Computer Vision

Computer Vision Initiation

Plan

- Context
  - Purpose
  - Problems
  - Potential
- Tasks
  - Classification
  - Object detection
  - Segmentation
  - Image creation
- Tools
  - OpenCV
  - Keras

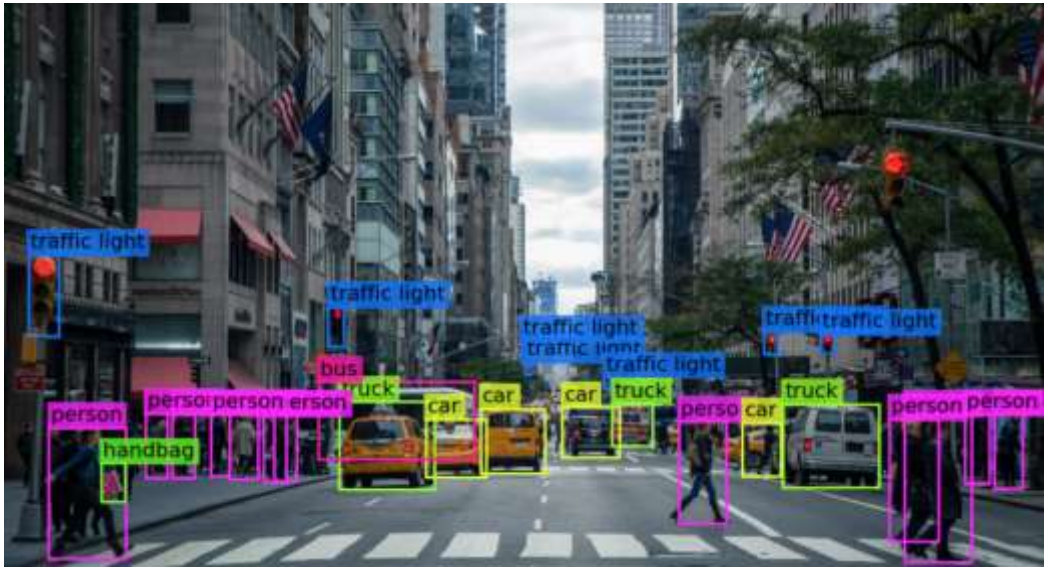
Context

Overview and… Why does it matter ?

Purpose

Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world. Using digital images from cameras and videos and **deep learning** models, machines can accurately **identify and classify objects** — and then **react** to what they “see.”

- Object detection
- Object tracking
- Image segmentation / search
- 3D scene modeling



Problems

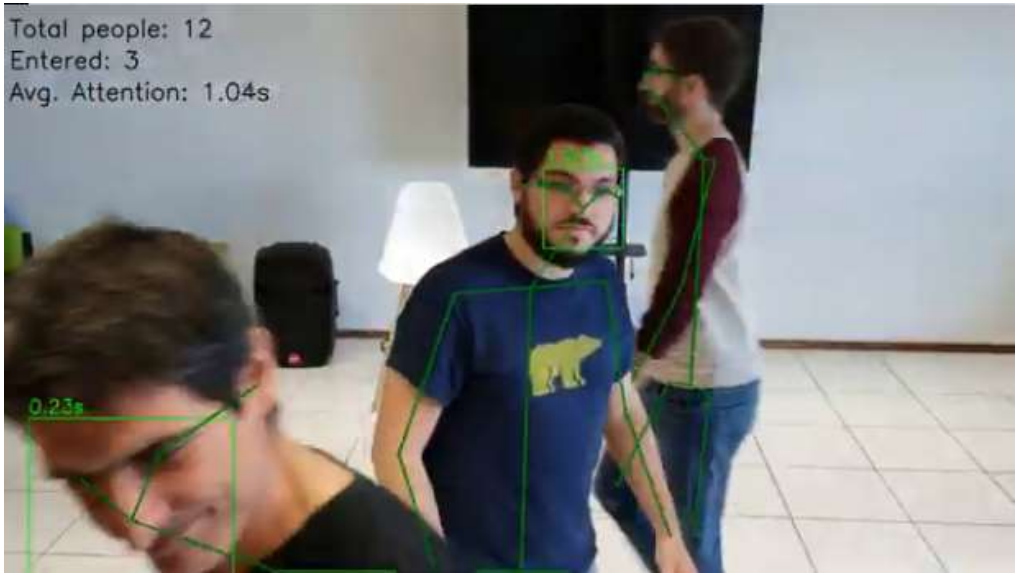
- Object orientation
- Lighting conditions
- Obstruction
- …

➔ **Too many possibilities !**

Potential

- Biomédical
- Robotics
- Airports
- Security / Surveillance
- Space exploration
- Customer experience
- ...

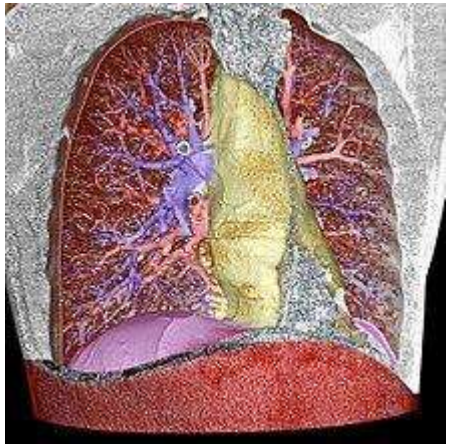
Behavioral tracking



Potential

- Biomédical
- Robotics
- Airports
- Security / Surveillance
- Space exploration
- Customer experience
- ...

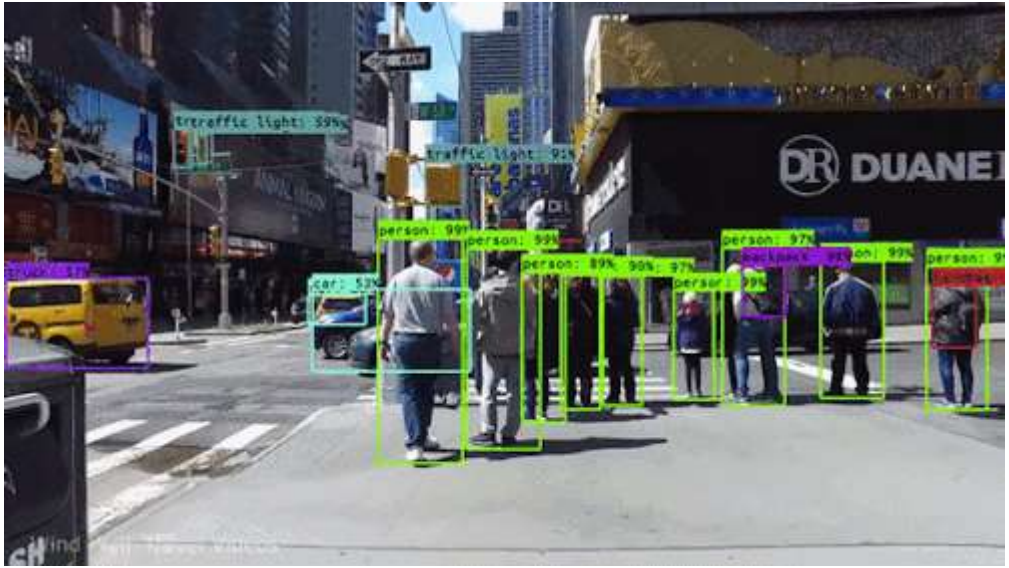
Healthcare



Potential

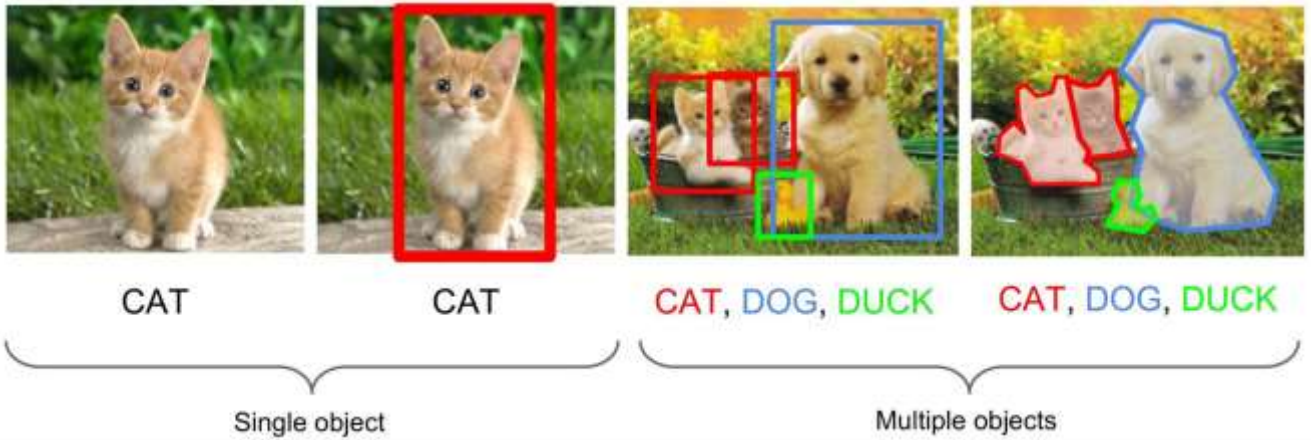
- Biomédical
- Robotics
- Airports
- Security / Surveillance
- Space exploration
- Customer experience
- ...

Autonomous vehicles

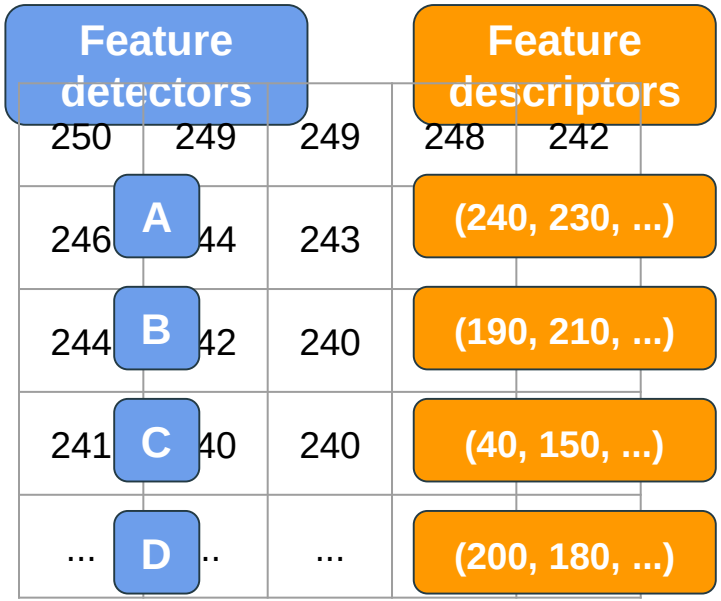


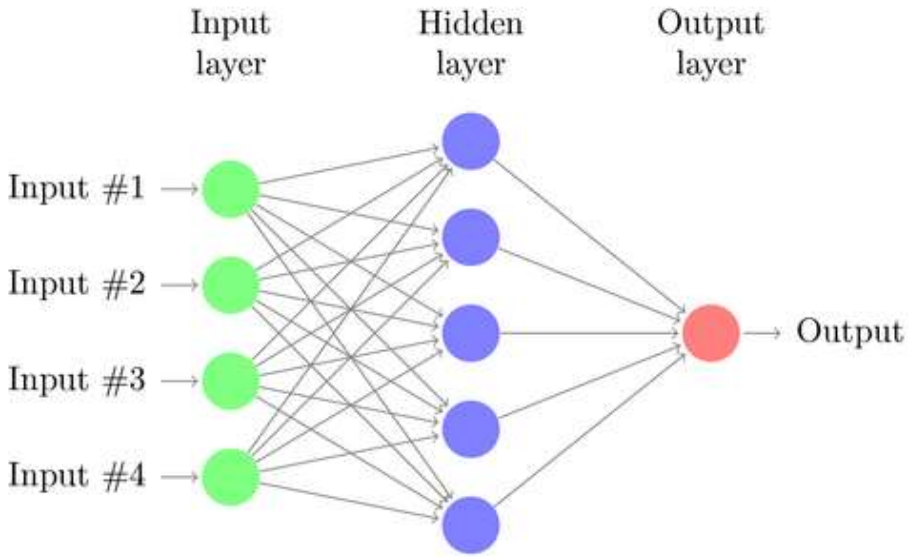
Tasks

Classification - detection - segmentation - creation

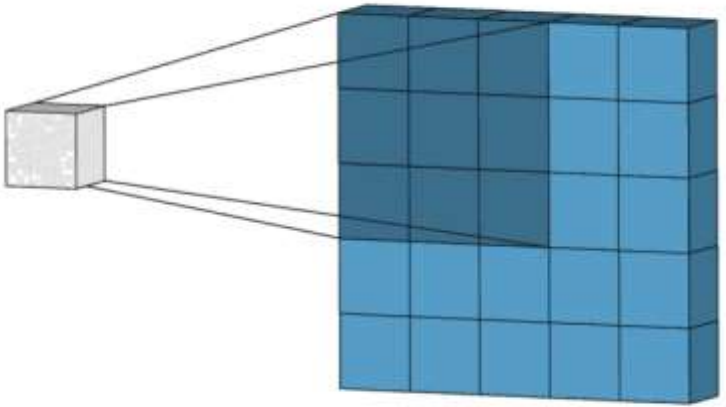


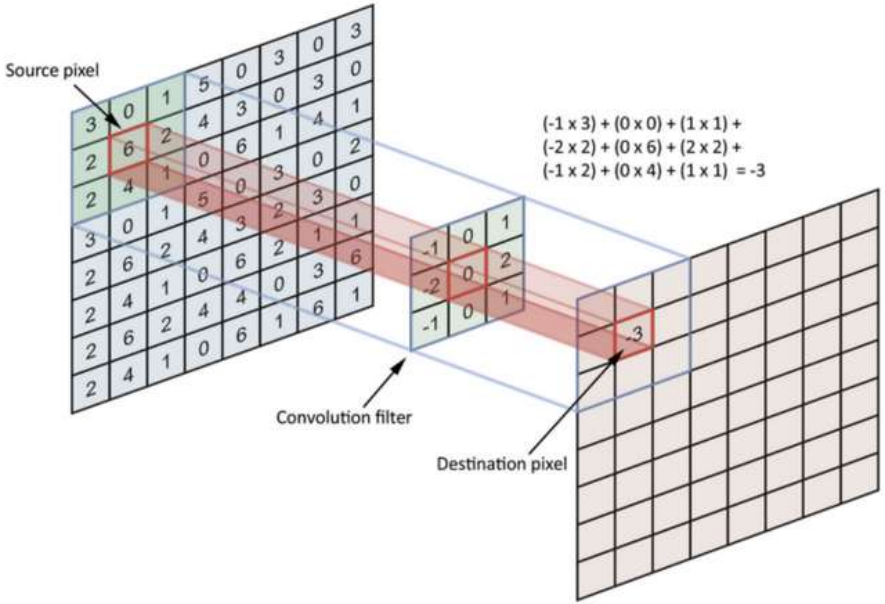
MLP


Feature detectors
Feature descriptors
250 249 249 248 242
A
246 44 243
(240, 230, ...)
B
244 42 240
(190, 210, ...)
C
241 40 240
(40, 150, ...)
D
... .
(200, 180, ...)

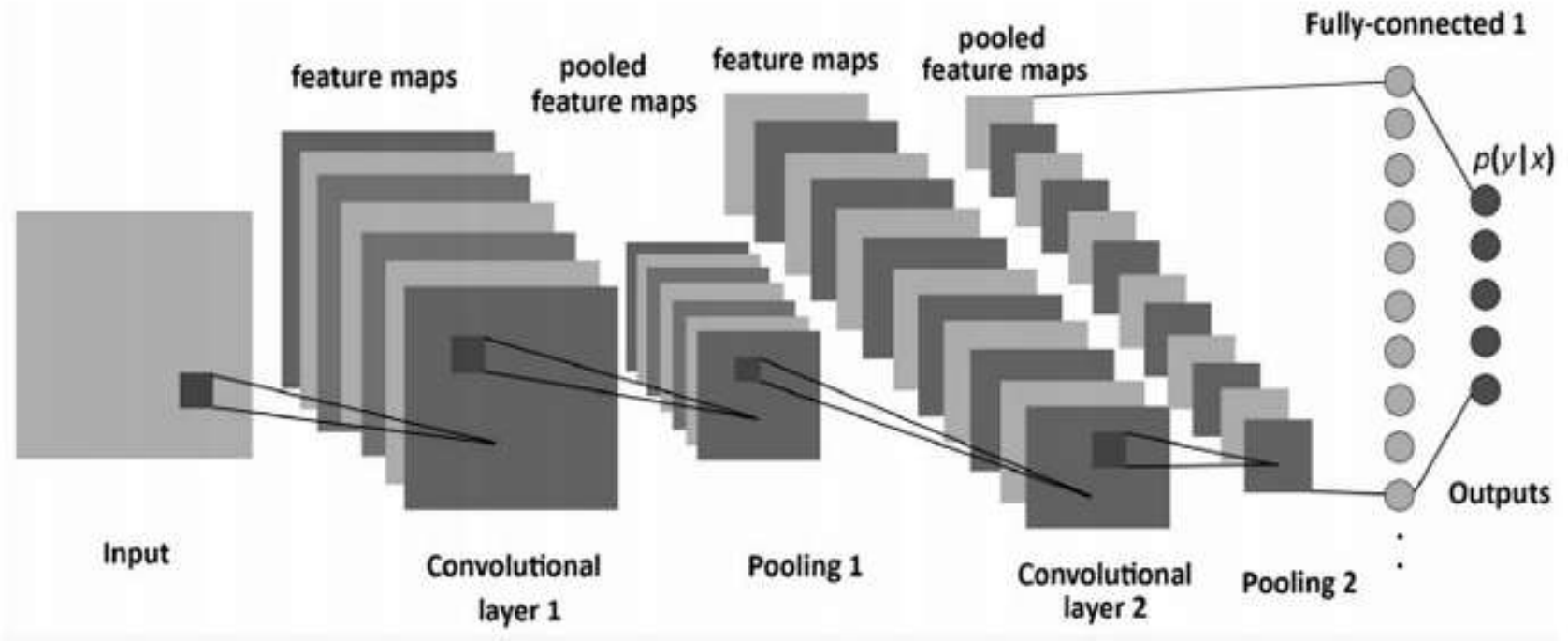

Input layer
Hidden layer
Output layer
Input #1
Input #2
Input #3
Input #4
Output

CNN




Source pixel
(-1 x 3) + (0 x 0) + (1 x 1) +
(-2 x 2) + (0 x 6) + (2 x 2) +
(-1 x 2) + (0 x 4) + (1 x 1) = -3
Convolution filter
Destination pixel

CNN


feature maps
pooled feature maps
feature maps
pooled feature maps
Fully-connected 1
p(y|x)
Outputs
Input
Convolutional layer 1
Pooling 1
Convolutional layer 2
Pooling 2

Classification

CAT

Particularity:

- N possible outputs with N = number of classes.

Algorithms

- Mark 1
- SVM
- CNN

Datasets

- ImageNet
- CIFAR 100
- ISIC
- MURA
- DermNet

Models

- LeNet5
- VGG
- GoogleNet / MobileNet
- RestNet

Object detection

Particularity:

- Is an object present ?
- Location:
  - x and y (top left of the object area, or center)
  - Area height and width

Algorithms

- Sliding window
- CNN

Datasets

- ImageNet
- COCO
- PASCAL

Models

- R-CNN
- YOLO
- SSD

Image segmentation

Particularity:

- Is an object present ?
- Object boundaries
- Give a value for each pixel

Algorithms

- CNN

Datasets

- COCO
- PASCAL
- BraTS
- Agriculture-Vision
- SpaceNet

Models

- U-net
- Mask R-CNN

Image creation

Particularity:

- Give a value for each pixel
- Unlike any training image

Algorithms

- CNN

Datasets

- Any...

Models

- GANs
- VAEs

CV Tools

Image processing and ML

OpenCV

What it can do

- Image processing
- 2D feature extraction
- Predefined ML
- GUI
- Object detection
- Video analysis and creation
- 3D reconstruction

Characteristics

- Written in C++
- Doc not great...
- First release 2006
- 40+k users / 18M download
- Supports Python, C++, Java, Matlab, CUDA
- All platformes (mobile included)
- Used by Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota

Keras

What it can do

- Built on top of tensorflow
- Deep learning framework
- Pre-trained model on ImageNet
- Multi-GPU / distributed training

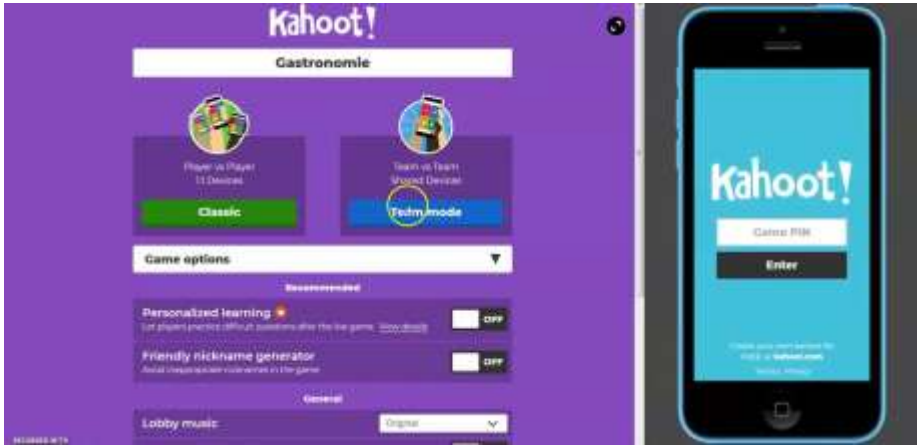
Characteristics

- Written in python
- Impressive documentation
- First release 2015
- 400+k users
- Supports Python, CUDA
- All platformes (mobile included)
- Used by Netflix, Uber, Instacart, CERN, NASA, NIH
- #2 after scikit-learn for ML usage on Kaggle competition

Exercices

Kaggle and your project

Quizz Kahoot !



- Link: https://kahoot.it/
- Pin: ...

Compete: digit recognizer

Link: kaggle.com/c/digit-recognizer

- ➢ **Create your team !**
- ➢ **Define your strategy**
- ➢ **Submit your model**