Introduction to Computer Vision

1

Introduction

Presentation of myself, Outline and Subject

Welcome!

I am Tom Genlis

EPITA (2024) ING2

SCIA Major - CV Speciality

tom.genlis@gmail.com

SCIA Major

AI & Data Science Major at EPITA

- Analysis & Optimization Maths
- Deep Learning
- CV & NLP
- Basics to state-of-the-art algorithms and models
- Data Engineering

Presentation Outline

- Quick Intro to Machine Learning & ML Concepts
- General Conference over Computer Vision
- Available Workshop (Learn)
- Kaggle Contest (Practice)

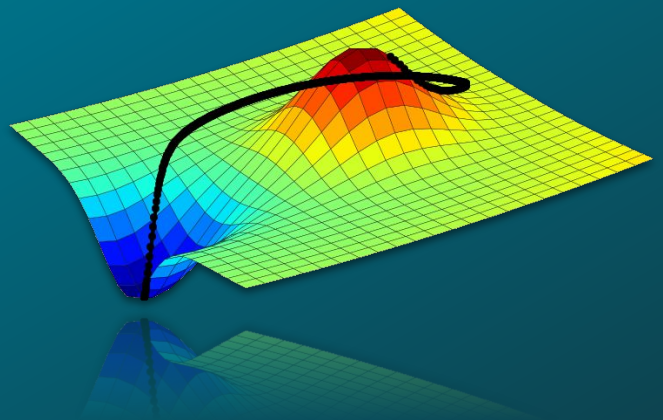
Feel free to ask any questions regarding this presentation !

Info Links

github.com/GDSC-EPITA/gdsc-computer-vision-workshop

discord.gg/3kbcZRFpbm

2



Quick Intro to Machine Learning

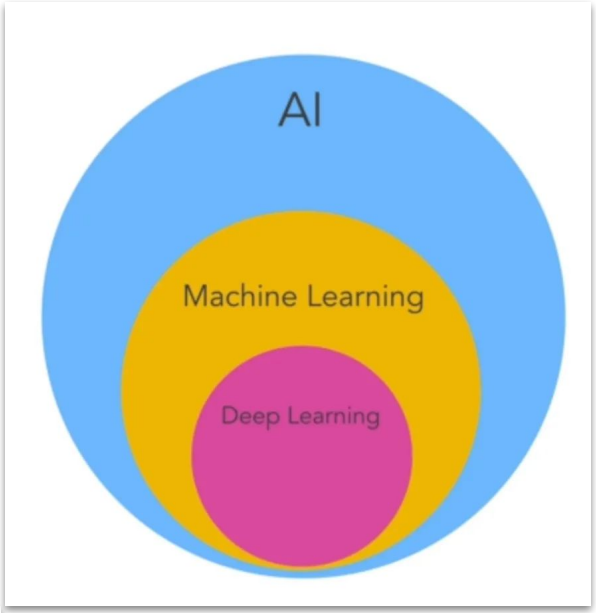
Definitions

Introduction Outline

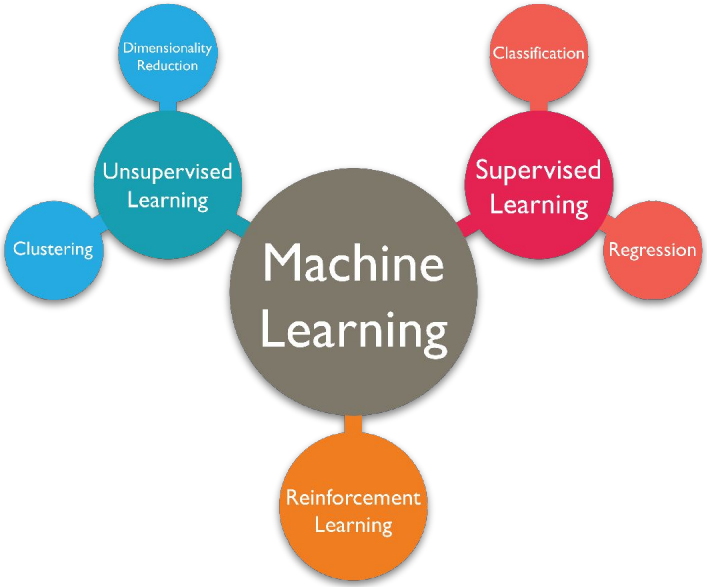
- What is AI ?
- What is Machine Learning ?
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- Deep Learning
- 8 Core ML Concepts

What is AI ?

- Mimic Human Behavior

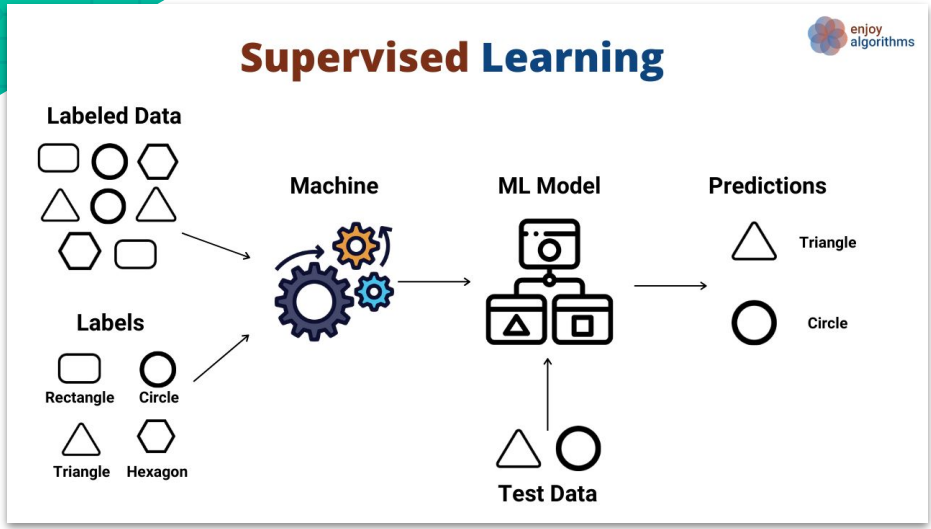


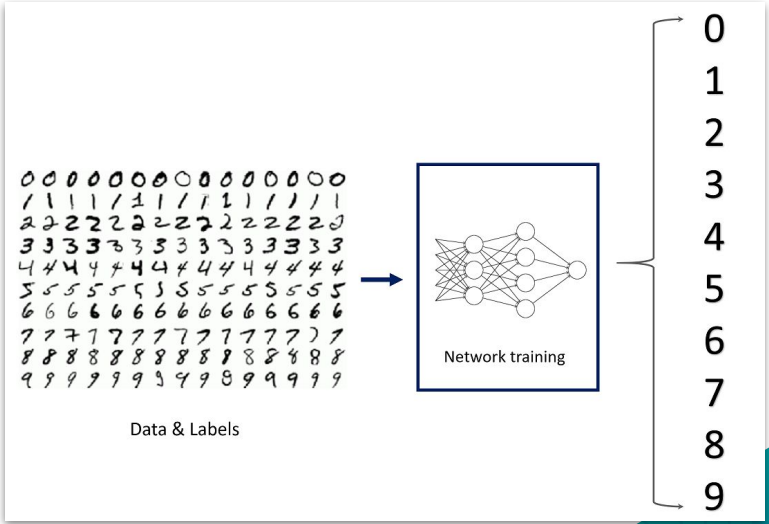
What is Machine Learning ?



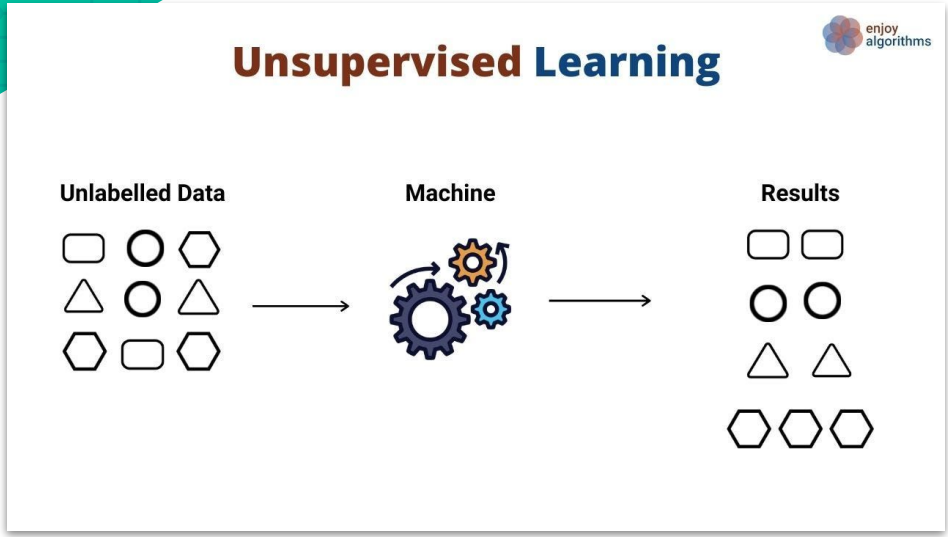
Supervised Learning

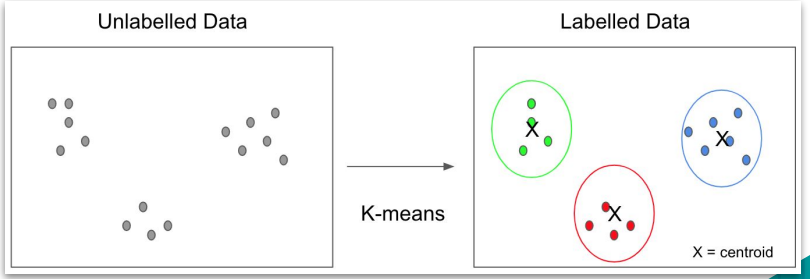
Value	Label
🍎	0
🍌	1
🍎	0
🍌	1
🍎	0



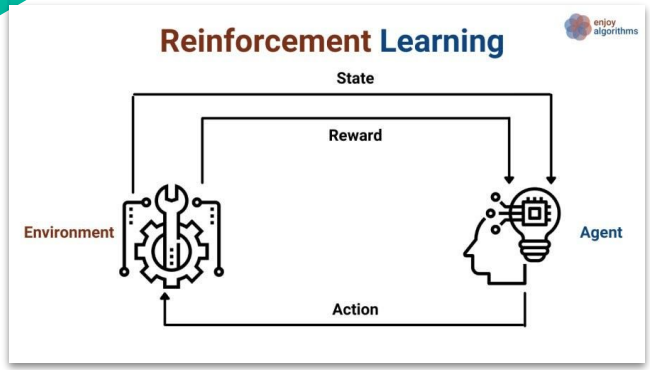


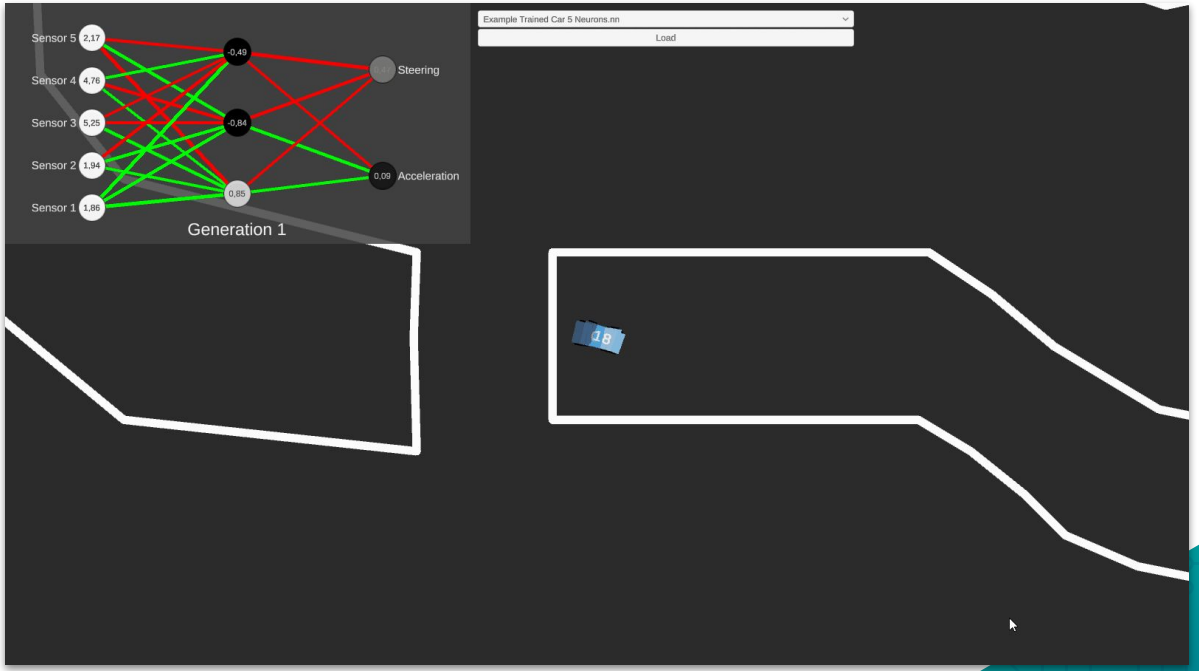
Unsupervised Learning



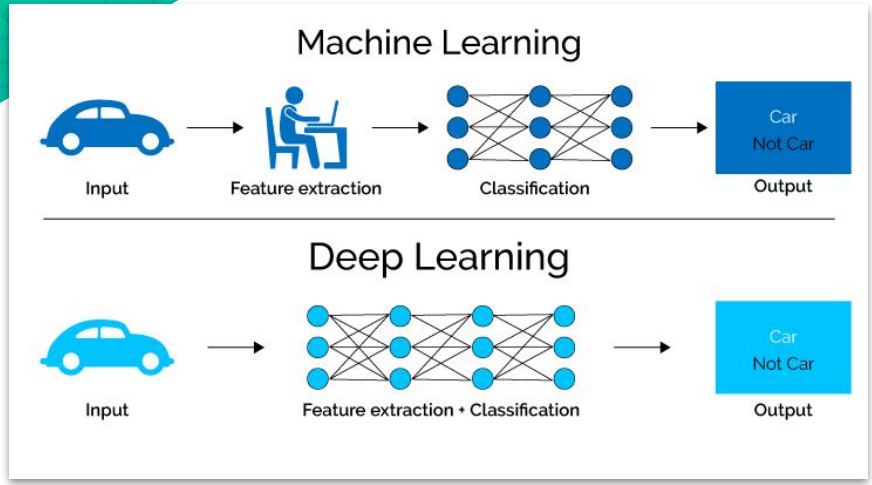


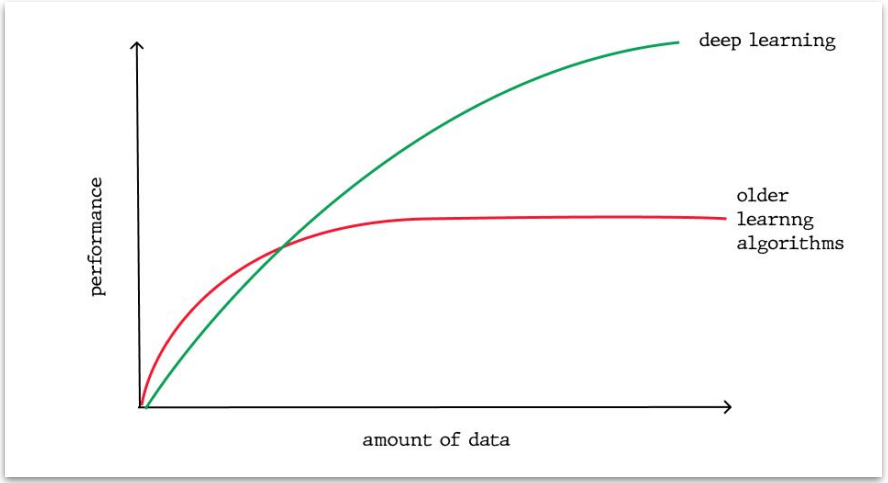
Reinforcement Learning





Deep Learning

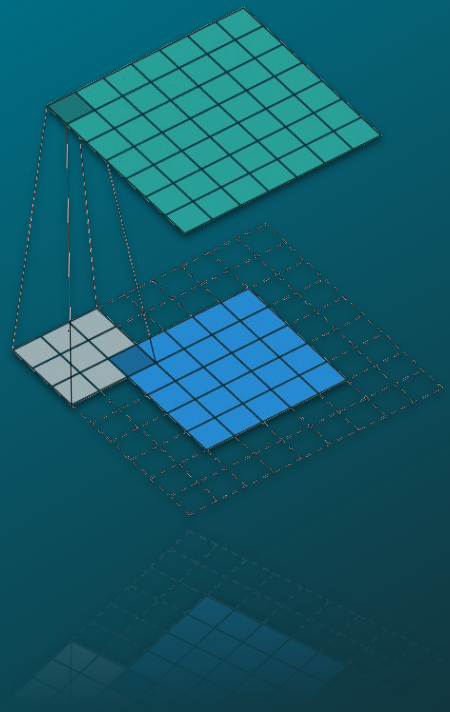




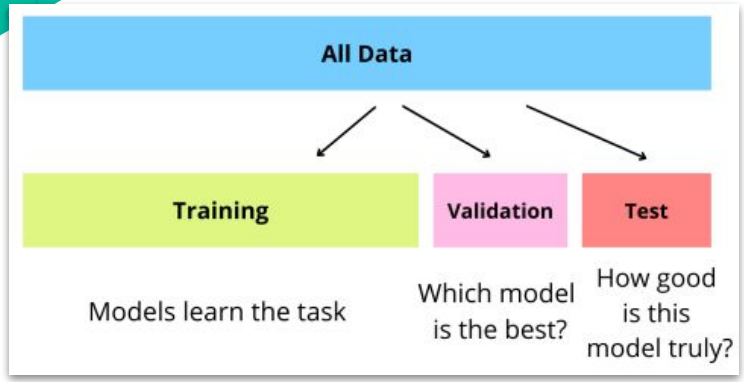
3

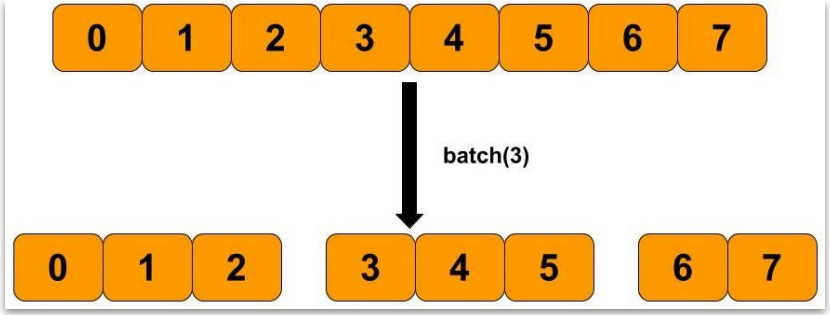
Core ML Concepts

Definitions, Examples



ML Concepts : Datasets & Batching





ML Concepts : Convolutions

Convolution

1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

Result

4		

- Kernel size (in ex :(3, 3))
- Strides (in ex : (1, 1))
- Padding (in ex : none)

- Spatial Context

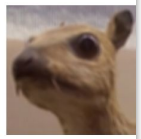
Kernels

Identity $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Sharpen $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$

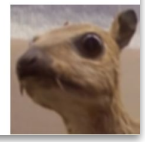
$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

Average Blur $\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$



Edge Detection $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$

Gaussian Blur $\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$



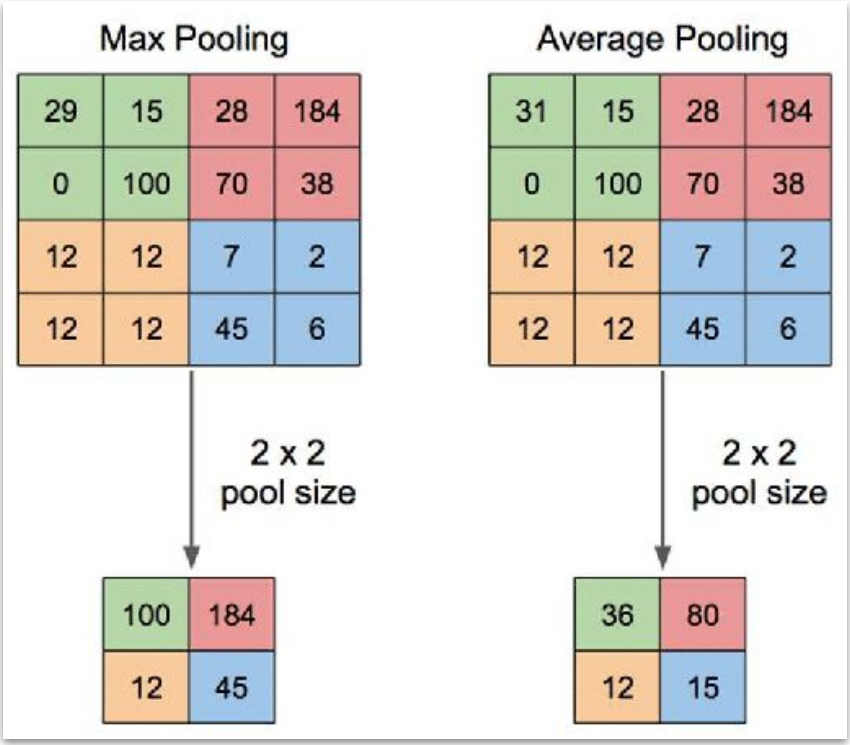
0	0	0	0	0	0	0
0	60	113	56	139	85	0
0	73	121	54	84	128	0
0	131	99	70	129	127	0
0	80	57	115	69	134	0
0	104	126	123	95	130	0
0	0	0	0	0	0	0

Kernel

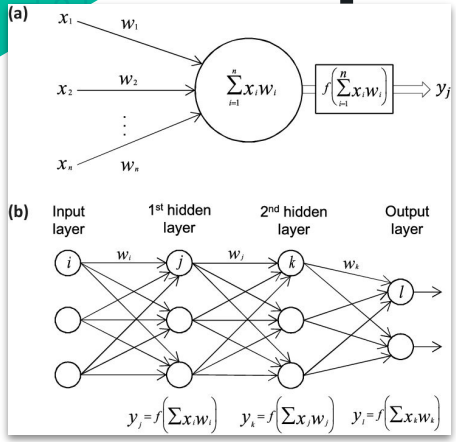
0	-1	0
-1	5	-1
0	-1	0

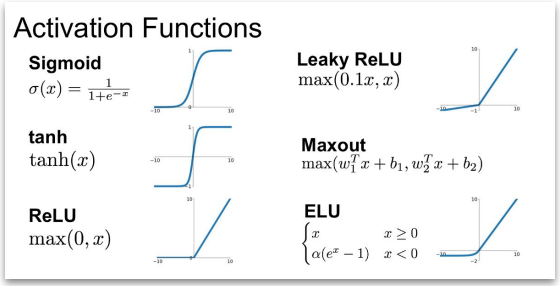
114				

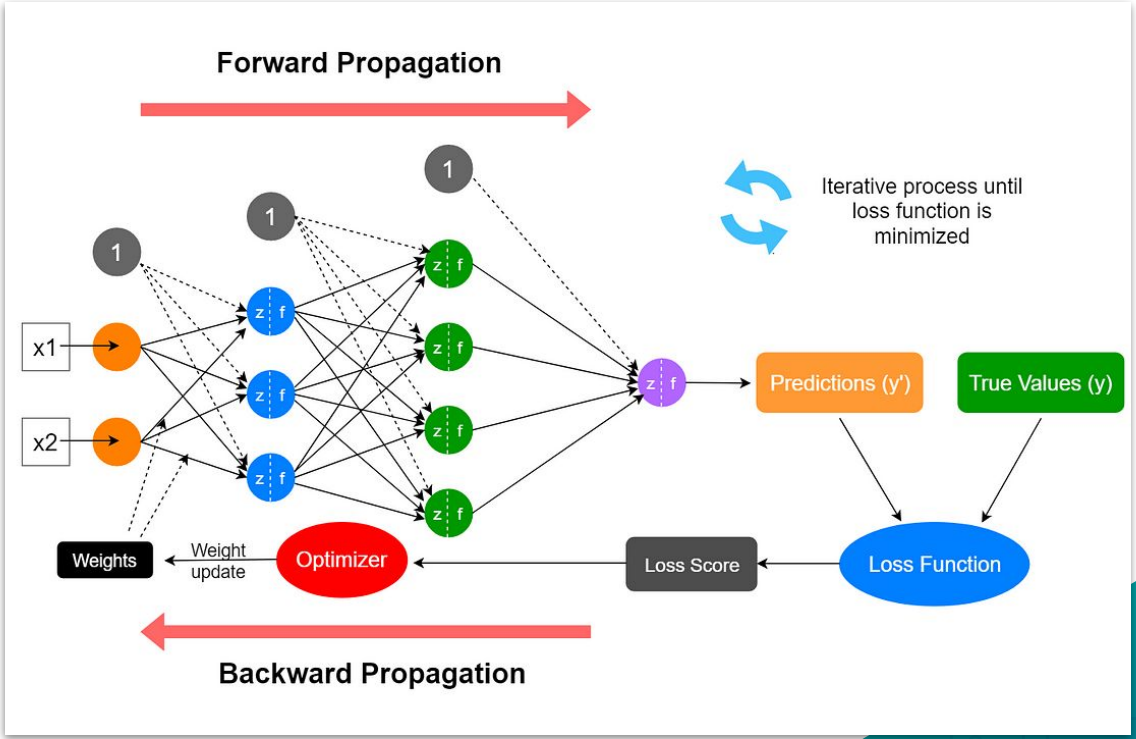
ML Concepts : Pooling



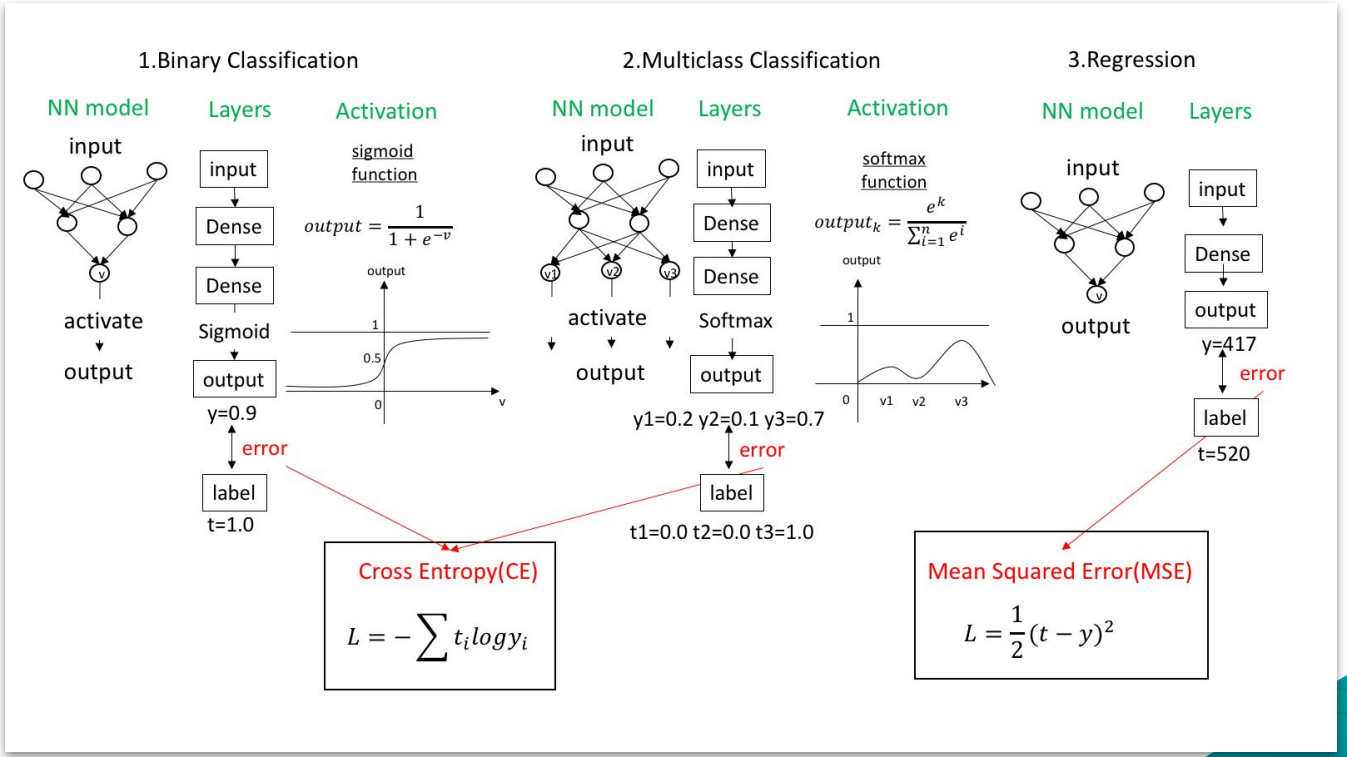
ML Concepts : Neural Networks



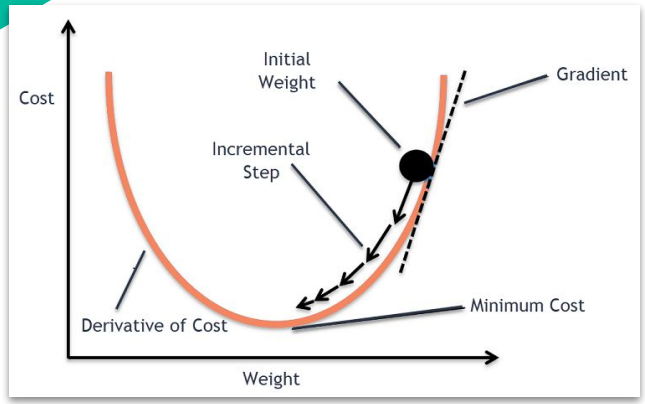


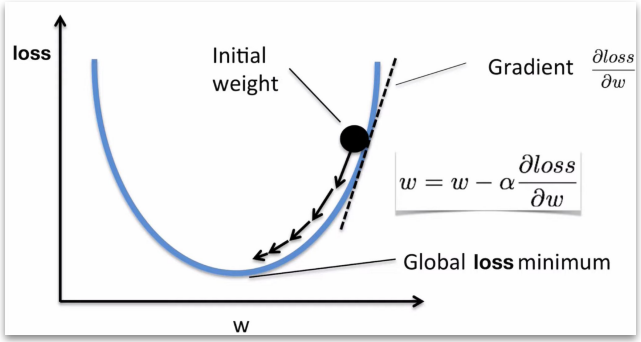


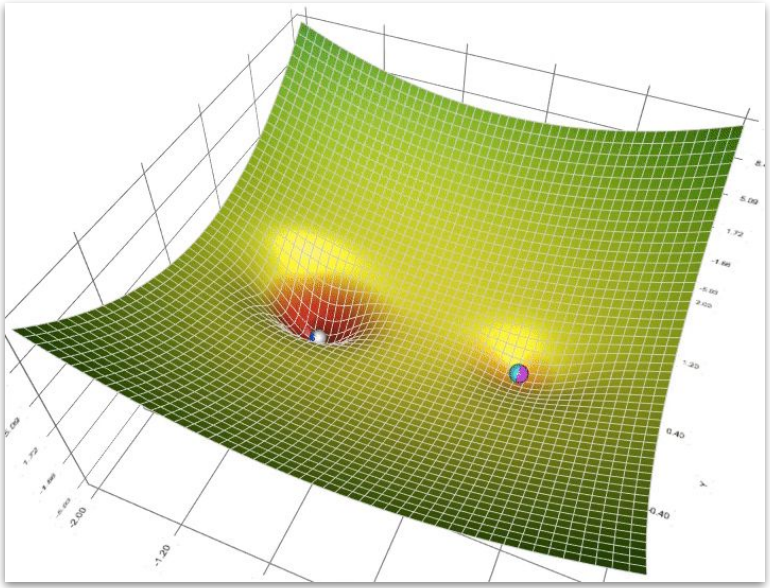
ML Concepts : Loss Function



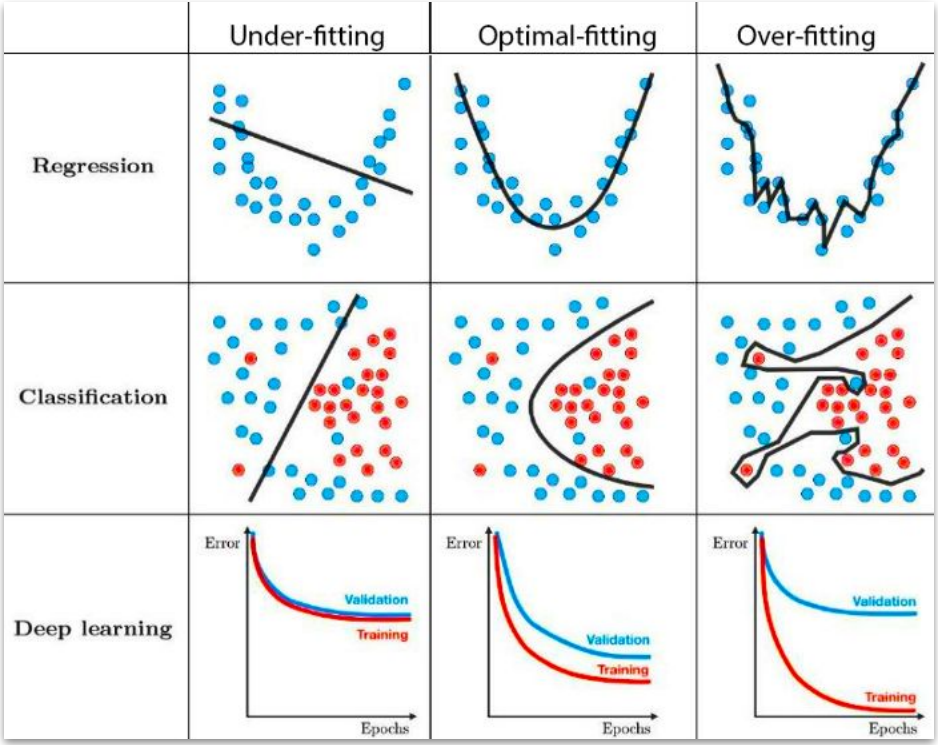
ML Concepts : Gradient Descent



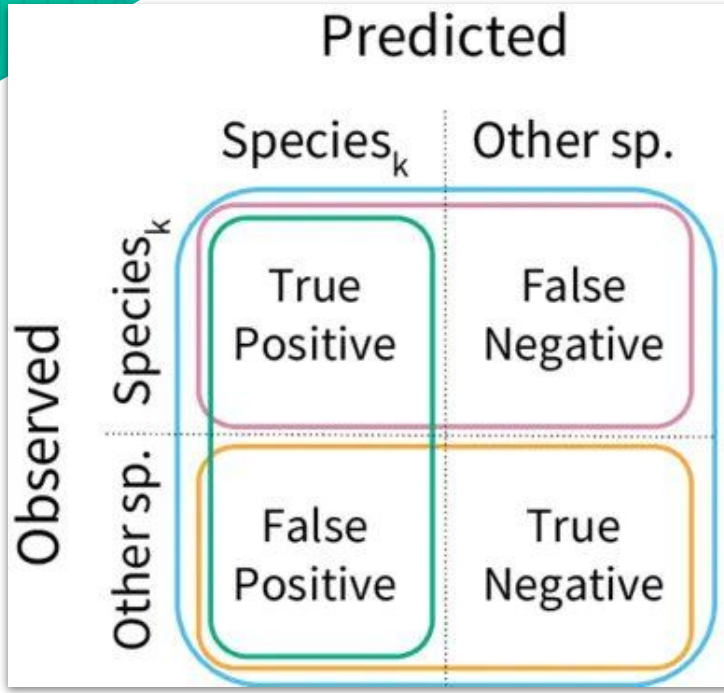




ML Concepts : Overfitting & Underfitting



ML Concepts : Metrics



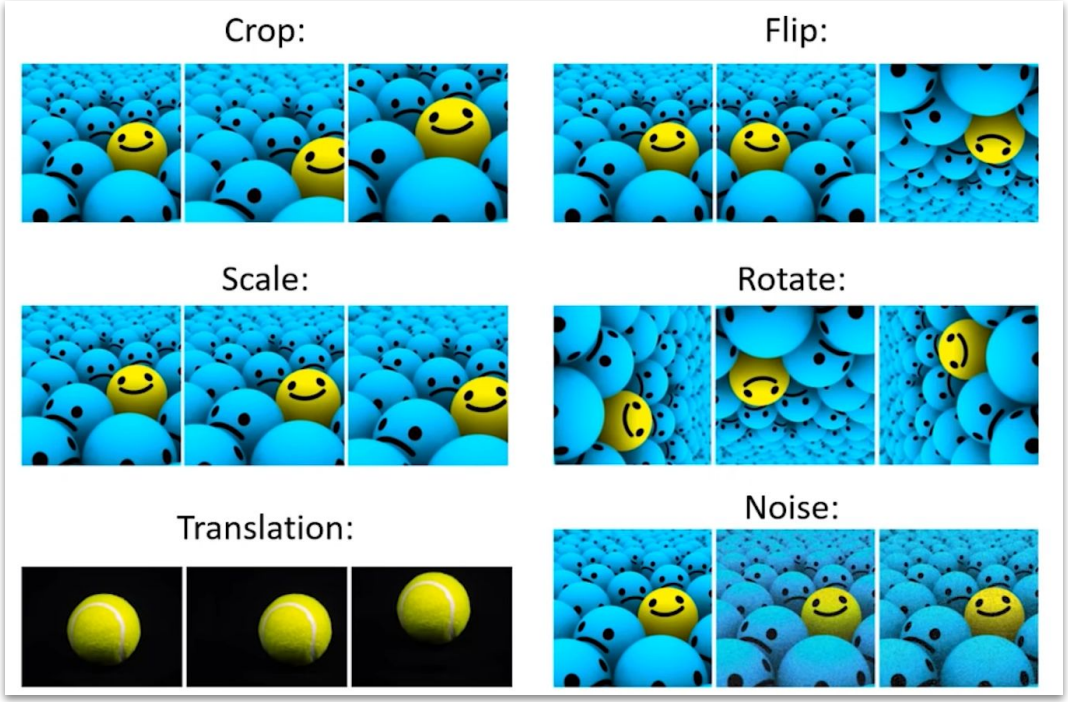
Accuracy $= \frac{TP + TN}{TP + TN + FP + FN}$

Specificity $= \frac{TN}{TN + FP}$

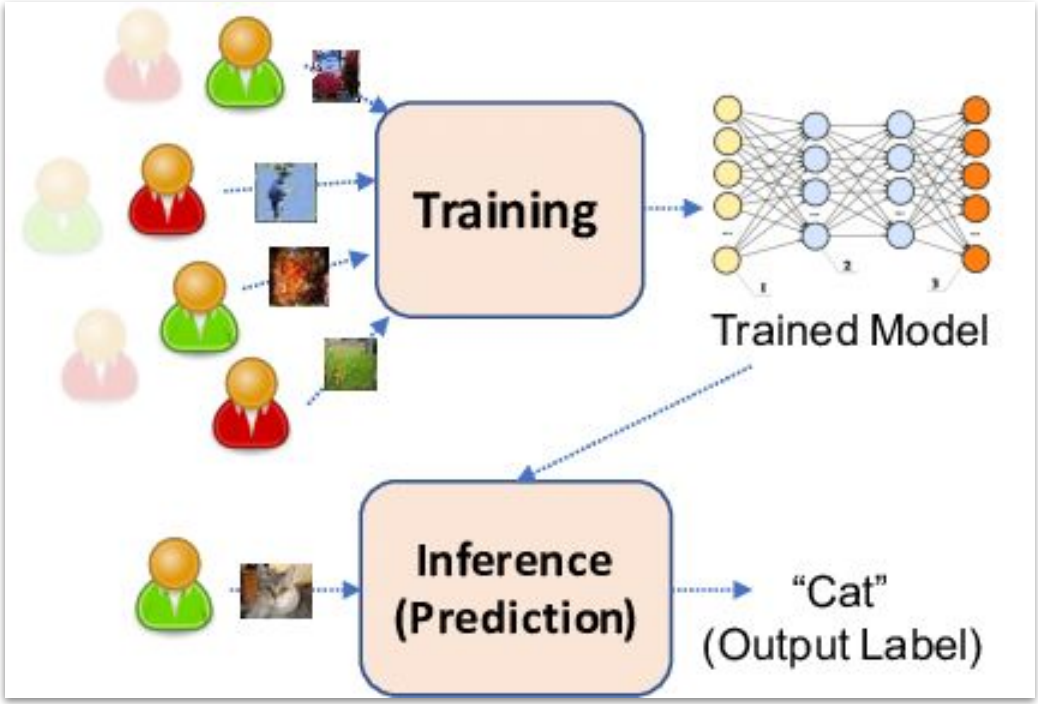
Precision $= \frac{TP}{TP + FP}$

Recall $= \frac{TP}{TP + FN}$

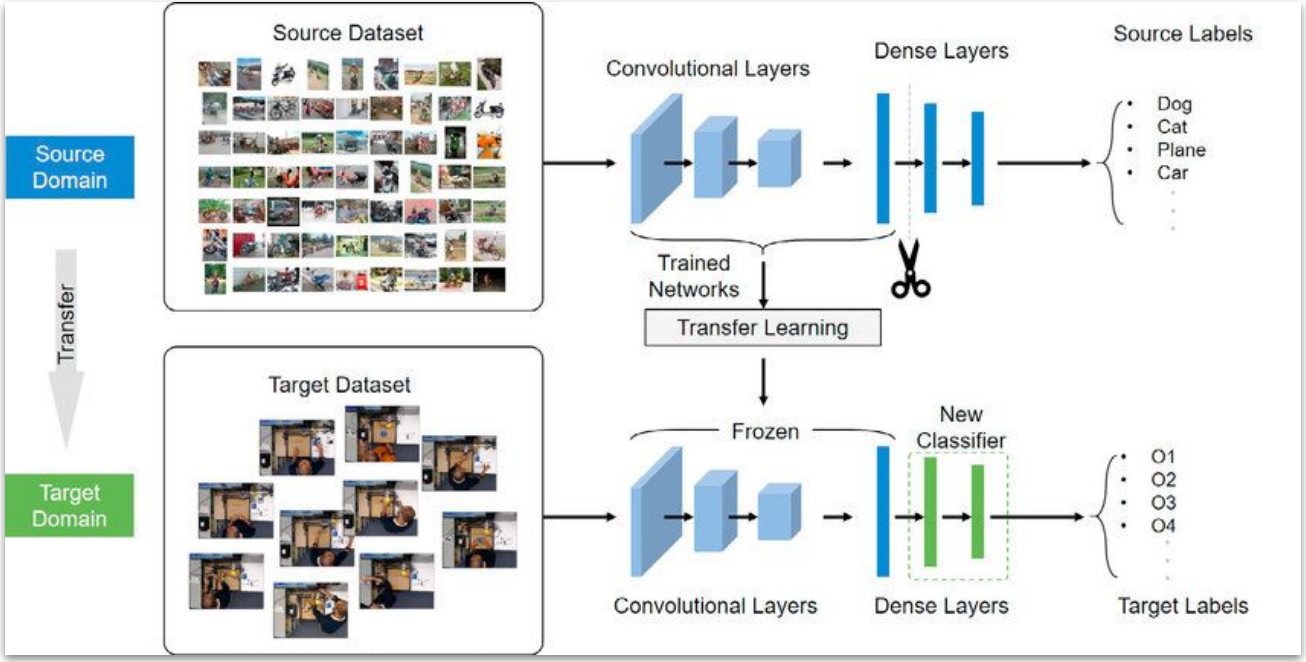
ML Concepts : Data Augmentation



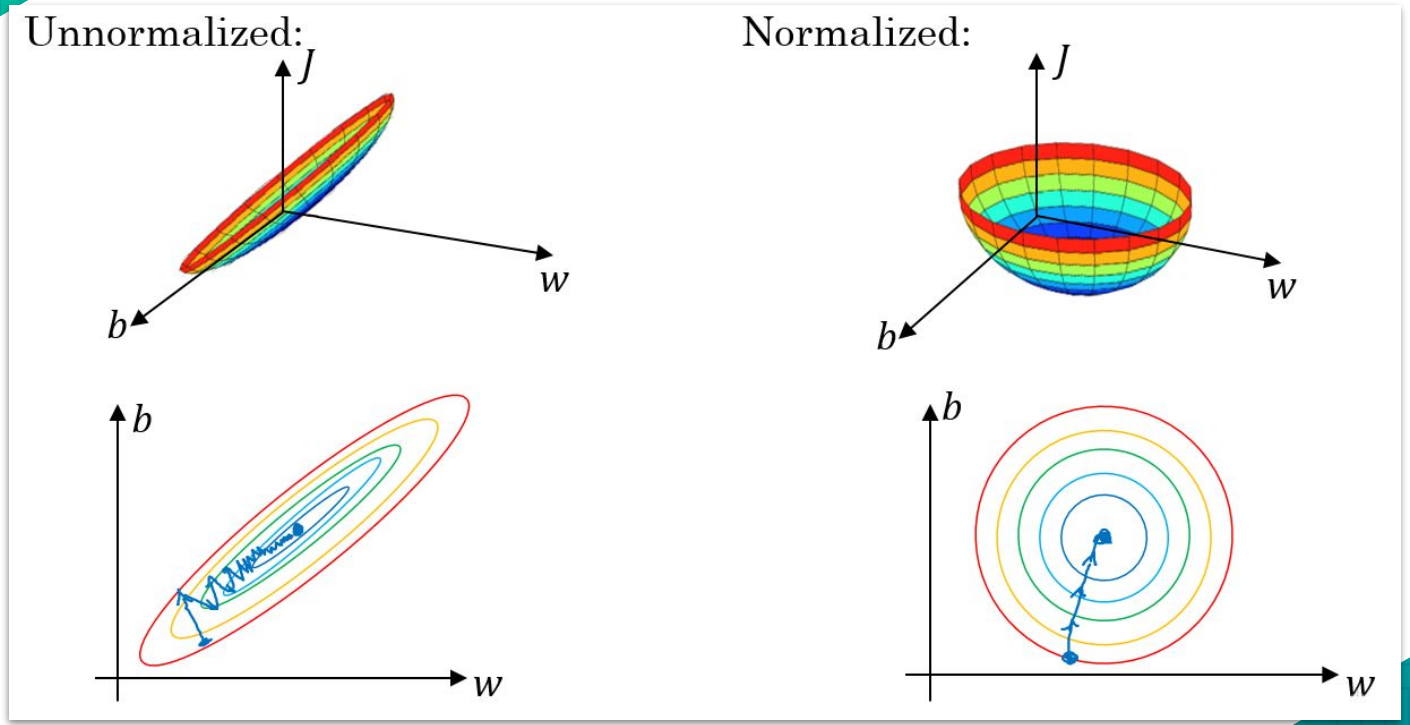
ML Concepts : Post Training Usage



ML Concepts : Transfer Learning



ML Concepts : Normalization



4

What is Computer Vision ?

Definition, Applications and Progress

Exploration Outline

- History Of Computer Vision
- Computer Vision Subfields
- Classification
- Localization
- Object Detection
- Segmentation

Computer Vision

Subfield of AI

History of Computer Vision

1963 – Larry Roberts, the father of CV, described the process of deriving 3D info about solid objects from 2D photographs.

1980 – Kunihiko Fukushima built the 'neocognitron', the precursor of modern Convolutional Neural Networks.

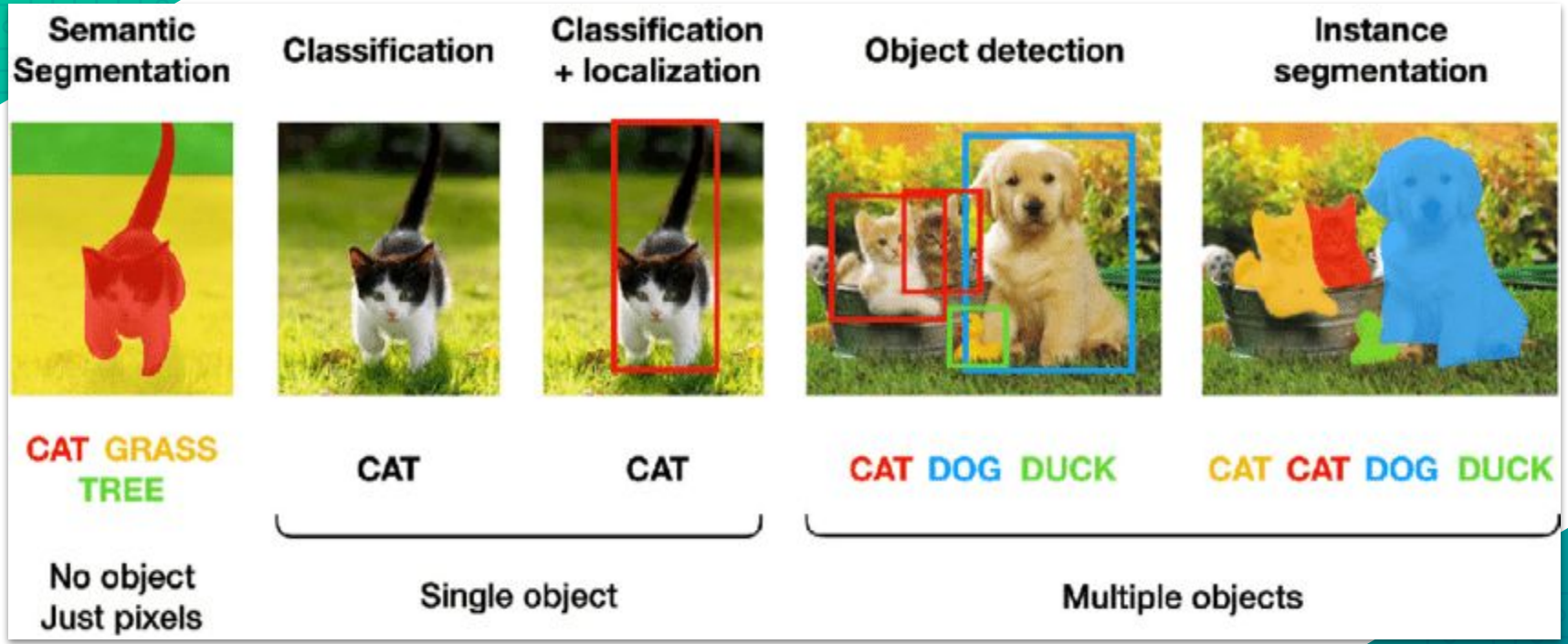
2001 – Two researchers at MIT introduced the first face detection framework (Viola-Jones) that works in real-time.

2012 – Google Brain's neural network recognized pictures of cats using a deep learning algorithm.

2017 – Apple released the iPhone X in 2017, advertising face recognition as one of its primary new features.

2018 – Alibaba's AI model scored better than humans in a Stanford University reading and comprehension test.

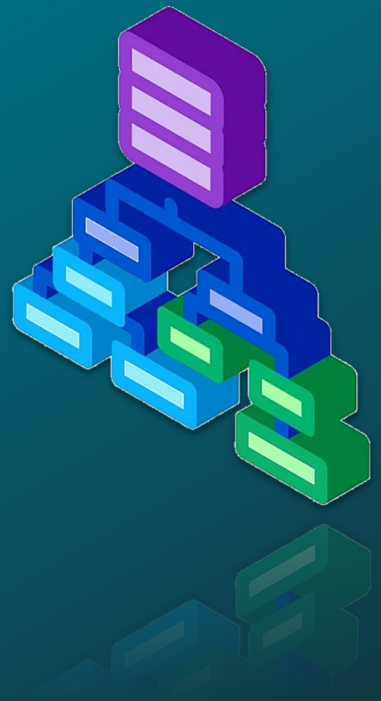
Computer Vision Subfields



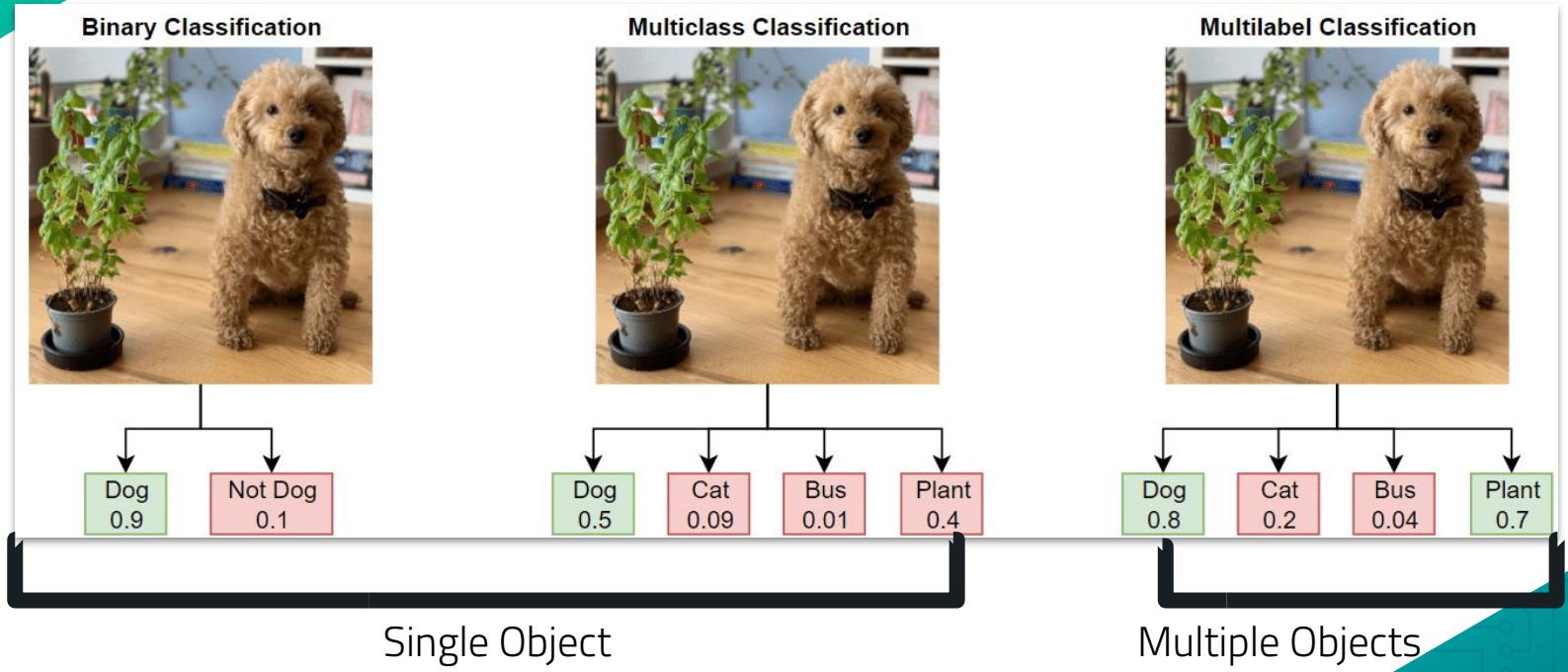
5

Computer Vision : Classification

Definition, Applications and Examples



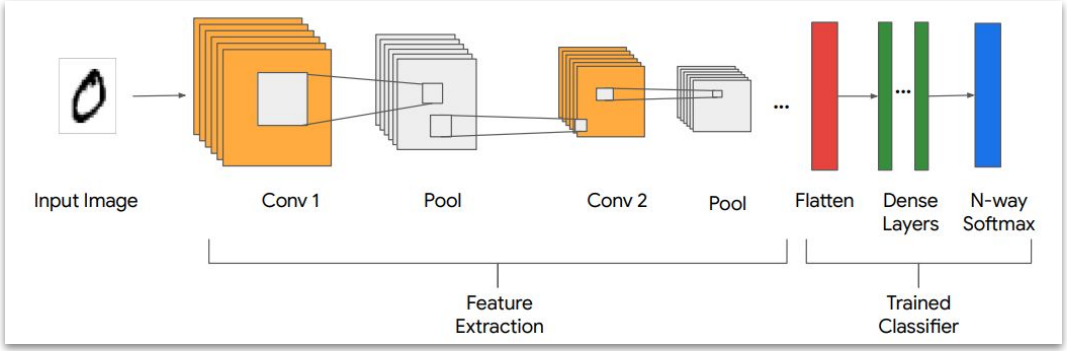
Classification Types

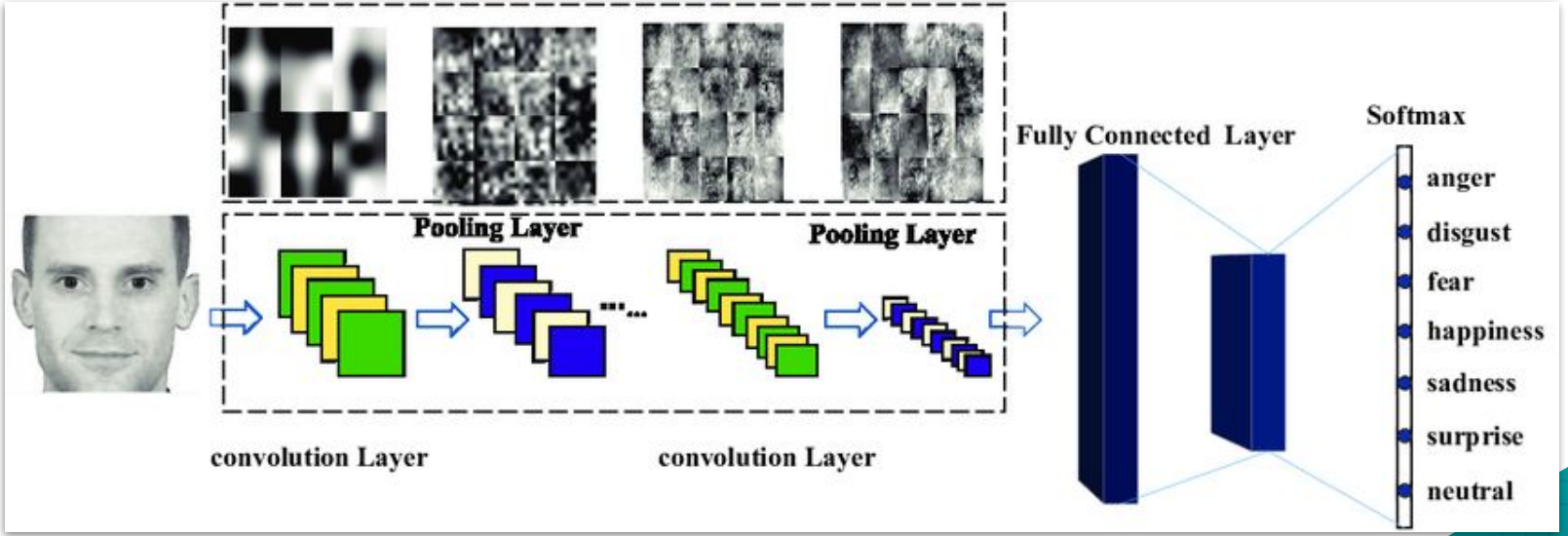


Classification Example

1.Classifying MNIST GitHub

Classification Model Example





Classification Example Loss

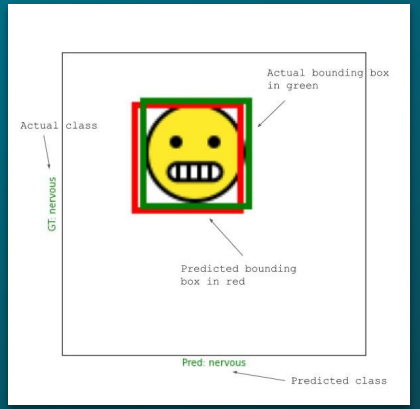
$$L_{\text{CE}} = -\sum_{i=1}^{n} t_i \log(p_i), \quad \text{for n classes,}$$

where t_i is the truth label and p_i is the Softmax probability for the i^{th} class.

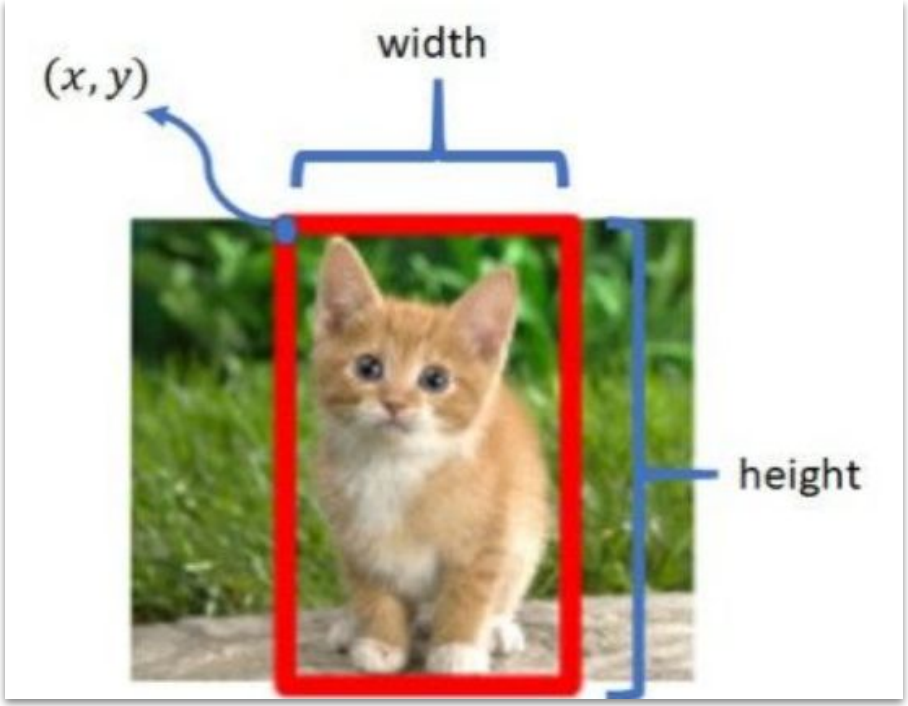
6

Computer Vision : Localization

Definition, Applications and Examples

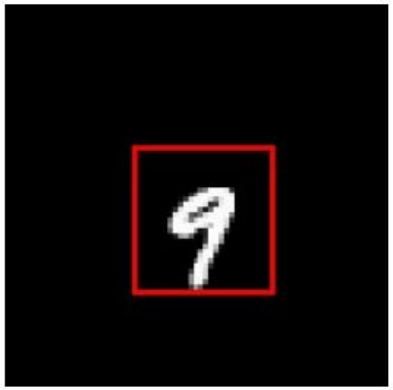


Localization



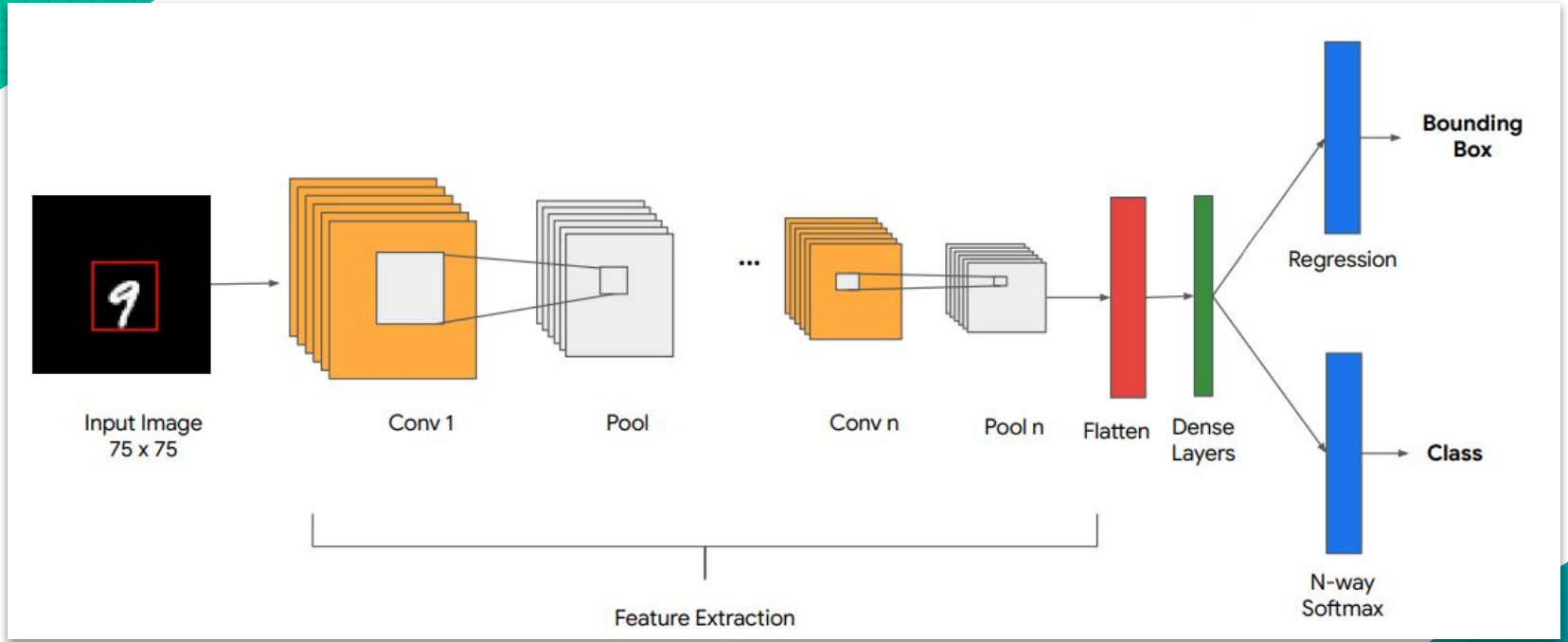
Localization Example

A sample from MNIST

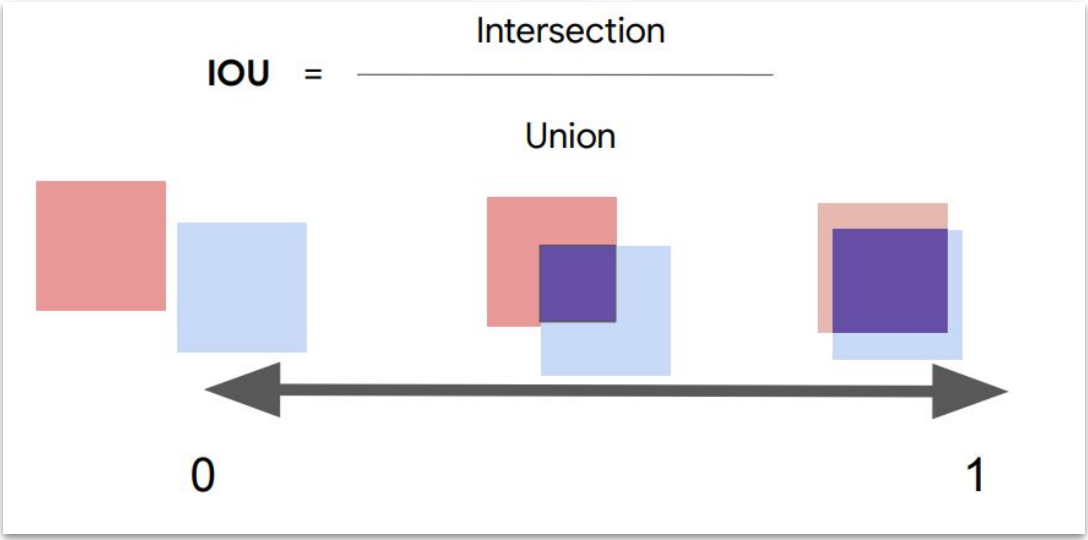


Sample pasted on a 75 x 75 black canvas

Localization Model Example



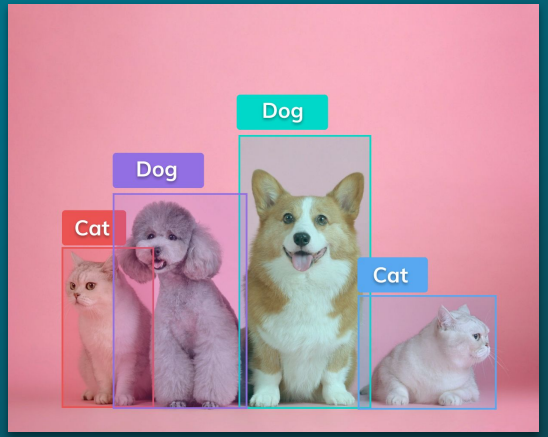
Localization Example Loss



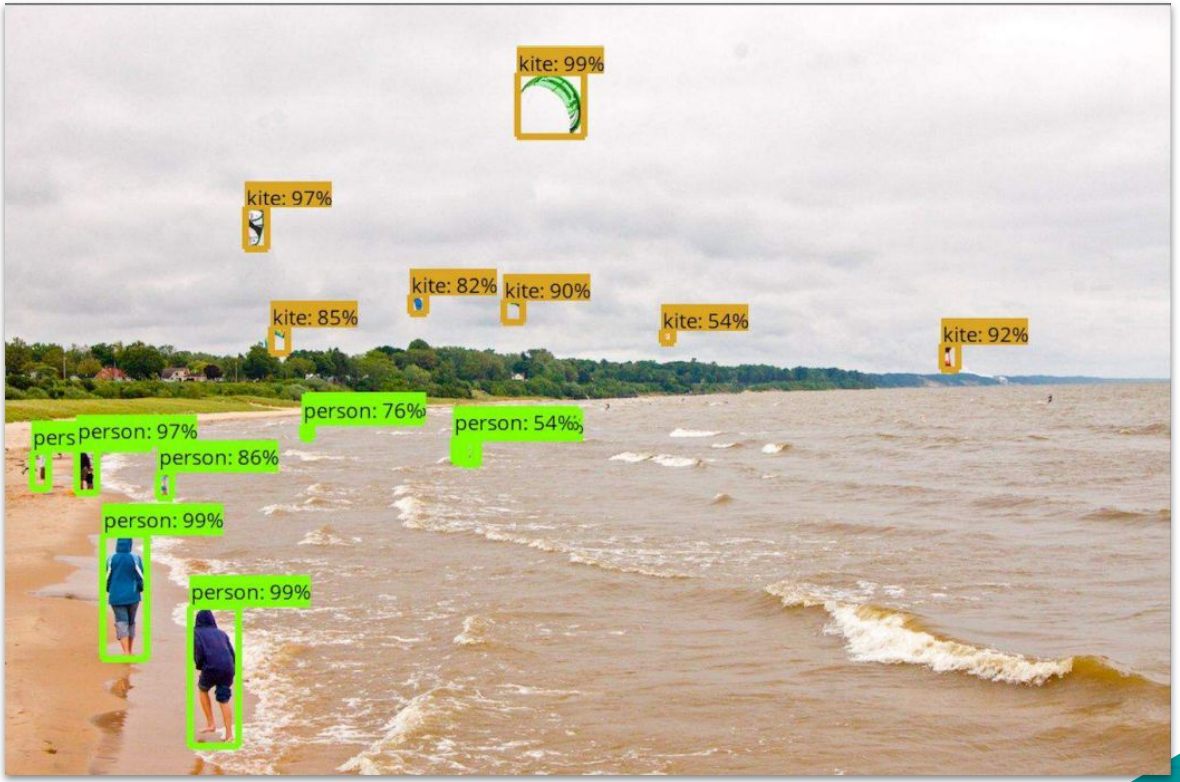
7

Computer Vision : Object Detection

Definition, Applications and Examples



Object Detection

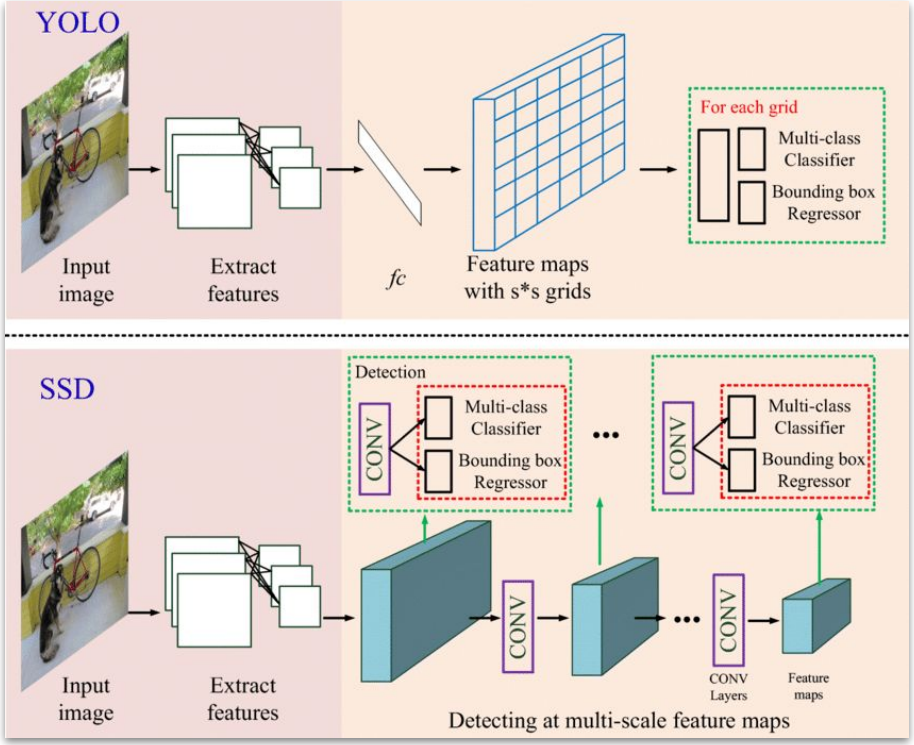


Object Detection Steps

Two phases :

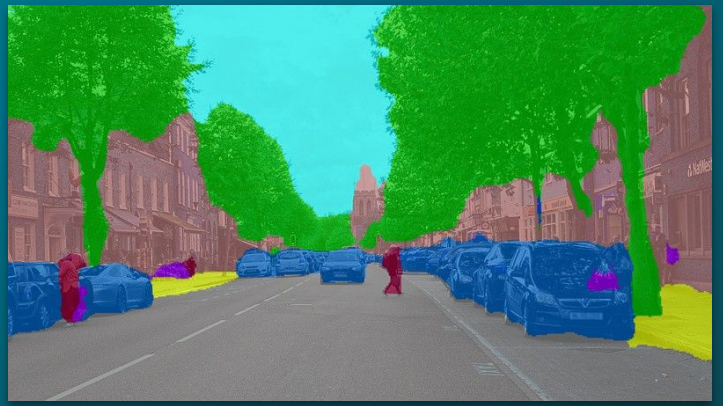
- Sliding Window & Non-Maximum Suppression
- Object Detection & Classification

Object Detection Model Examples



1. YOLOV5 GitHub
2. SSD Paper

8



Computer Vision : Segmentation

Definition, Applications and Examples

Segmentation Architecture

224 x 224 x 3 (RGB image)

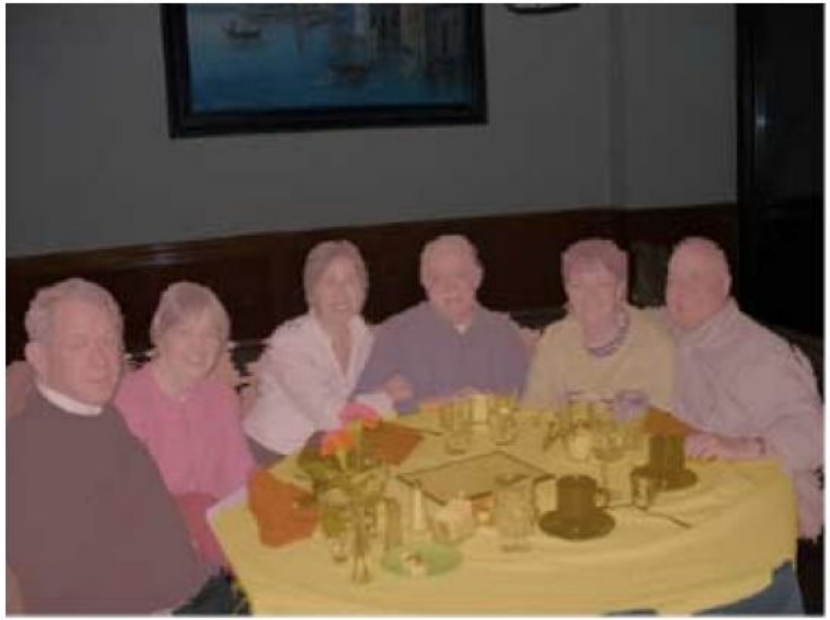
Encoder
(Feature Extractor, Down samples Image)

Downsampled Feature Map

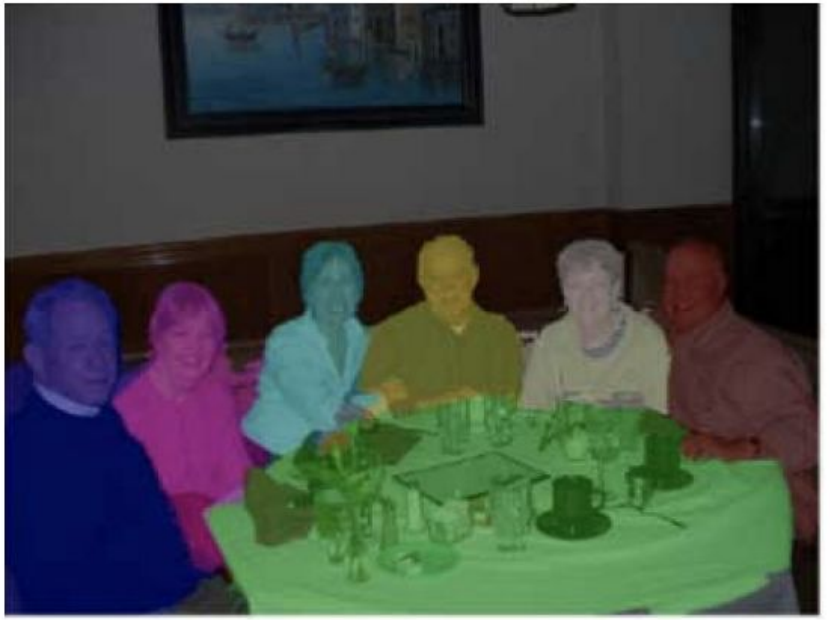
Decoder
(Up samples feature map, generates pixel-wise label map

224 x 224 x n (Pixel map)
n = number of classes

Semantic Segm. VS Instance Segm.



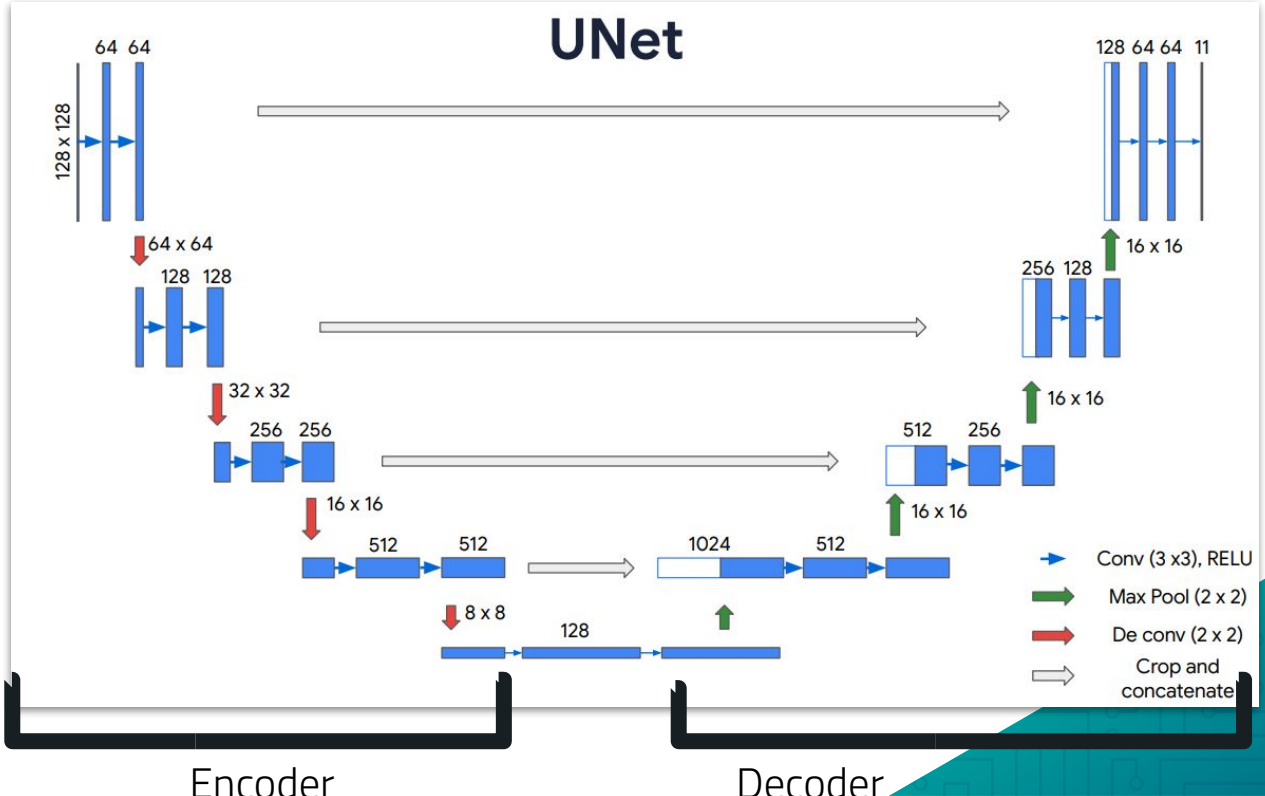
Semantic Segmentation



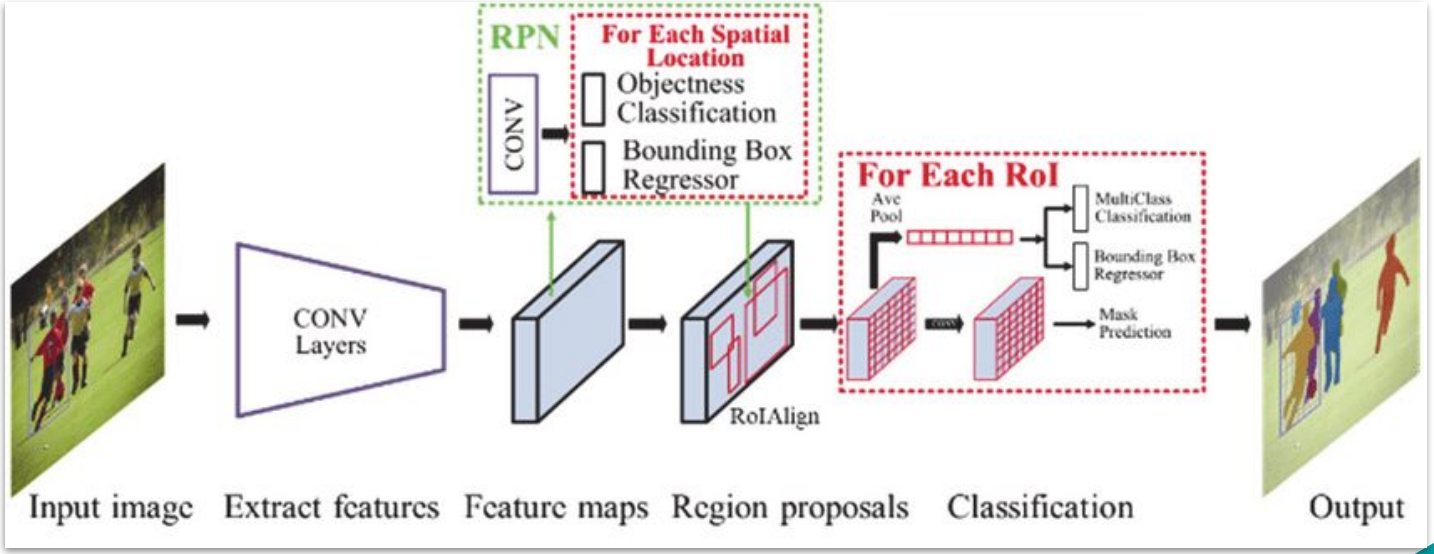
Instance Segmentation

1. UNet Paper

Semantic Segmentation Model Example



Instance Segmentation Example Model

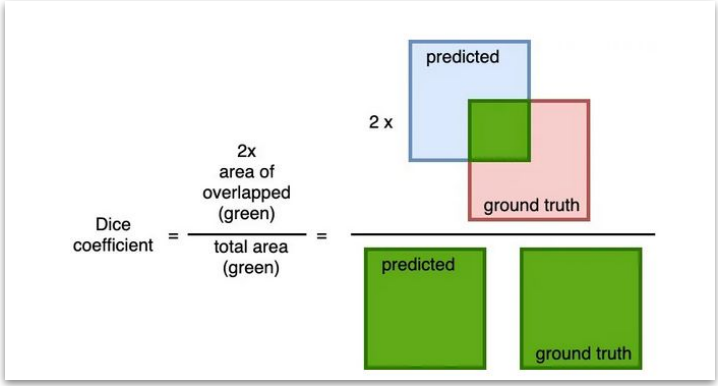


Segmentation Example Loss

$$L_{\text{CE}} = -\sum_{i=1}^{n} t_i \log(p_i), \quad \text{for n classes,}$$

where t_i is the truth label and p_i is the Softmax probability for the i^{th} class.

Or



9

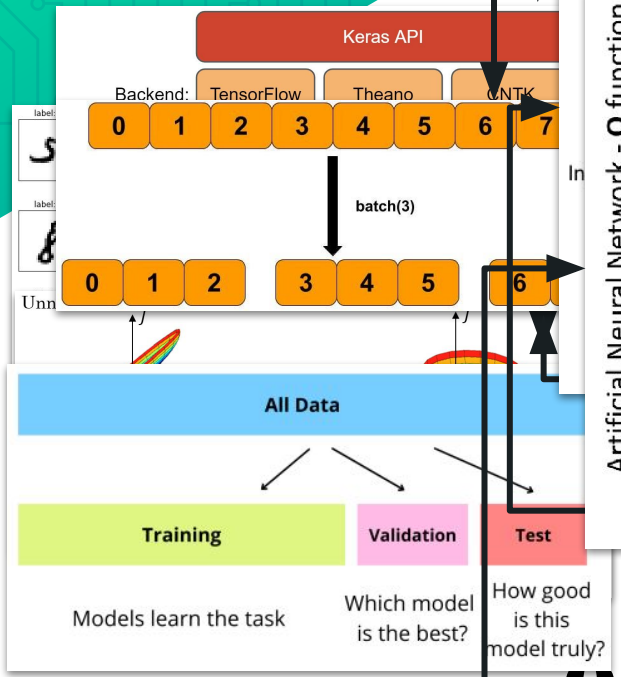
Workshop Presentation

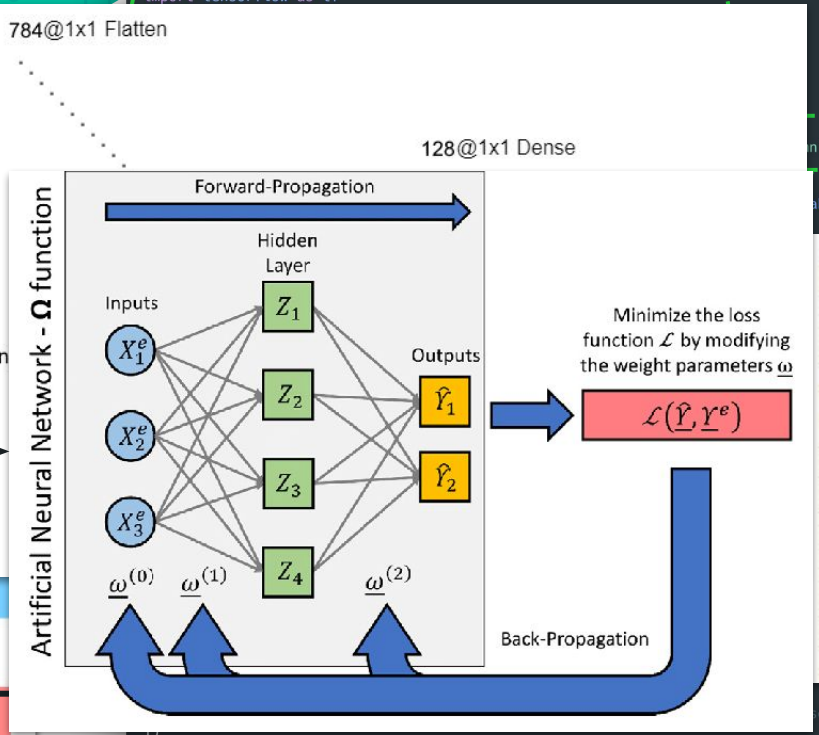
Explanation of the workshop

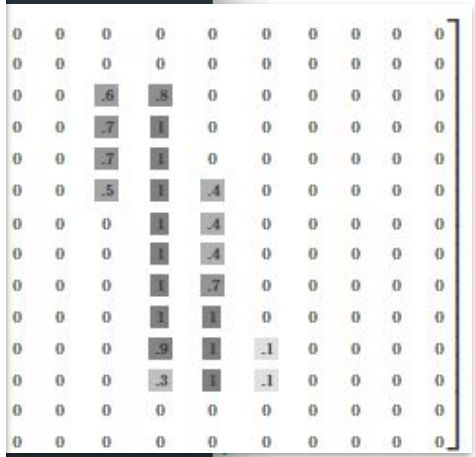
TensorFlow

Google’s open-source library for machine learning

TensorFlow Example



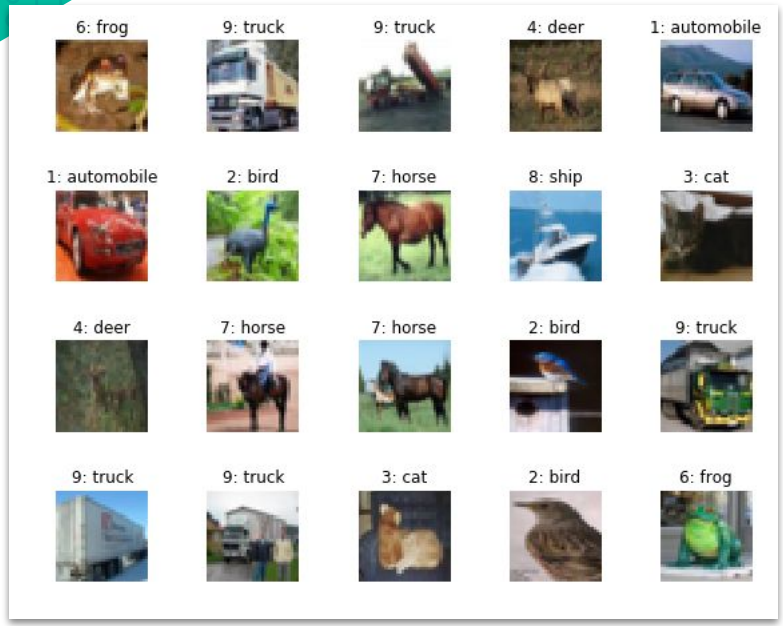




```
import tensorflow as tf
nist.load_data()
abels))
es) and softmax AF
# Define the loss function, the optimizer, and the metrics
model.compile(optimizer=Adam(),
              loss=SparseCategoricalCrossentropy(),
              metrics=[SparseCategoricalAccuracy()])
# Train the model
model.fit(train_dataset, epochs=5, validation_data=validation_dataset)
# Evaluate the model
test_loss, test_acc = model.evaluate(test_dataset)
print(f'Test accuracy: {test_acc}')
```

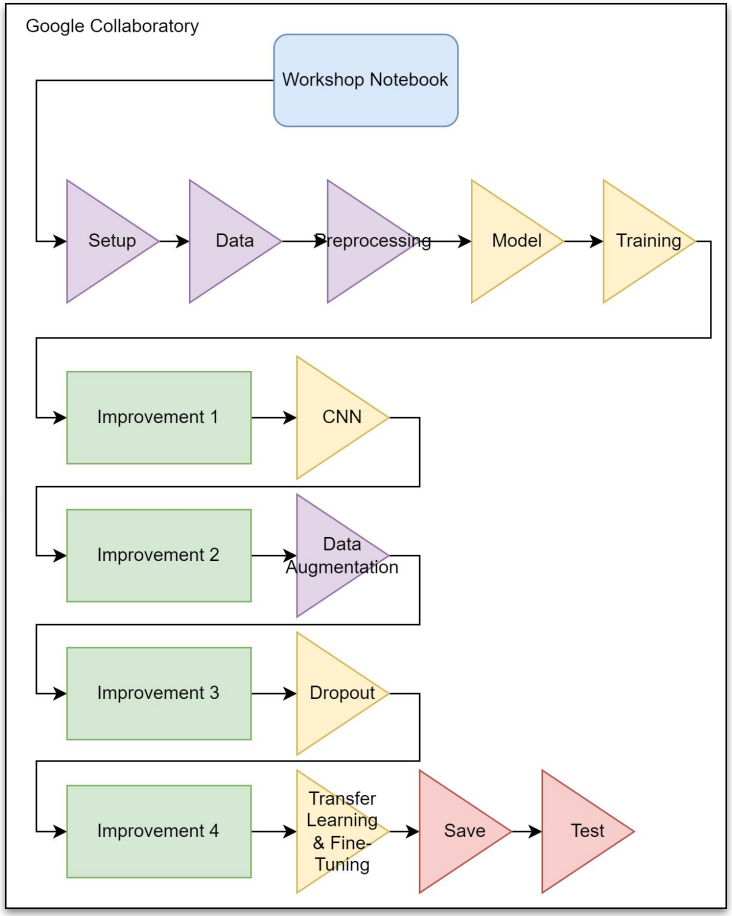
8
9
10

Workshop Subject



- Classification problem
- Model improvements
- Data improvements
- Model testing
- Model exporting

Workshop Relationships

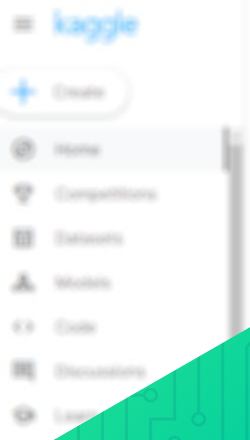


10

k

Kaggle Contest

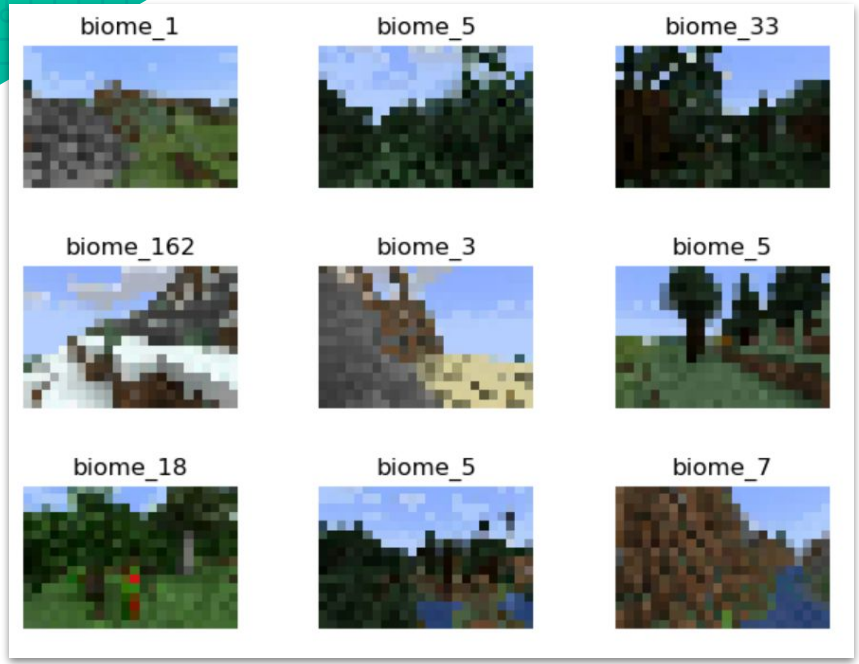
Contest subject & rules



Kaggle

Google's platform for data science competitions

Contest Subject



- Classification problem
- Minecraft Biome Dataset
- No transfer learning
- Maximum 22 layers
- Skeleton provided
- Prizes will be announced

22nd June (in 1 week)
Competition Starts

2023

JUNE JULY

20th July (in 5 weeks)
Competition Ends

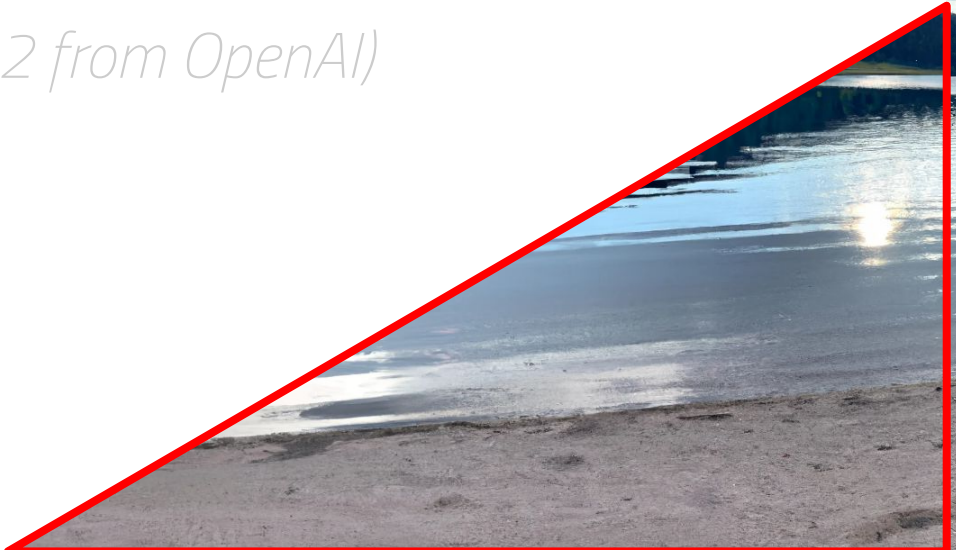
Contest Tips

- Use **successive convolutions**
- Take a look at the architecture of **VGG**
- **Validation dataset** is closest to reality
- Data is very **unbalanced**
- Don’t **overfit** too much !
- Don’t forget to create your **CSV submission** during execution !
- Inspire yourself from the **workshop**
- Don’t hesitate **to contact** SCIA folks !

Did you notice ?

20% of the photo is computer generated !

(DALLE-2 from OpenAI)



THANKS!

Any questions?

You can find me at:

- OrdinaryDev83
- tom.genlis@gmail.com
- Tom Genlis
- OrdinaryDev83#9338