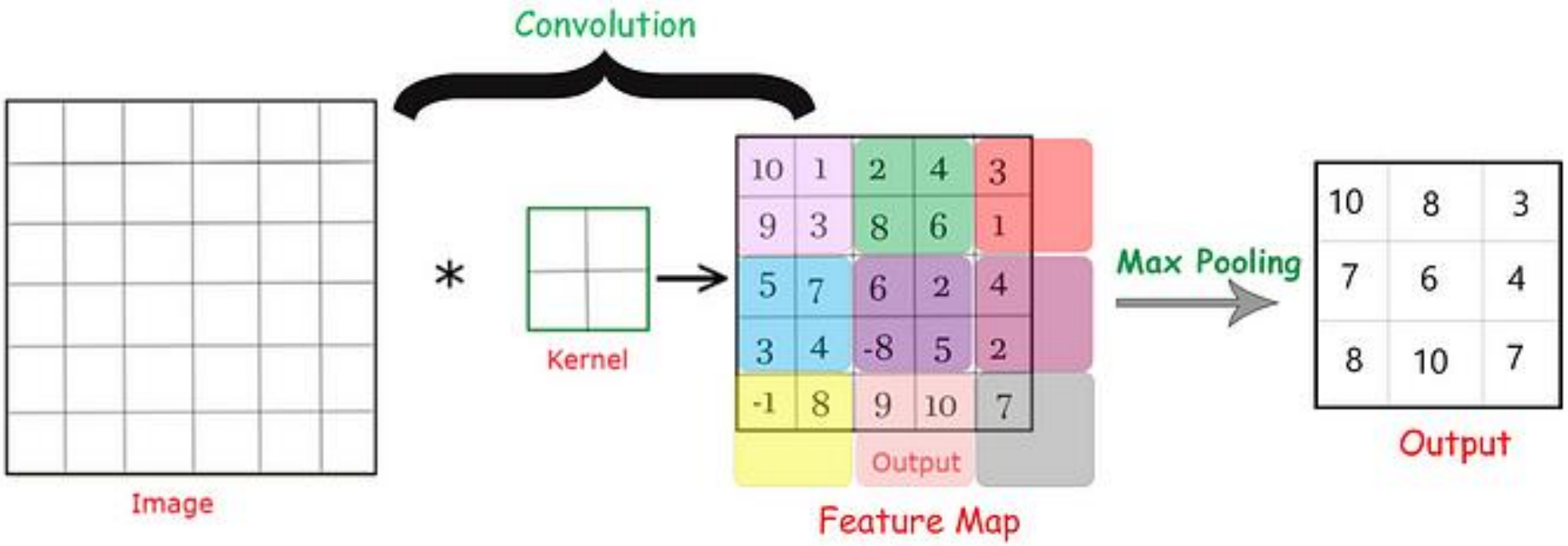
Output

	x	

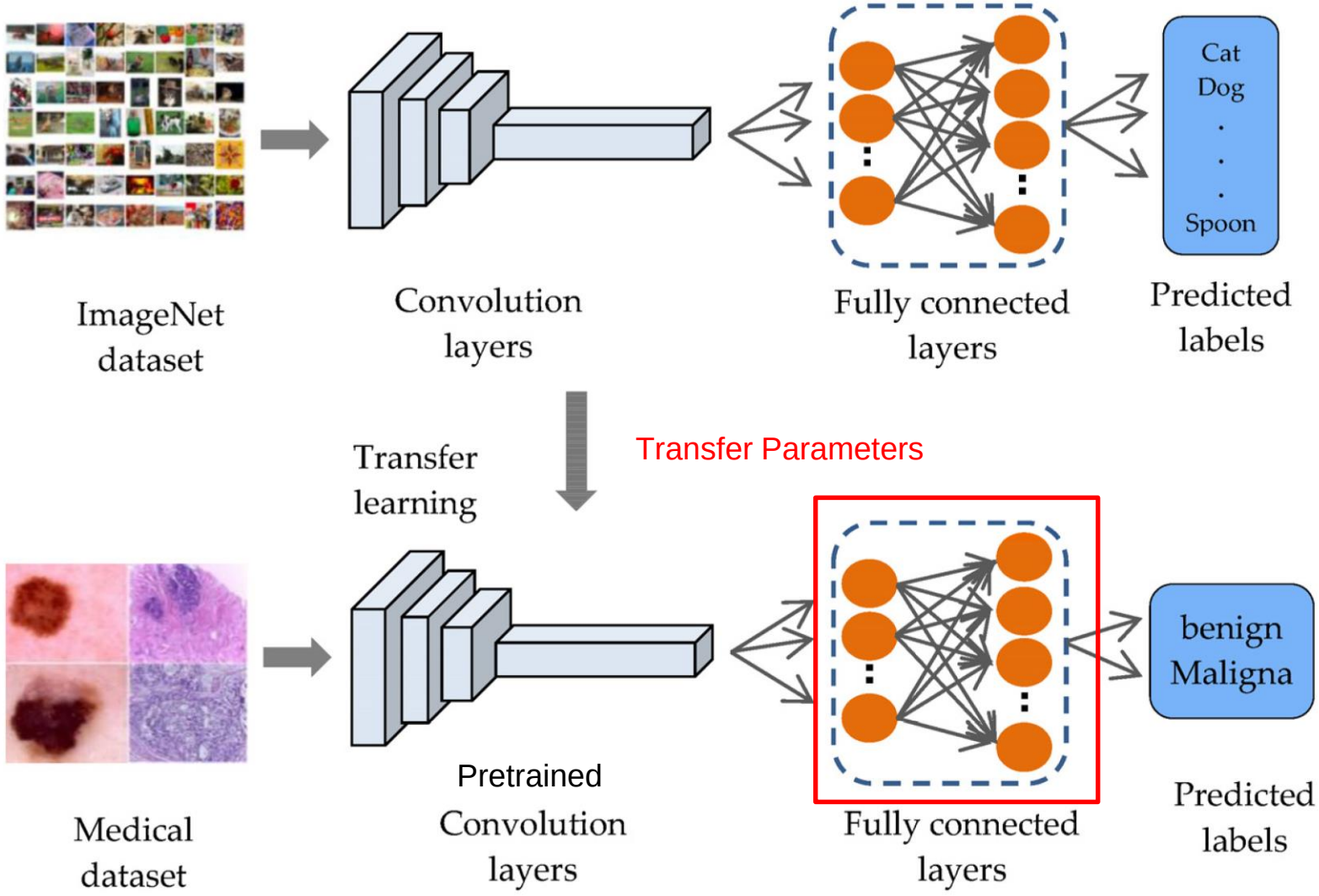
Pooling



Max Pooling – Example



Transfer Learning



https://www.mdpi.com/1424-8220/23/2/570

Referencias

https://www.coursera.org/learn/computer-vision-with-embedded-machine-learning/

https://hannibunny.github.io/mlbook/neuralnetworks/03ConvolutionNeuralNetworks.html

https://wandb.ai/ayush-thakur/dl-question-bank/reports/Intuitive-understanding-of-1D-2D-and-3D-convolutions-in-convolutional-neural-networks---VmlldzoxOTk2MDA

https://www.tensorflow.org/tutorials/keras/classification?hl=es-419

https://www.linkedin.com/pulse/convolutional-neural-networks-ahtesham-iqbal/

Institución Universitaria
Reacreditada en Alta Calidad

¡Gracias!

Somos Innovación Tecnológica con Sentido Humano

Alcaldía de Medellín