



Institución
Universitaria
Reacreditada en Alta Calidad

Computer Vision - Convolution

Aprendizaje Automático Embebido

Somos Innovación Tecnológica con *Sentido Humano*



Alcaldía de Medellín

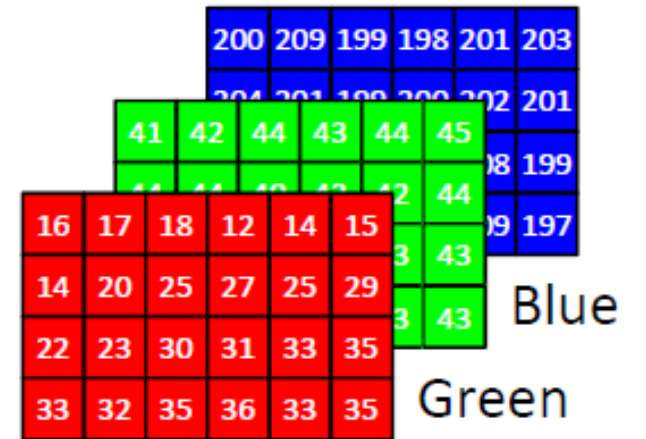


Institución
Universitaria
Reacreditada en Alta Calidad

Outline

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







Institución
Universitaria
Reacreditada en Alta Calidad



Somos Innovación Tecnológica con *Sentido Humano*



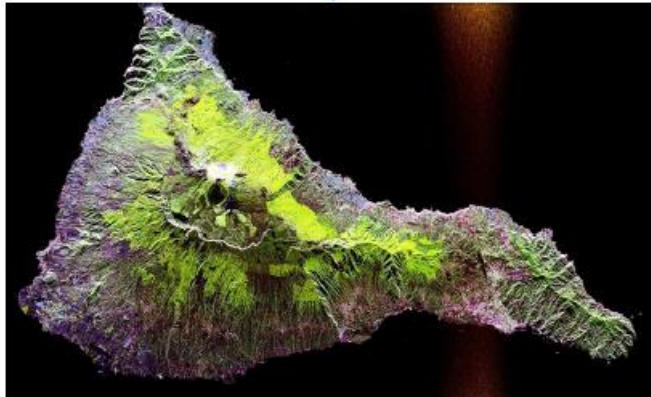
"Thermal image of a building", by Marco Verch (CC BY 2.0)



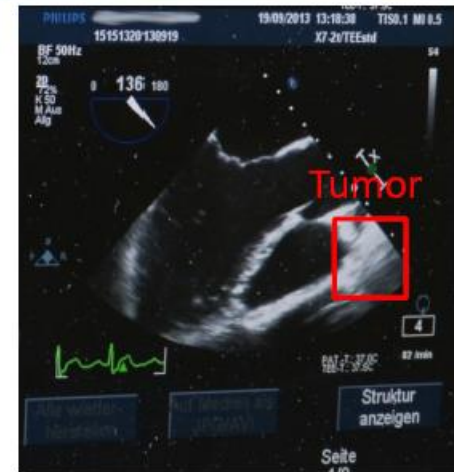
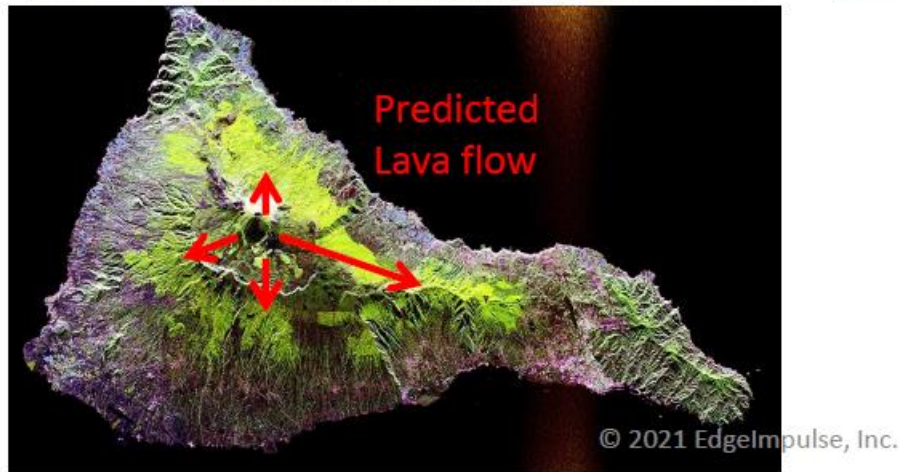
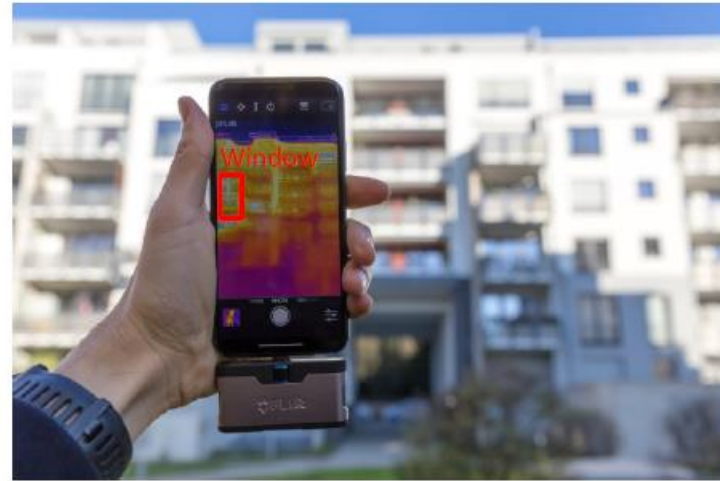
SIR-C/X-SAR imaging radar
on Space Shuttle Endeavour



Teide volcano

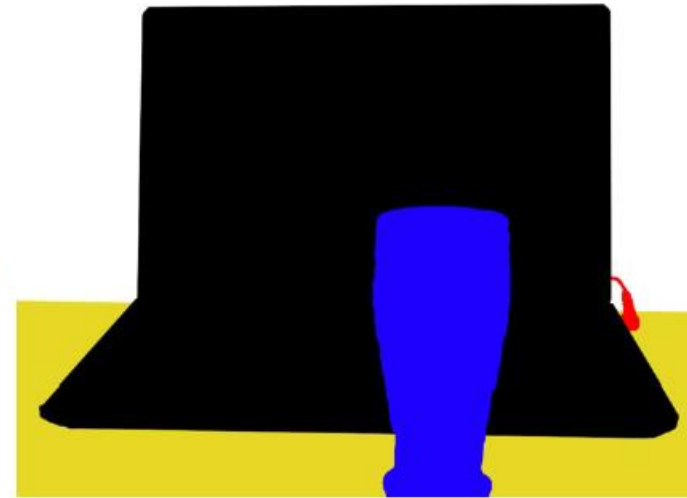


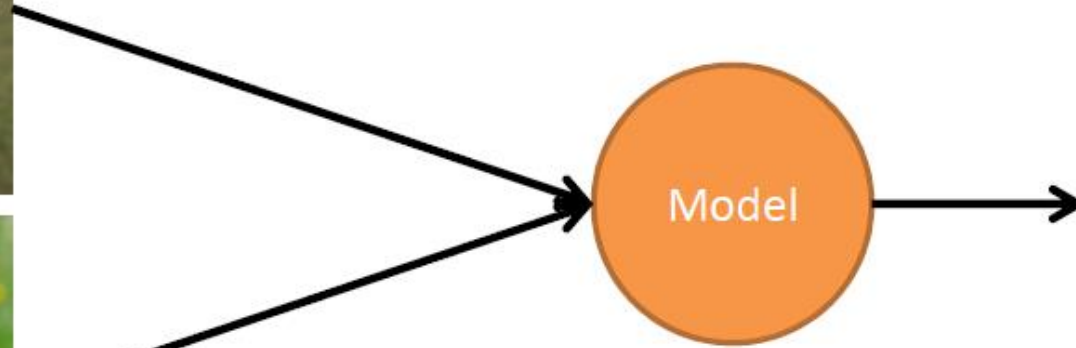
mpulse, Inc.



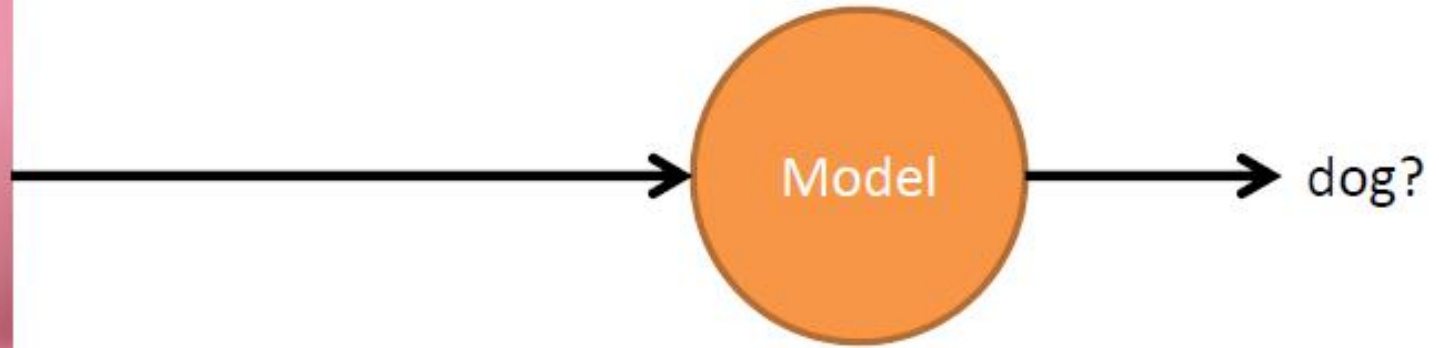


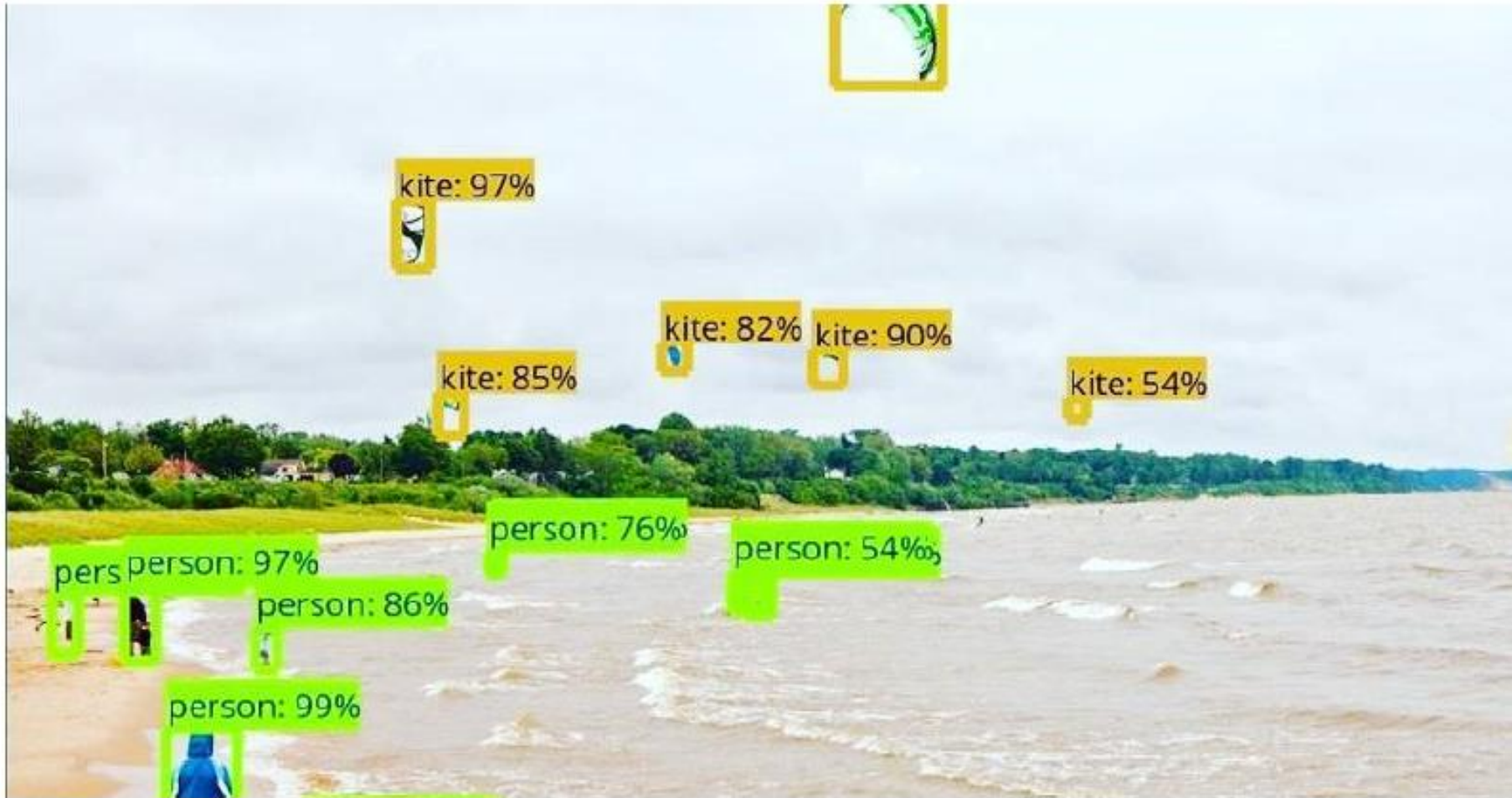
Institución
Universitaria
Reacreditada en Alta Calidad





© 2021 EdgeImpulse, Inc.





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



$$240 \times 240 \times 8 \times 30 \times 1 = 13.824 \text{ Mbps}$$

Number of pixels / Num. cameras
Bits per pixel / Frames per second



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

$$1 \times 30 \times 1 = 30 \text{ bps}$$

Num. cameras
Person or not person / Frames per second



Institución
Universitaria
Reacreditada en Alta Calidad



Somos Innovación Tecnológica con *Sentido Humano*



Institución
Universitaria
Reacreditada en Alta Calidad

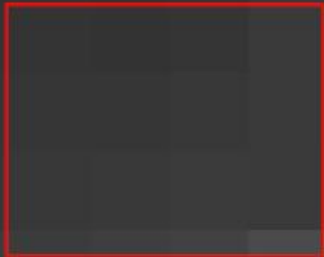


Somos Innovación Tecnológica con *Sentido Humano*

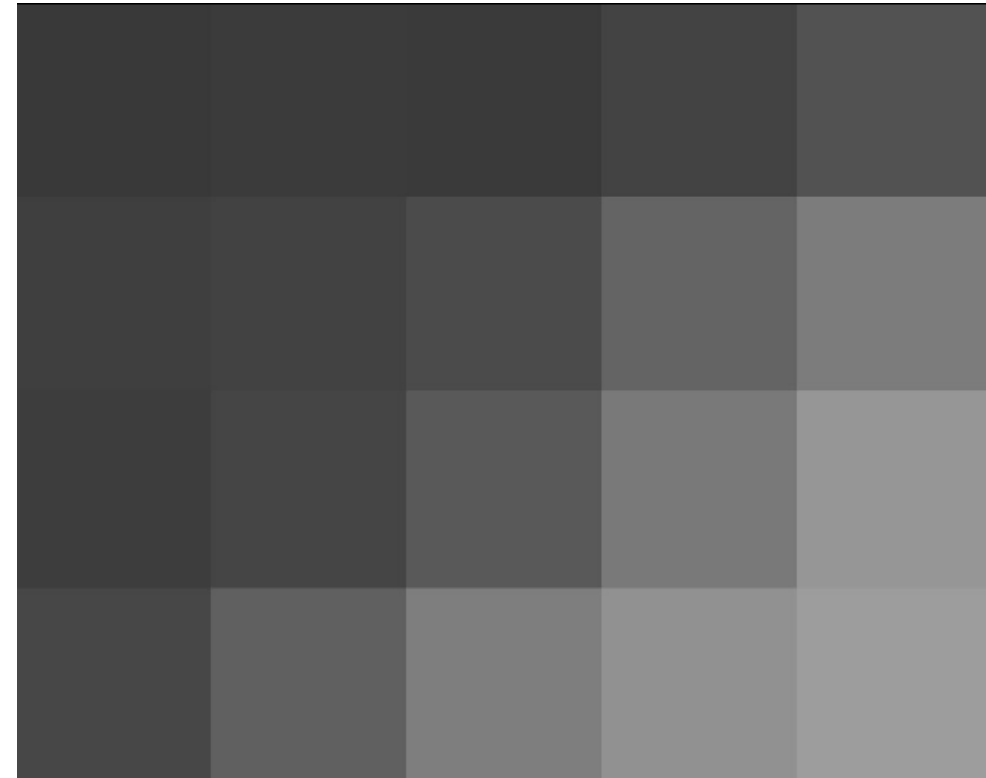


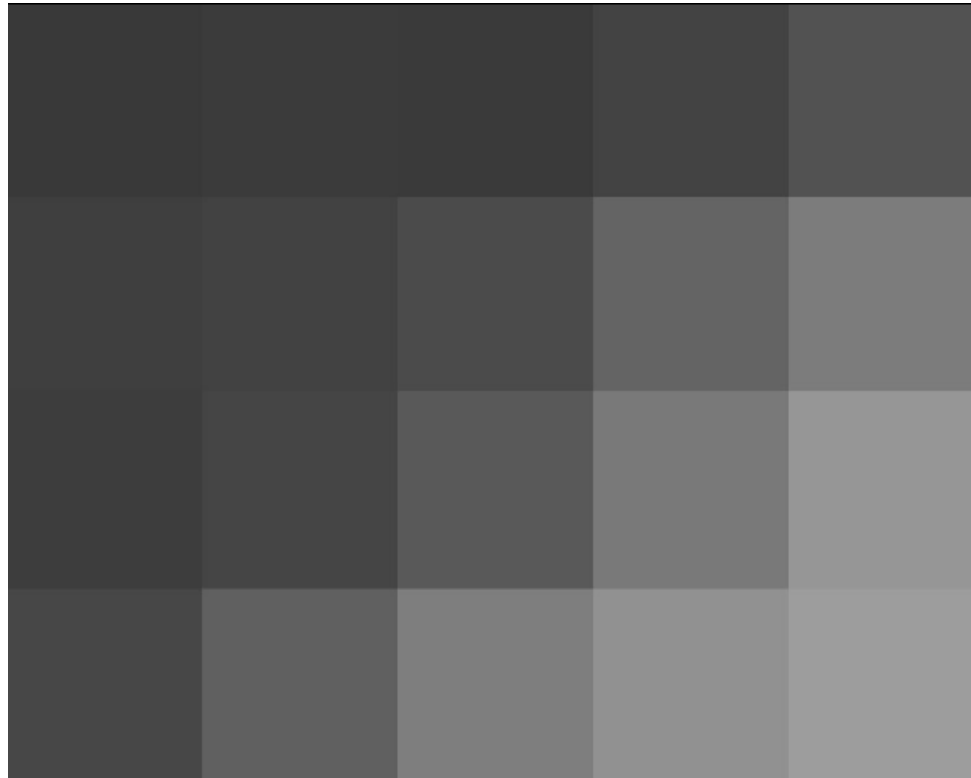
Institución
Universitaria
Reacreditada en Alta Calidad

picture element ("pixel")



© 2021 EdgeImpulse, Inc.



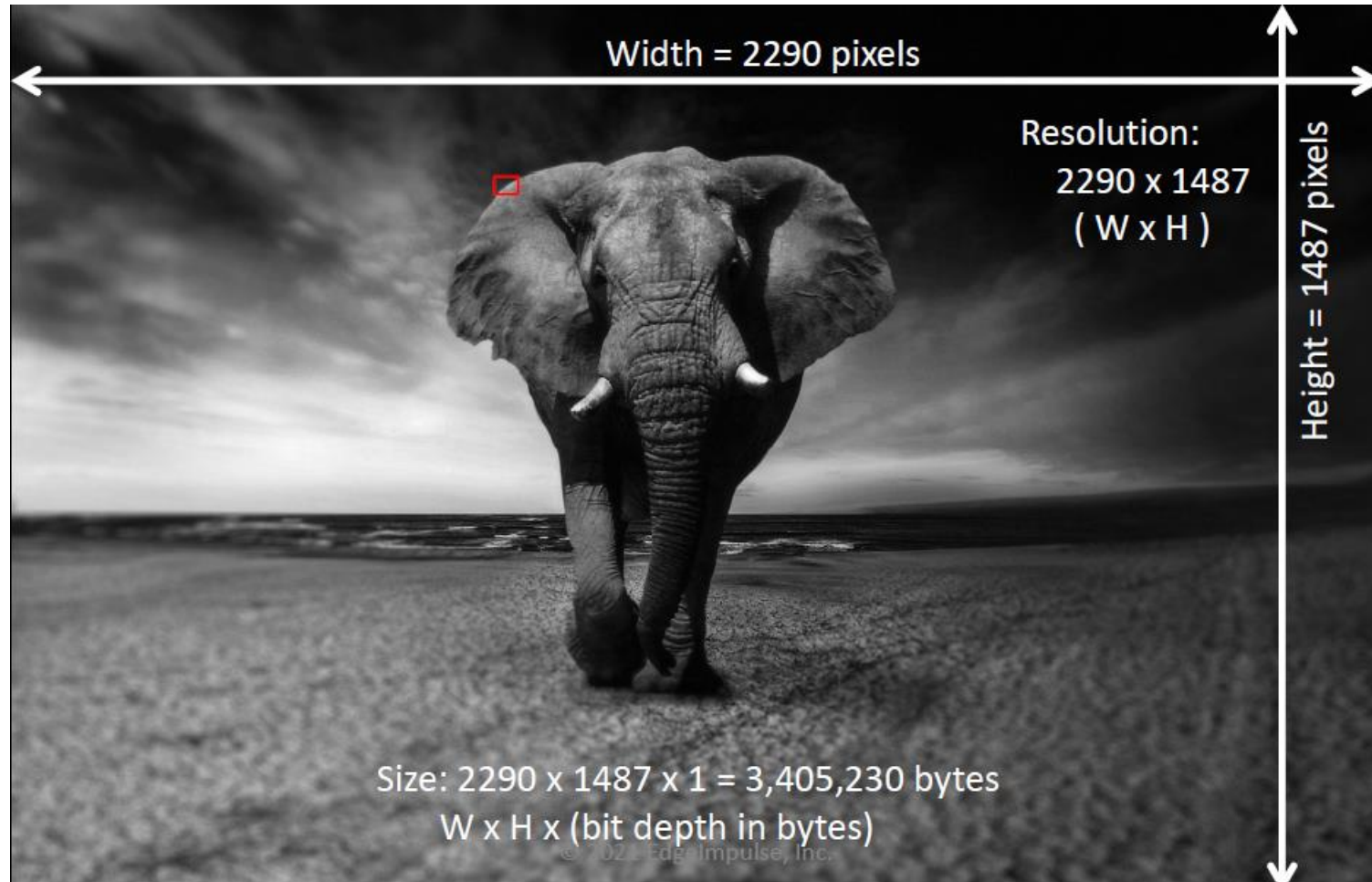


0.22	0.23	0.23	0.26	0.32
0.25	0.26	0.29	0.39	0.49
0.24	0.27	0.35	0.47	0.59
0.28	0.38	0.49	0.57	0.62

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





Institución
Universitaria
Reacreditada en Alta Calidad



Somos Innovación Tecnológica con *Sentido Humano*

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: $8 \times 3 = 24$ bits

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

Optional

- 1 byte: Alpha

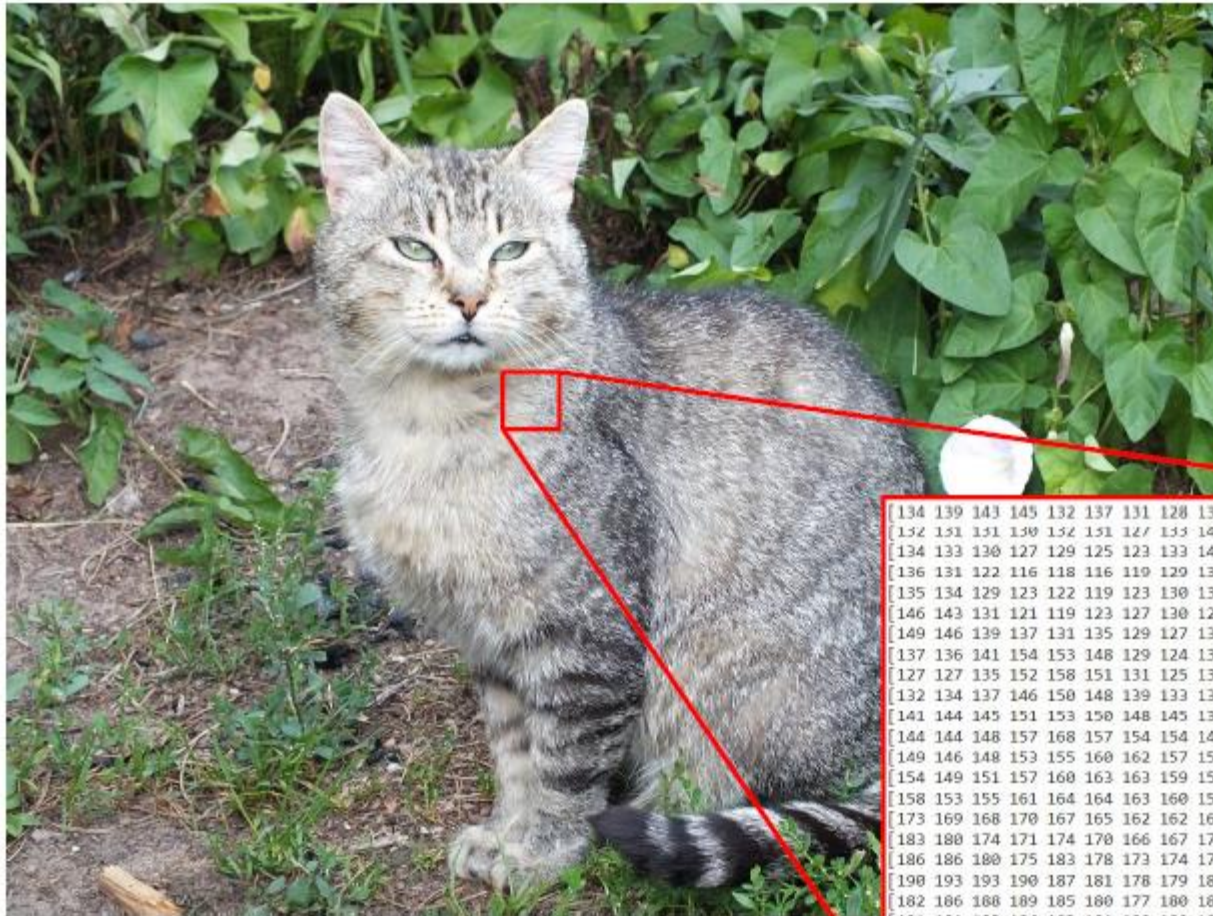


Institución
Universitaria
Reacreditada en Alta Calidad

Outline

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

Image Classification



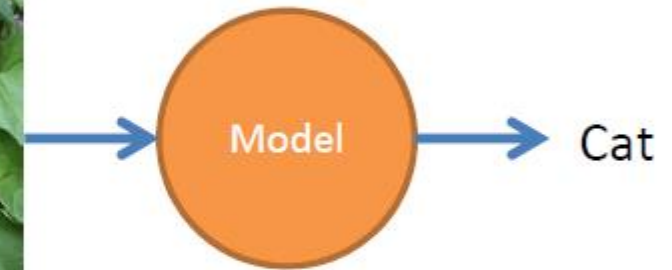
Resolution: 2048 x 1536 pixels

Bit depth: 24 bits

Numpy array shape: (1536, 2048, 3)

```
[134 139 143 145 132 137 131 128 134 117 127 132 117 124 11
[152 151 151 150 152 151 147 153 145 150 158 125 111 127 11
[134 133 130 127 129 125 123 133 140 136 155 131 124 138 14
[136 131 122 116 118 116 119 129 132 138 174 155 146 154 15
[135 134 129 123 122 119 123 130 133 139 145 145 148 155 16
[146 143 131 121 119 123 127 130 129 127 130 136 154 163 16
[149 146 139 137 131 135 129 127 131 126 122 131 149 163 17
[137 136 141 154 153 148 129 124 137 133 124 132 145 153 16
[127 127 135 152 158 151 131 125 135 129 120 128 144 142 15
[132 134 137 146 150 148 139 133 130 118 113 125 133 140 14
[141 144 145 151 153 150 148 145 134 122 124 134 134 146 15
[144 144 148 157 168 157 154 154 145 139 143 149 155 152 15
[149 146 148 153 155 160 162 157 153 153 155 156 162 167 17
[154 149 151 157 160 163 163 159 155 156 157 157 177 173 17
[158 153 155 161 164 164 163 160 159 161 163 162 182 177 17
[173 169 168 170 167 165 162 162 165 168 170 169 174 177 18
[183 180 174 171 174 170 166 167 171 175 176 176 176 180 18
[186 186 180 175 183 178 173 174 178 179 180 180 188 182 17
[190 193 193 190 187 181 178 179 181 181 181 182 186 176 16
[182 186 188 189 185 180 177 180 181 181 182 184 172 169 16
[181 181 183 186 183 181 180 182 185 185 181 176 172 170 16
[180 181 183 185 180 182 184 188 188 184 177 172 175 171 16
[182 184 186 188 181 183 187 189 188 183 177 173 174 170 16
[185 186 188 188 185 186 187 186 183 180 179 179 174 170 16
```


Image Classification



Binary Classification

One-vs-One



VS.



One-vs-Rest



VS.



Multiclass Classification



VS.



VS.



VS.



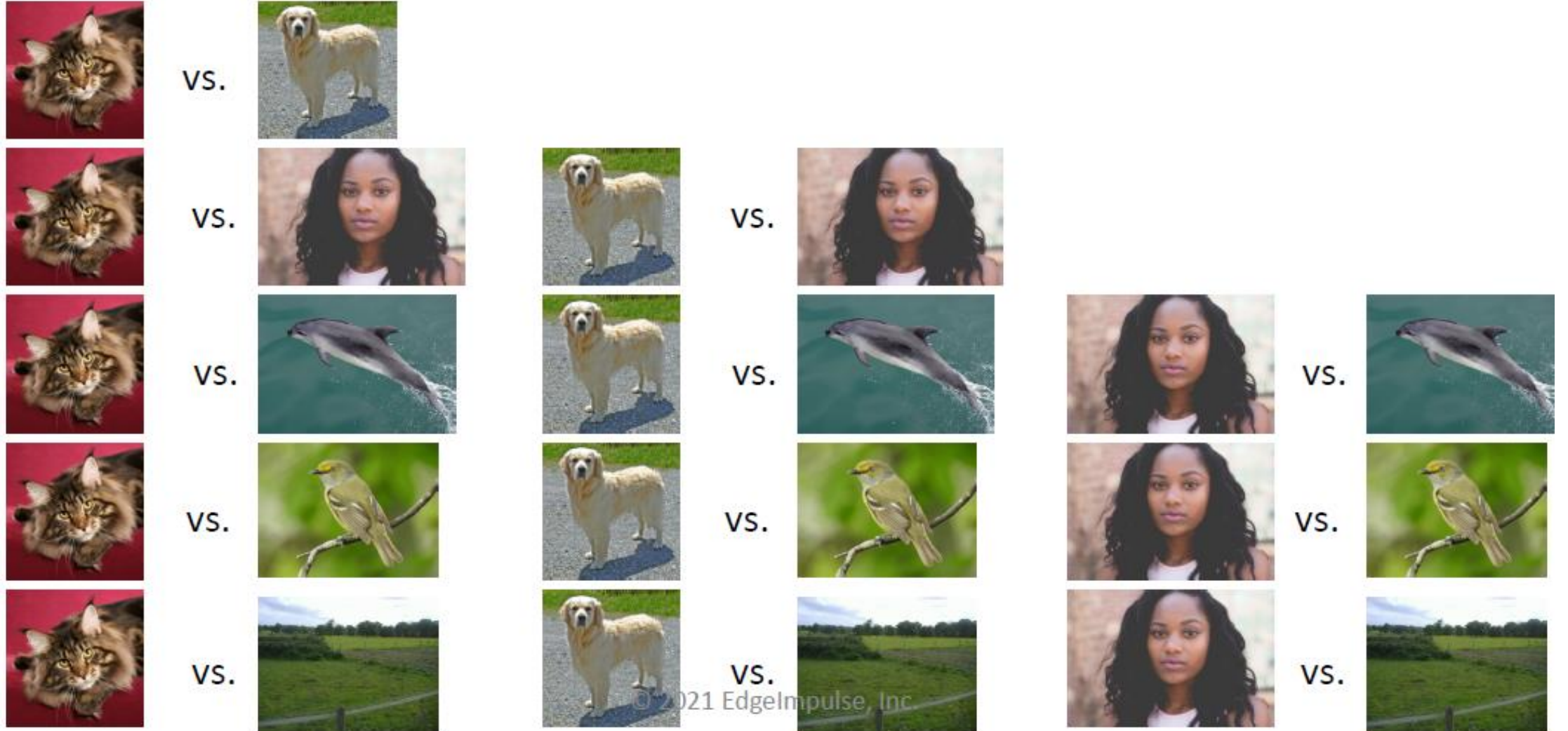
VS.



VS.



Multiclass: One-vs-One



© 2021 Edgelpulse, Inc.

Somos innovación tecnológica con

genio humano

Multiclass: One-vs-Rest



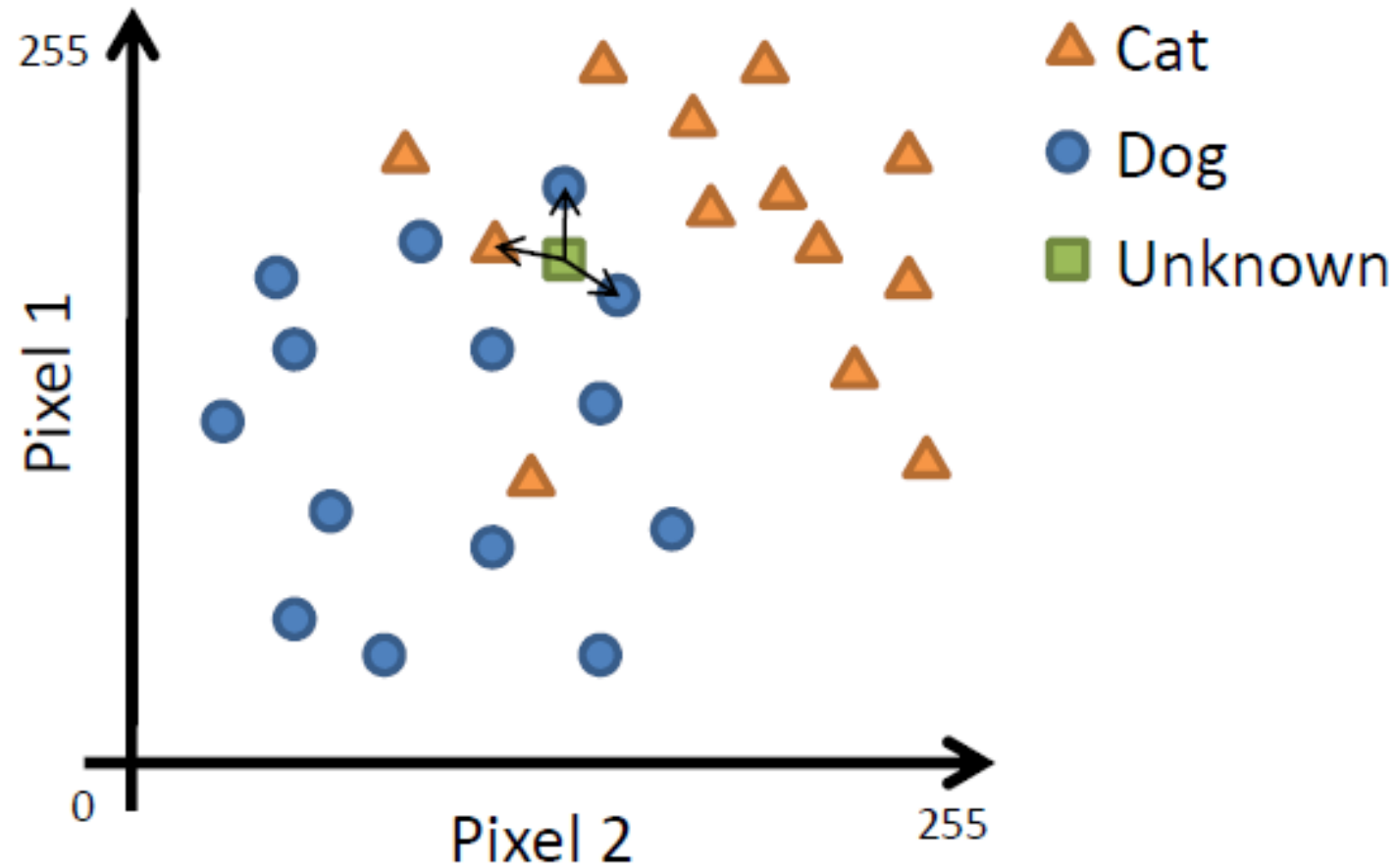
VS.



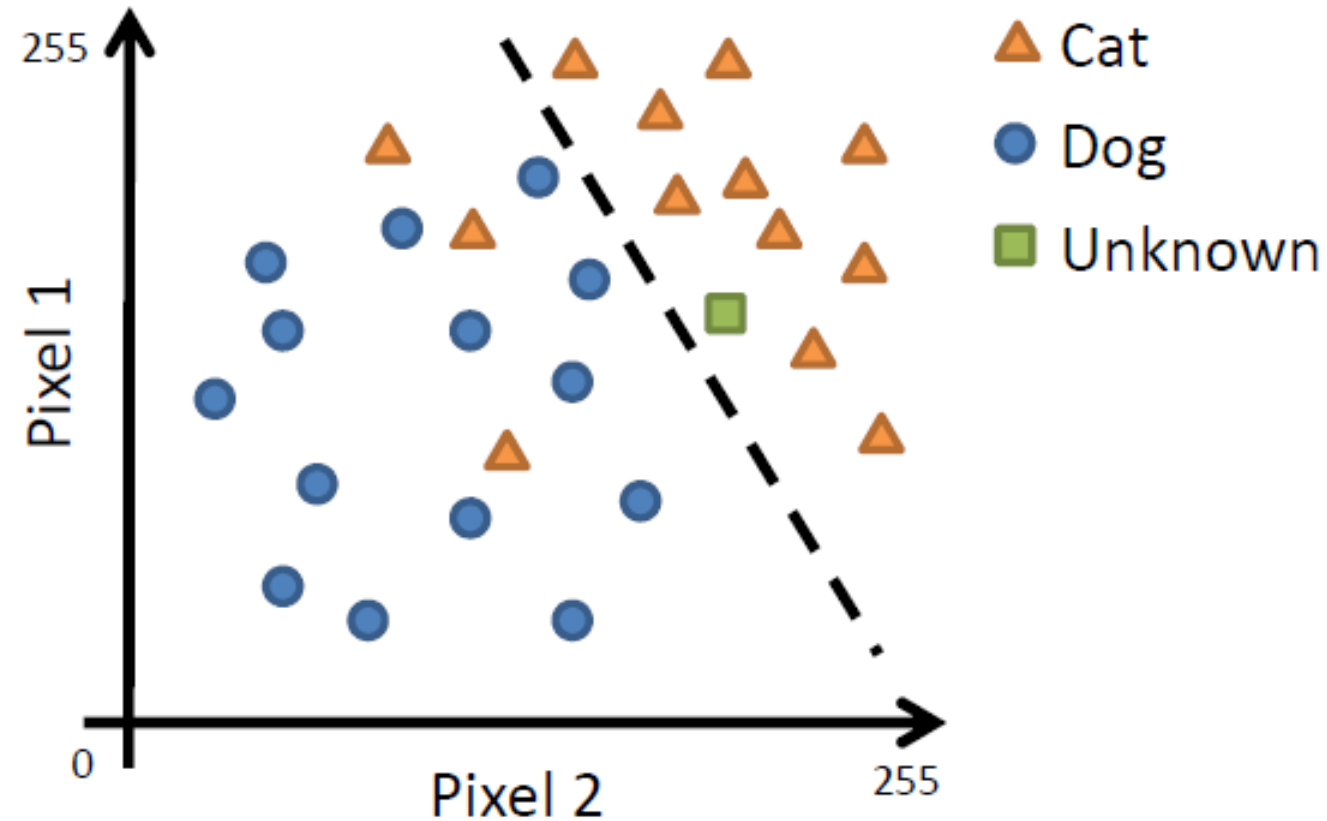
VS.



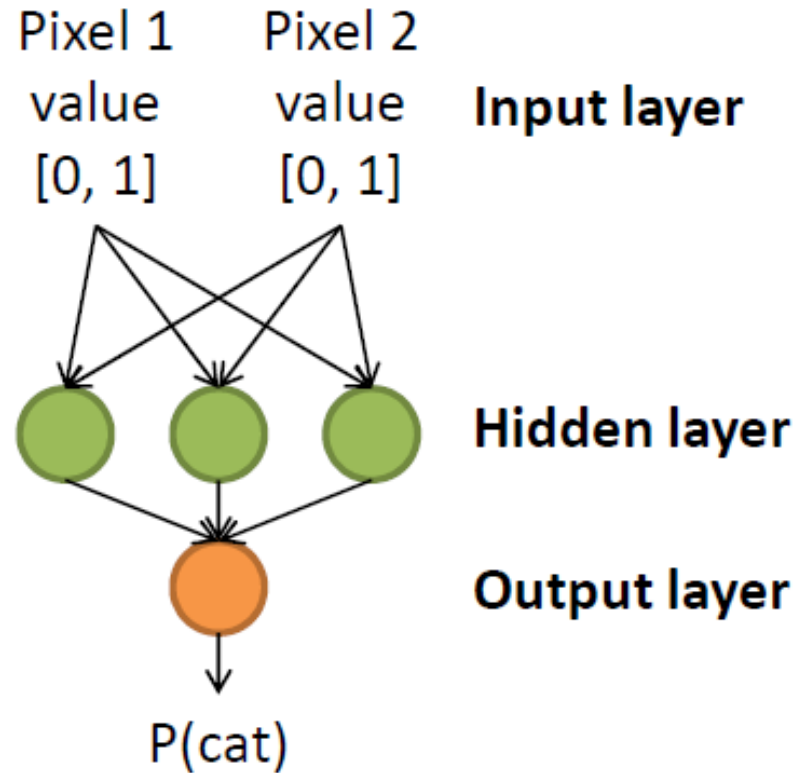
k-Nearest Neighbors (k-NN)



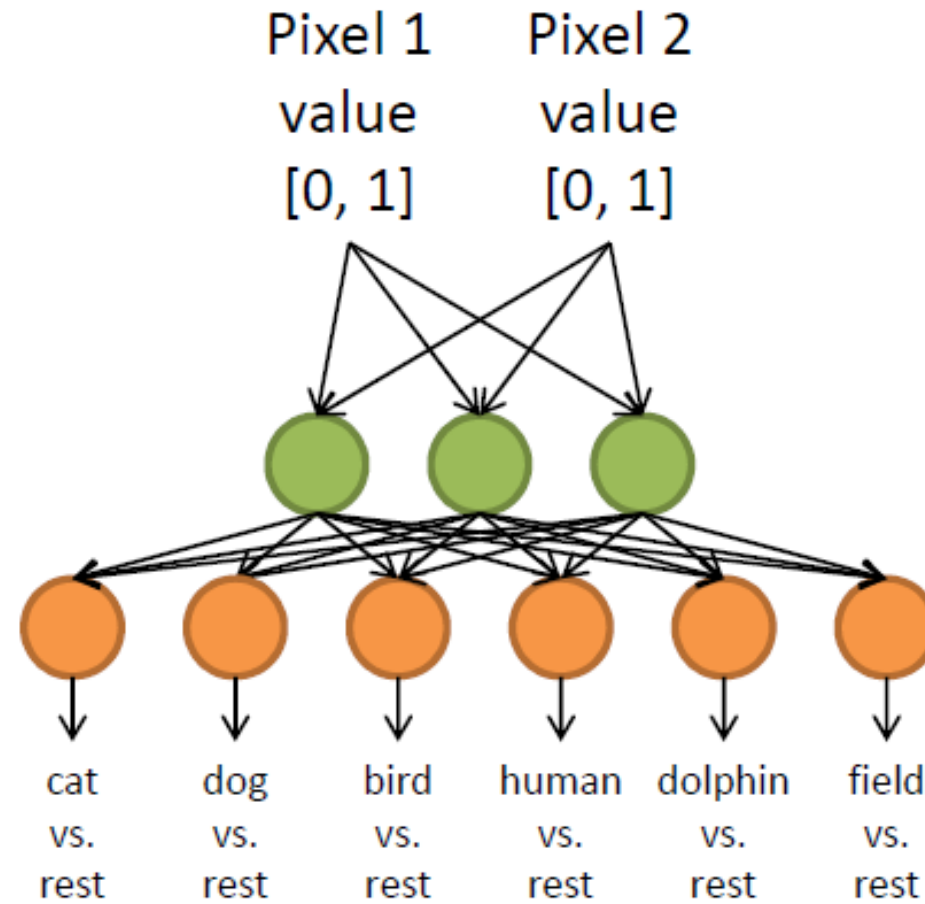
Support vector machine (SVM)



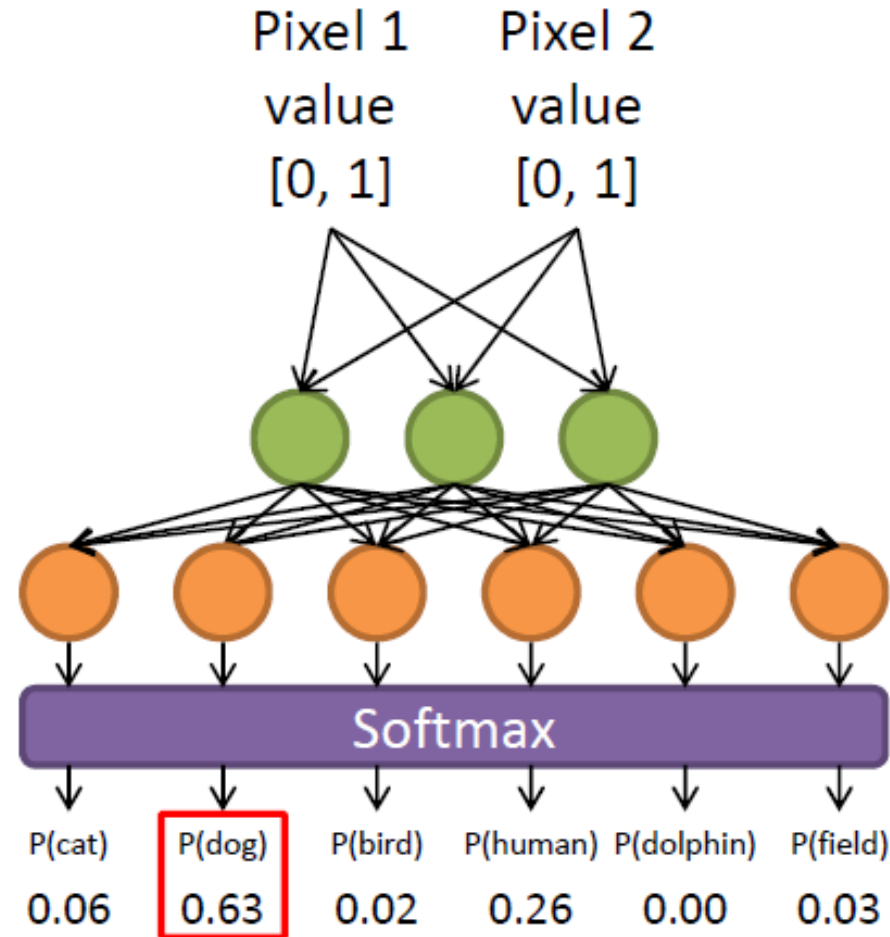
Neural Network



Neural Network



Neural Network



$$P(\text{cat}) + P(\text{dog}) + P(\text{bird}) + P(\text{human}) + P(\text{dolphin}) + P(\text{field}) = 1.0$$



Institución
Universitaria
Reacreditada en Alta Calidad

Challenge



Challenge: deformation



© 2021 Edge Impulse, Inc.

Challenge: Occlusion





Institución
Universitaria
Reacreditada en Alta Calidad

Outline

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

Raw Data

Example of sample x_0

x_{00}	x_{01}	x_{02}	x_{03}	x_{04}
x_{05}	x_{06}	x_{07}	x_{08}	x_{09}
x_{010}	x_{011}	x_{012}	x_{013}	x_{014}
x_{015}	x_{016}	x_{017}	x_{018}	x_{019}

$$x_{00} = 0.22$$

$$x_{01} = 0.23$$

$$x_{02} = 0.23$$

$$x_{03} = 0.26$$

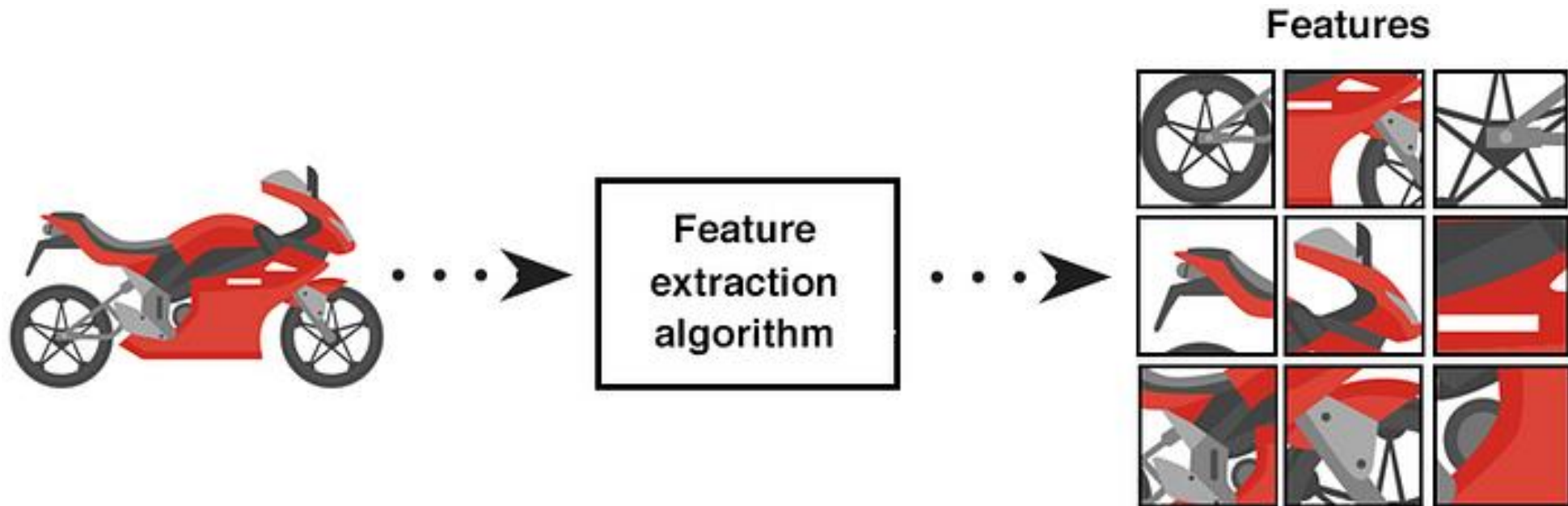
$$x_{04} = 0.32$$

$$x_{05} = 0.25$$

$$x_{06} = 0.26$$

⋮

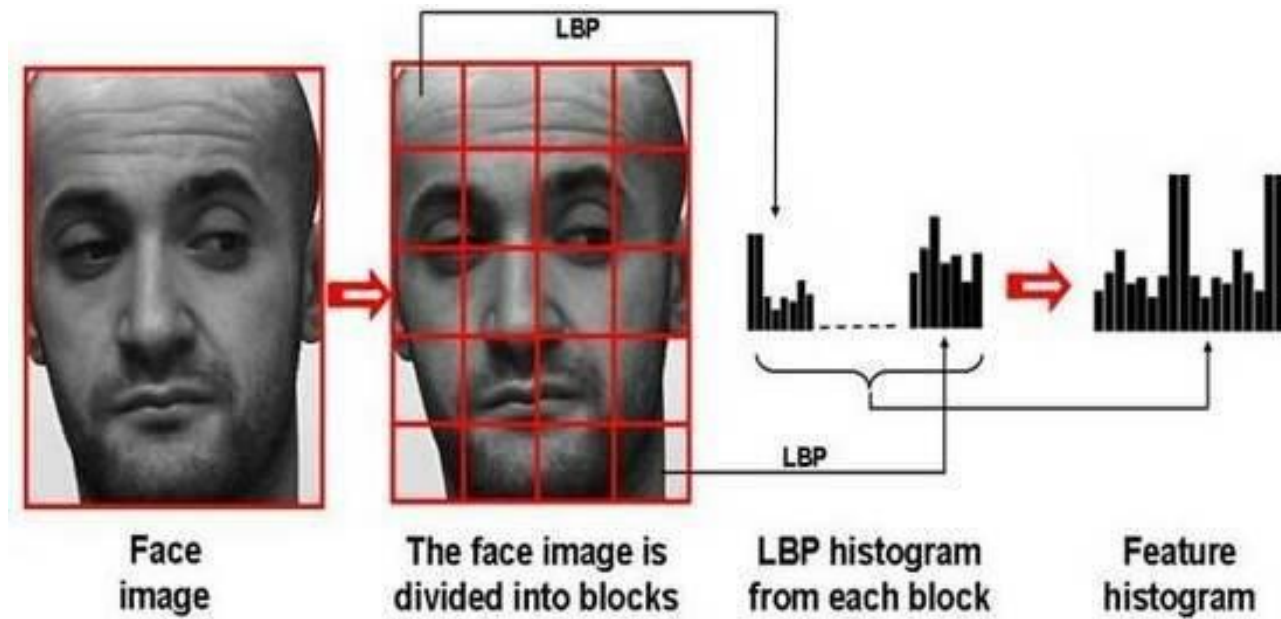
Feature extraction



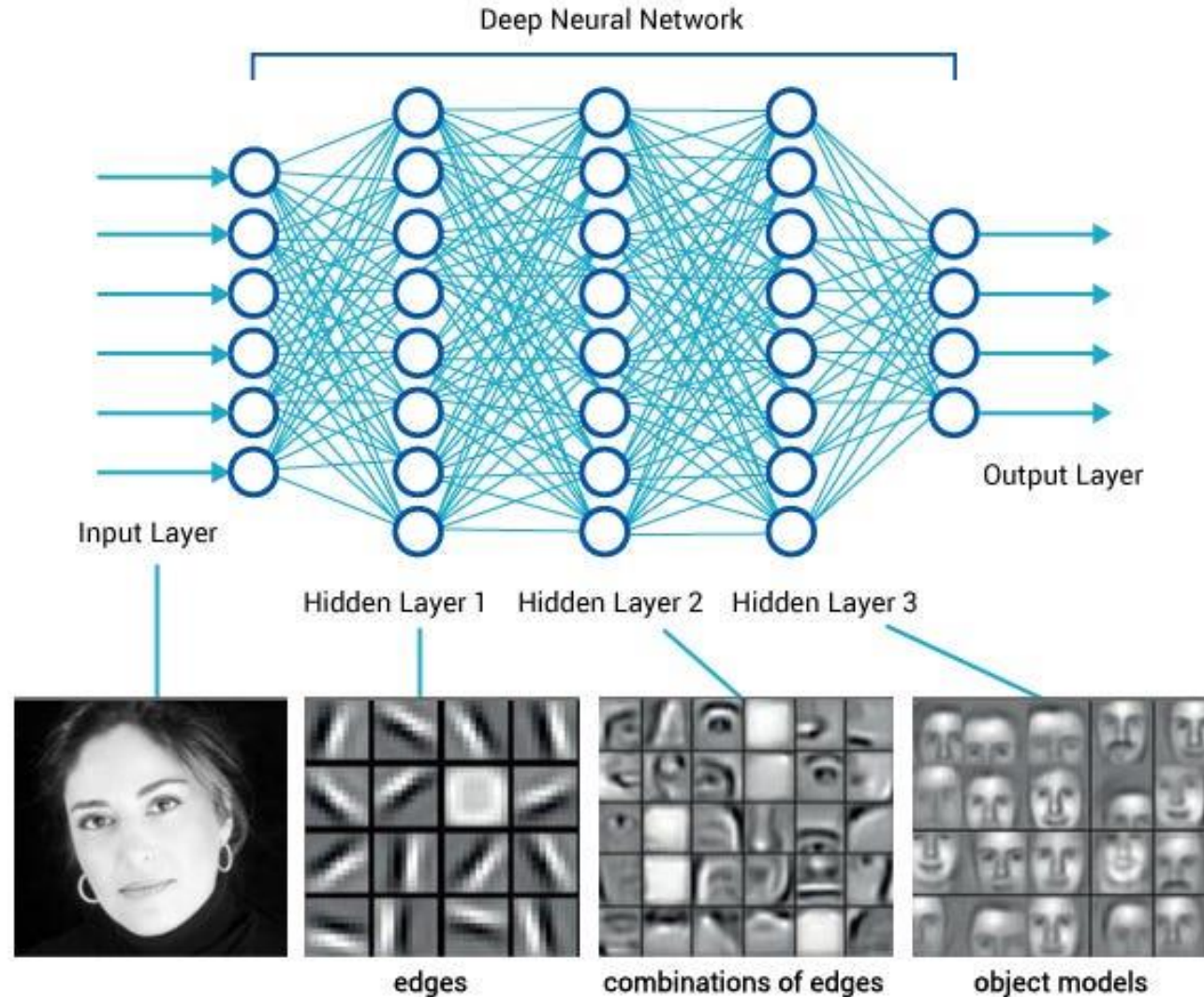
Feature Extraction Methods

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 channel.
4. Edge Features.
5. AutoEncoders.
6. Histogram of Oriented Gradients.
7. Scale-Invariant Feature Transform (SIFT).
8. Local Binary Patterns (LBP).
9. Frequency-based features.
10. Color based features.

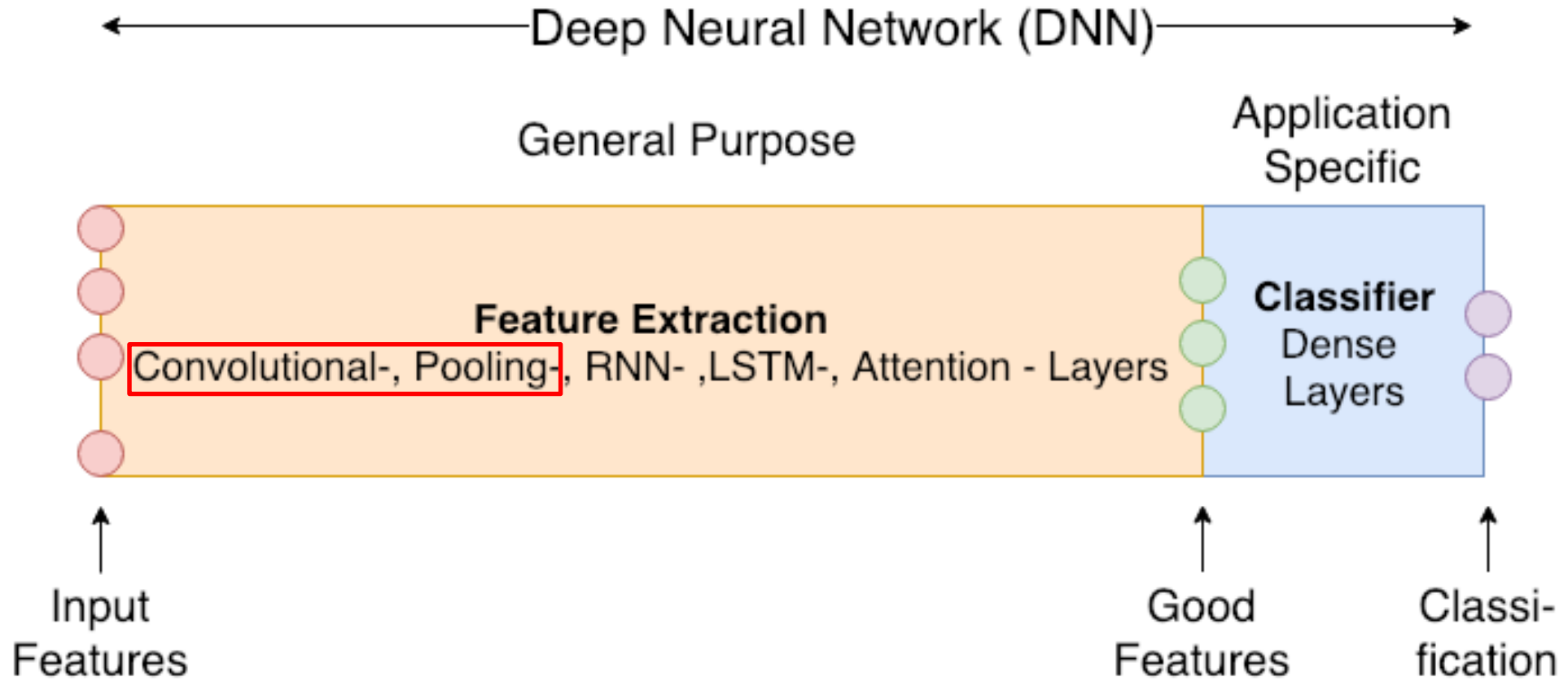
Features - LBP



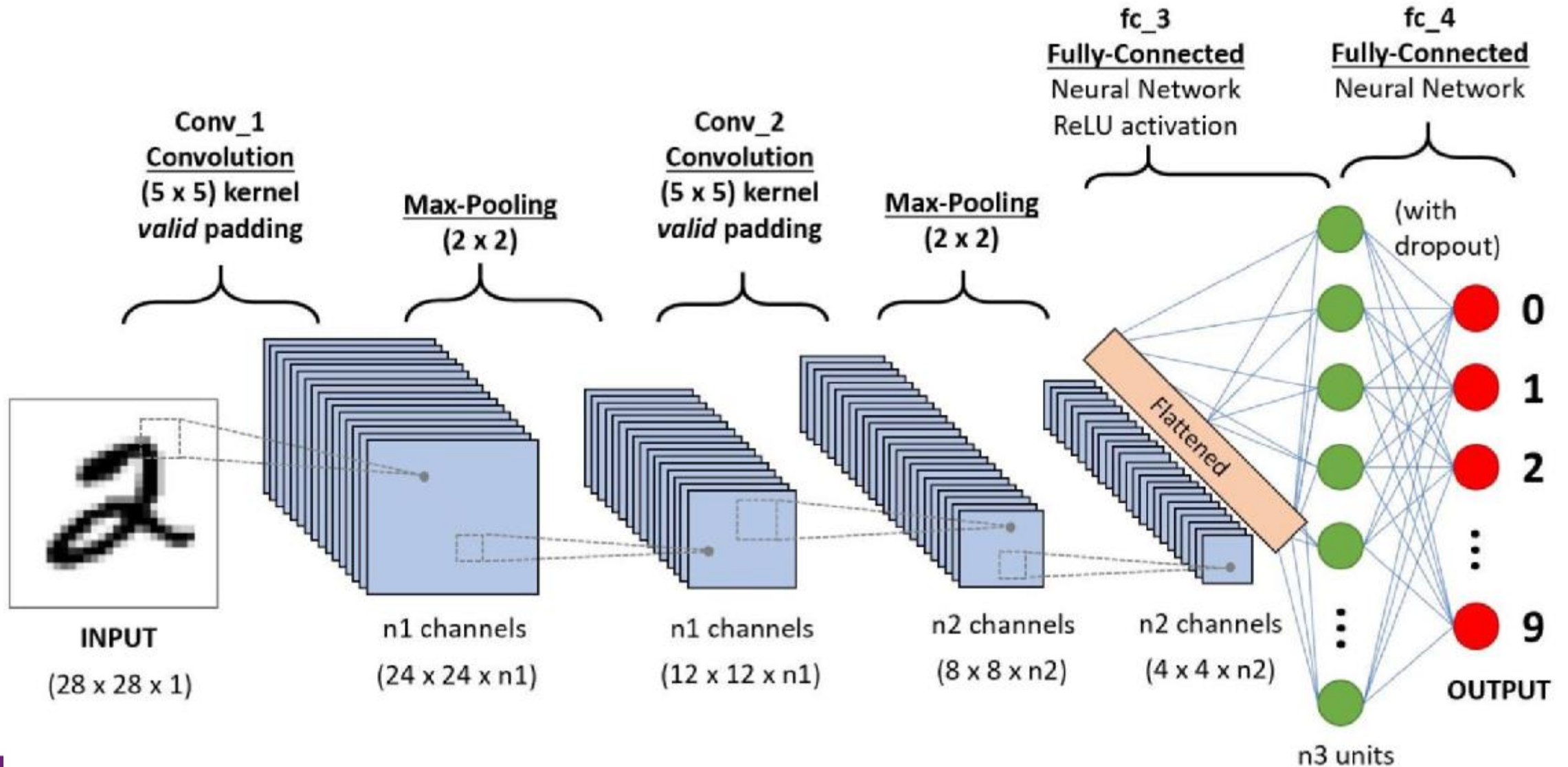
Deep learning



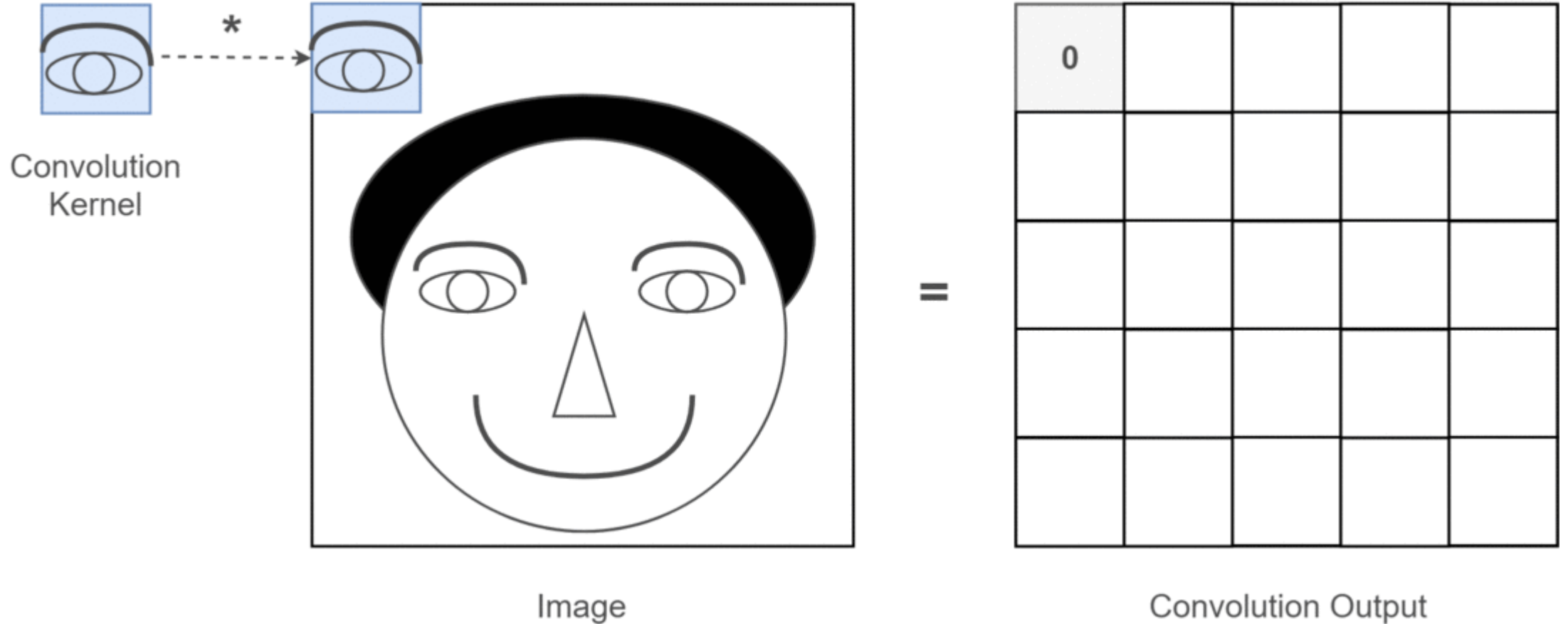
Deep Learning



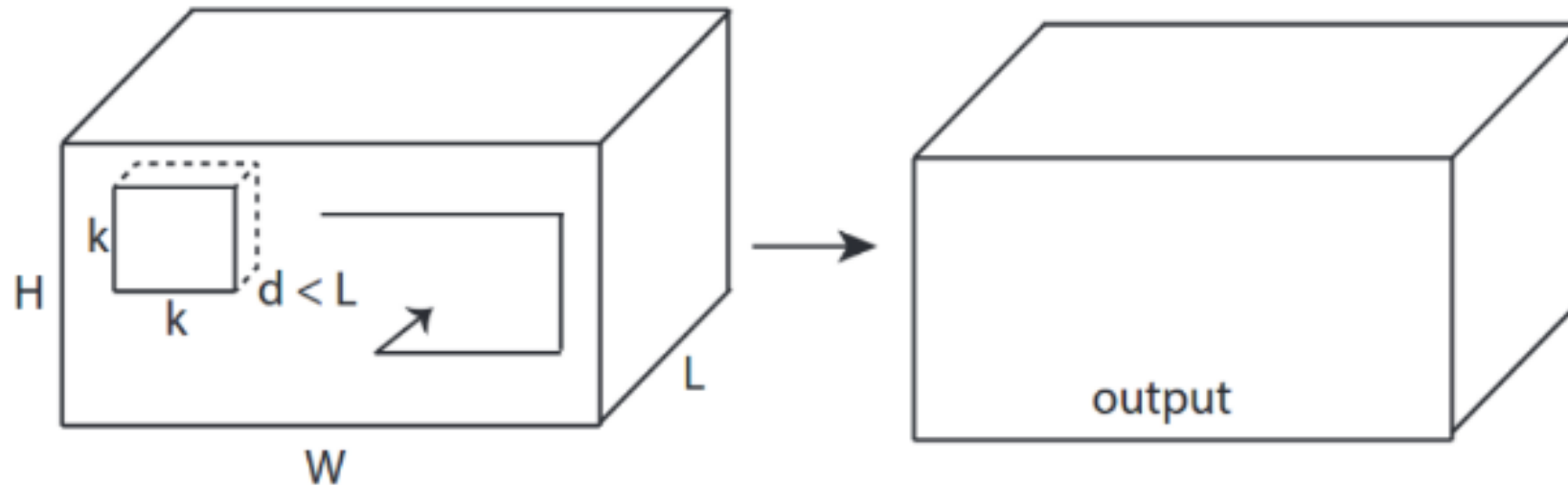
Convolutional Neural Network

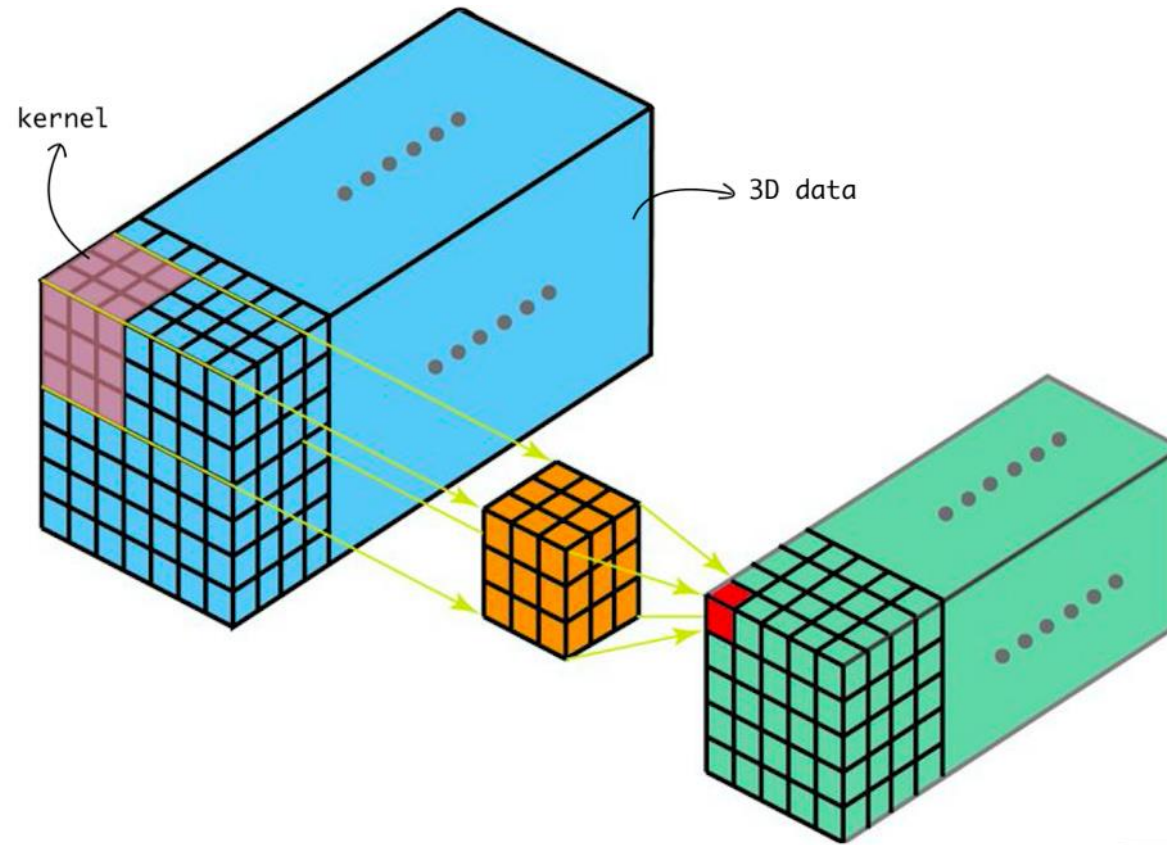


Convolution 2D

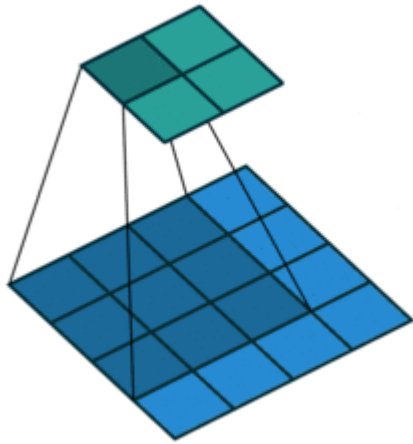


3D CNN

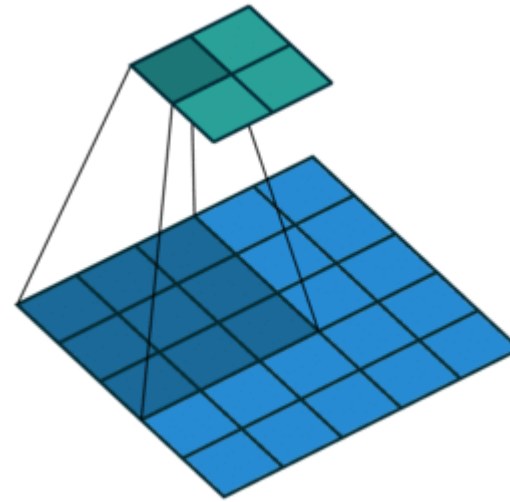




Convolution 2D

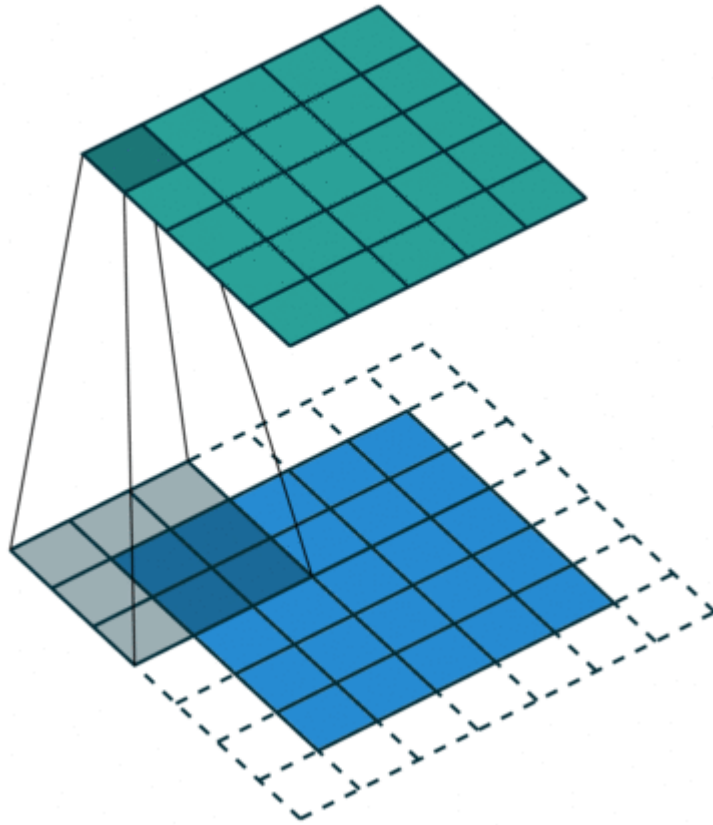


padding = 0, stride = 1

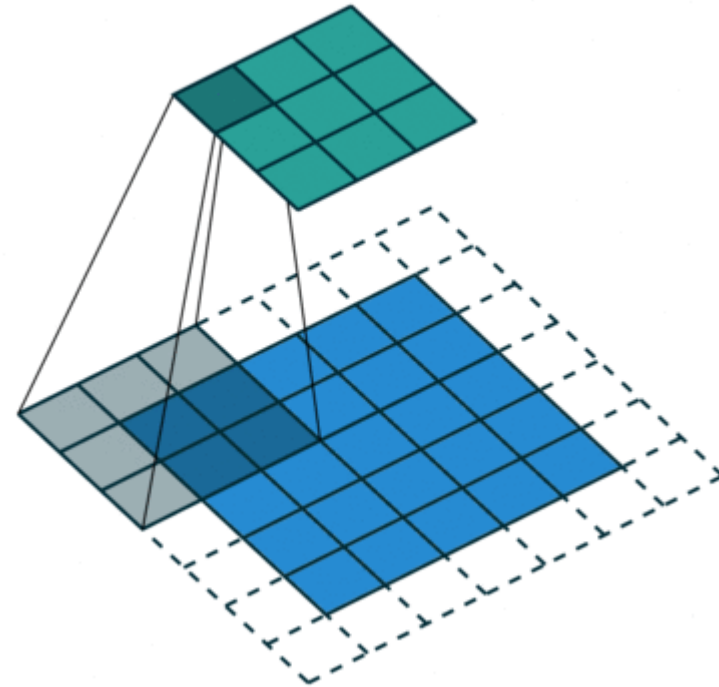


padding = 0, stride = 2

Convolution 2D

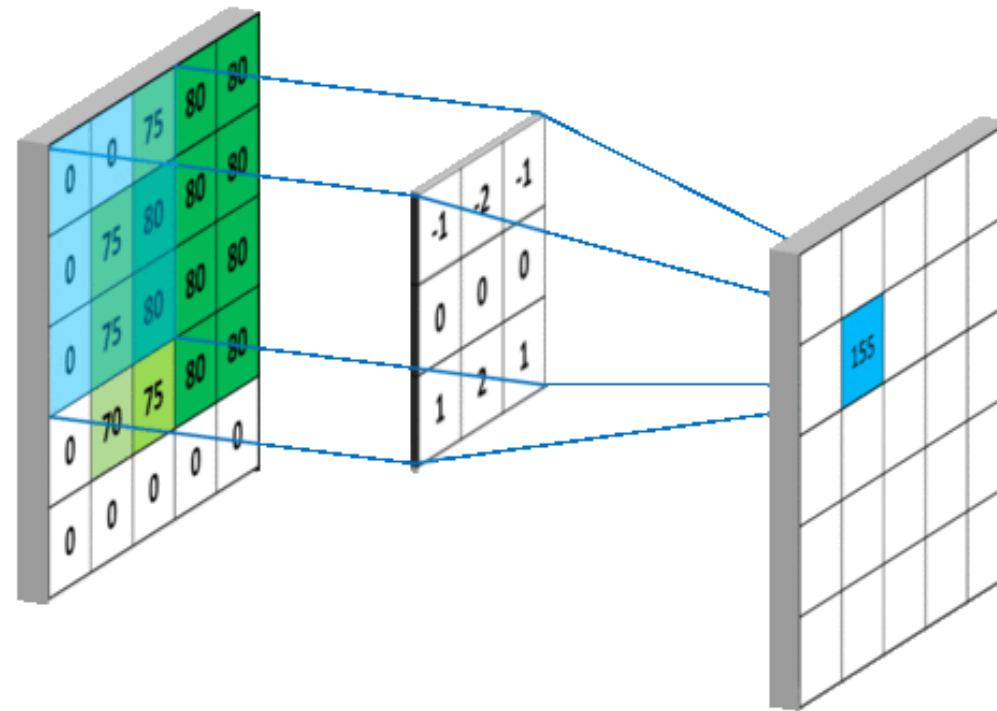


padding = 1, stride = 1



padding = 1, stride = 2

Convolutional layer - example



Exercise

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

29	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

2 x 2
pool size

100	184
12	45

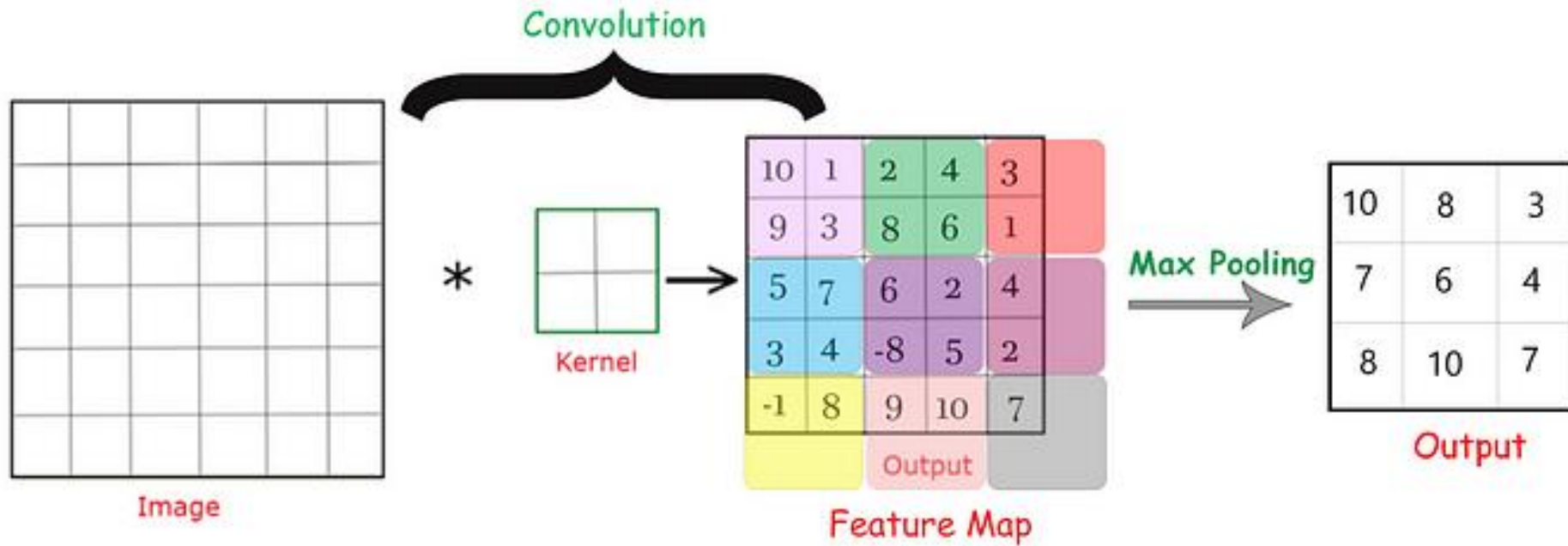
Average Pooling

31	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

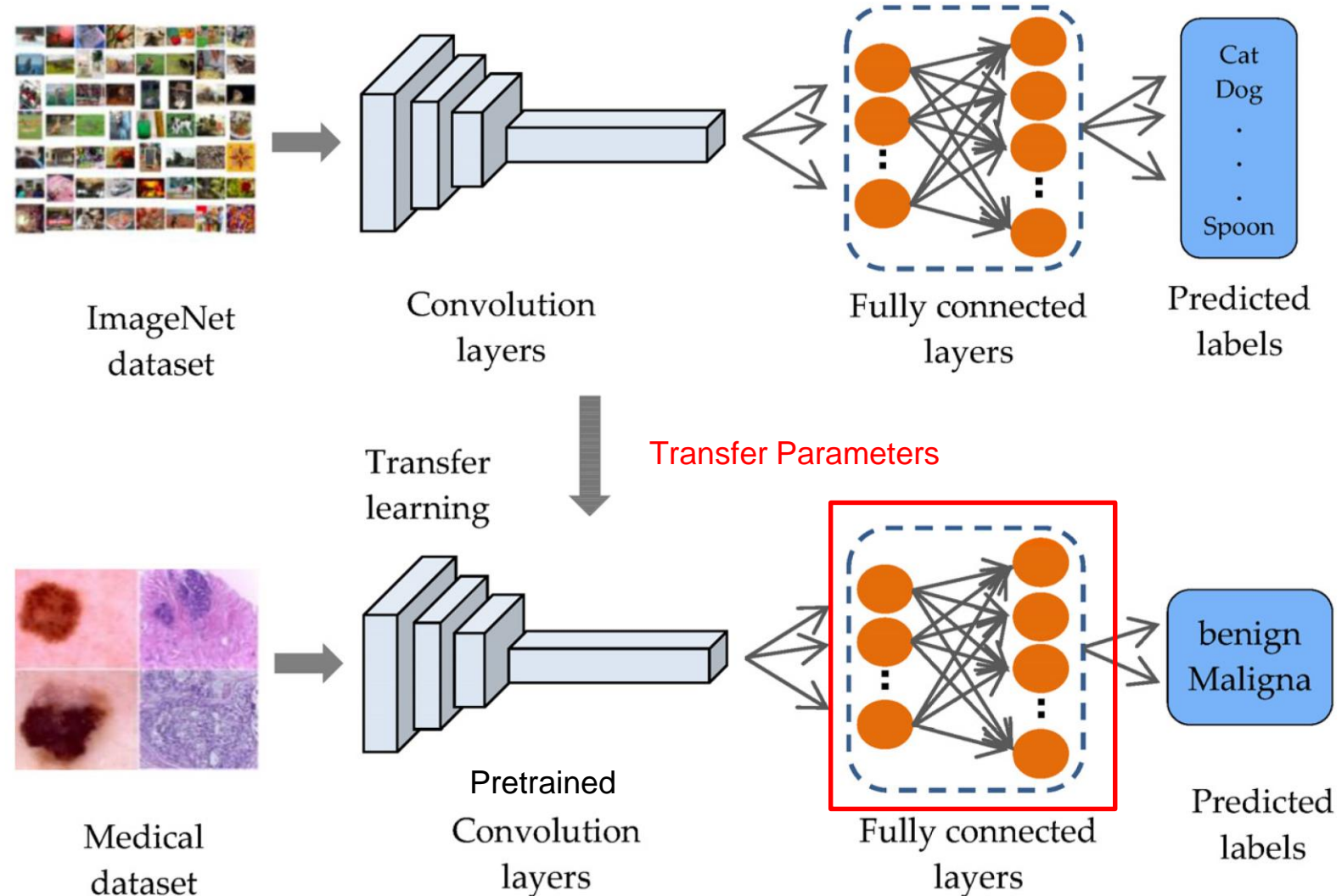
2 x 2
pool size

36	80
12	15

Max Pooling – Example



Transfer Learning



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---VmldzoxOTk2MDA>

<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