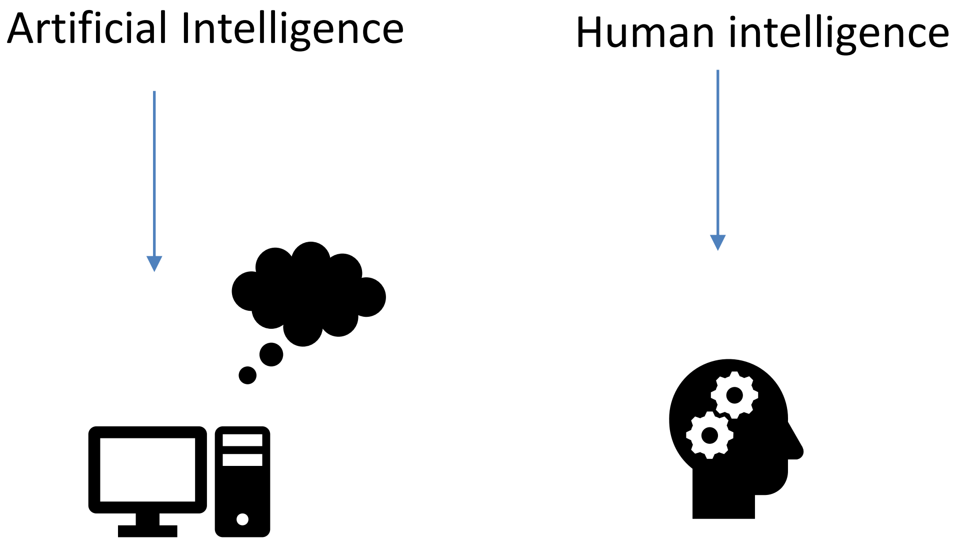
Cours Systèmes Intelligents (Part: Computer Vision)

Corentin Domken

Contact: corentin.domken@henallux.be

Introduction

What is Artificial Intelligence ?



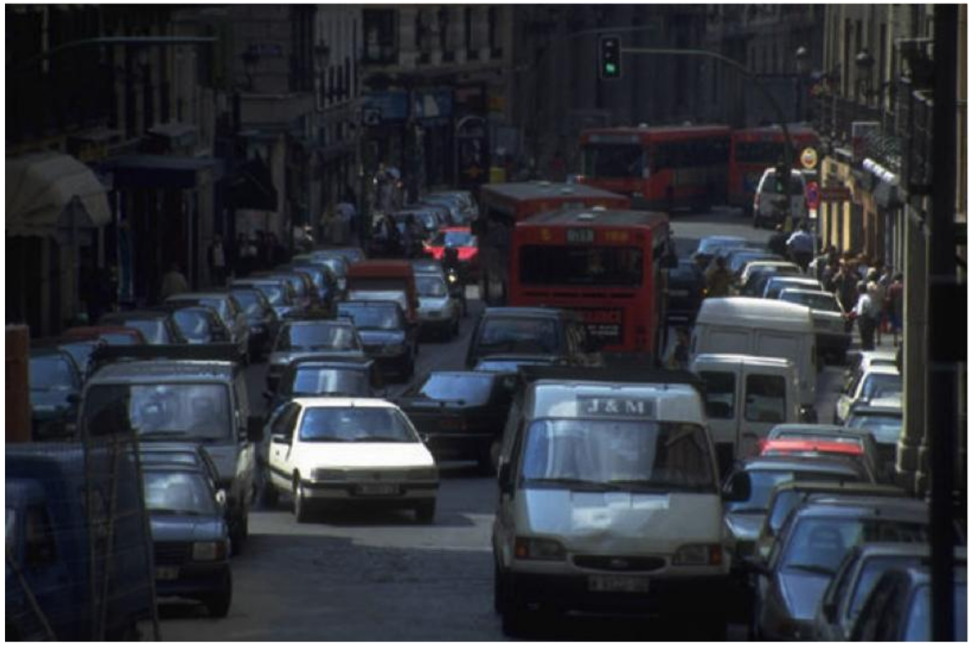
Artificial Intelligence

Human intelligence

- Making computers that think?
- The automation of activities we associate with human thinking, like decision making, learning ... ?

What is Computer Vision?

- Make computers understand images and videos.



What kind of scene?

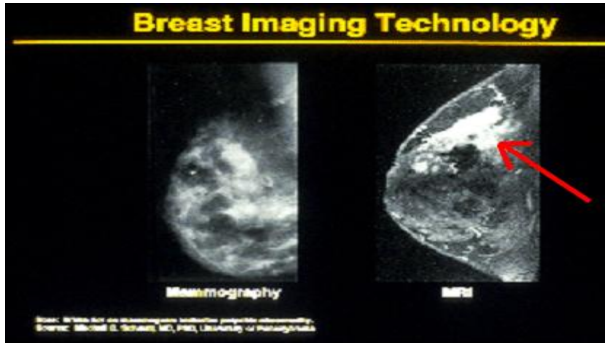
Where are the cars?

How far is the building?

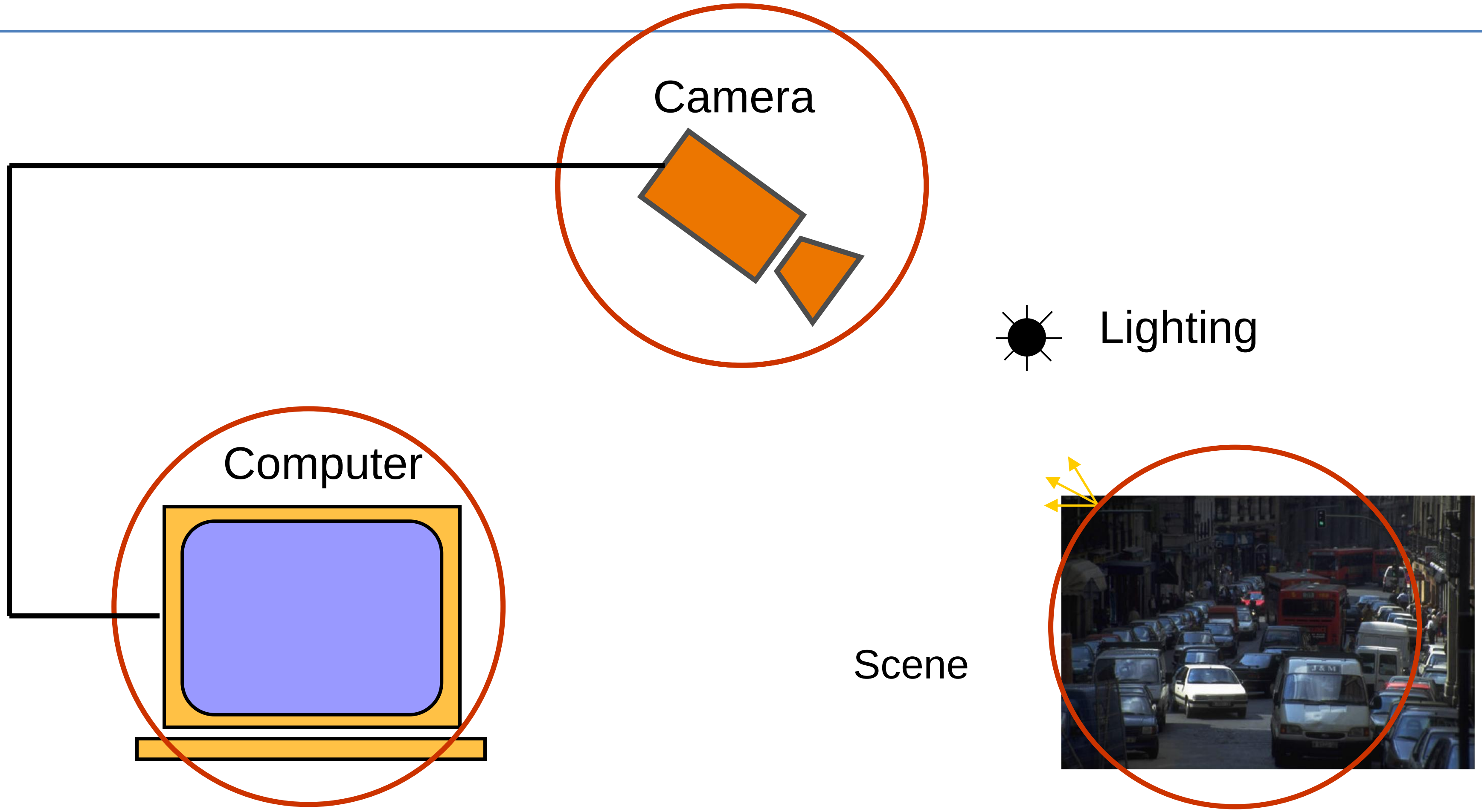
…

Computer Vision

- **Computer vision** is the science and technology of machines that see
- Concerned with the theory for building artificial systems that obtain information from images.
- The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medical scanner



Components of a computer vision system



Camera

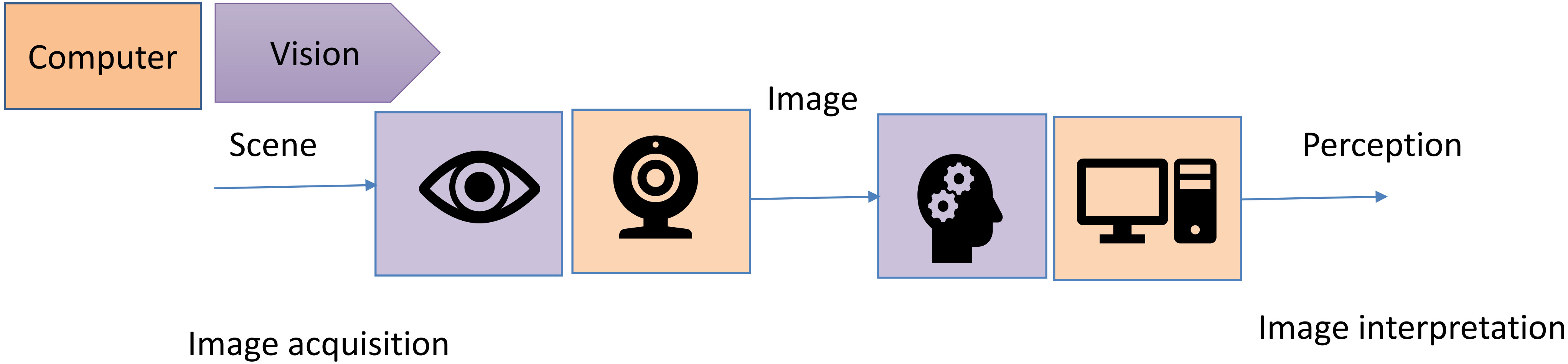
Lighting

Computer

Scene

Scene Interpretation

Computer vision vs human vision



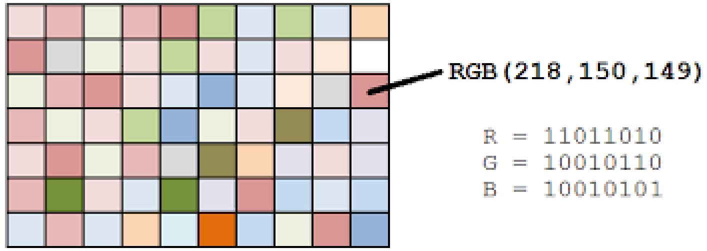
Computer vision vs human vision

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Computer vision vs human vision



- Take a picture with a camera, it is just a bunch of colored dots (pixels)
- Given an image or more, extract properties of the 3D world

-Traffic scene

- Number of vehicles
- Type of vehicles
- Location of closest obstacle
- Assessment of congestion
- Location of the scene captured
- …

→ Need Machine Learning

What is machine learning

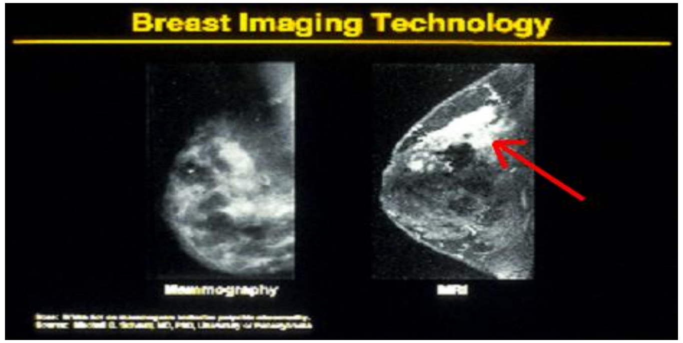
- ML is a **subset of the larger field of artificial intelligence (AI)** that "focuses on teaching computers how to learn without the need to be programmed for specific tasks,"
- Automatically learn to recognize complex patterns and make intelligent decisions based on data

Computer Vision vs. Machine Learning

- In Machine Learning, it usually does not care about how to obtain the data or sensors
- Machine Learning is very useful for Computer Vision (e.g., learning for vision)

Why computer vision matters

Agriculture



Health

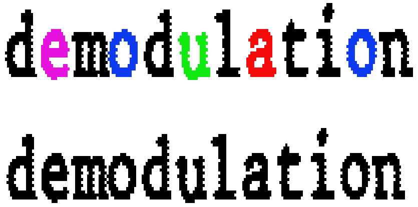
Security

Comfort

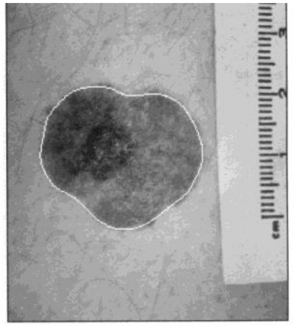
Fun

Access

Applications



Character Recognition



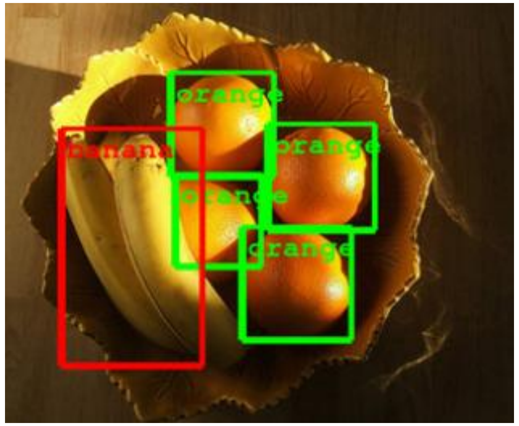
skin cancer

breast cancer

Traffic Monitoring

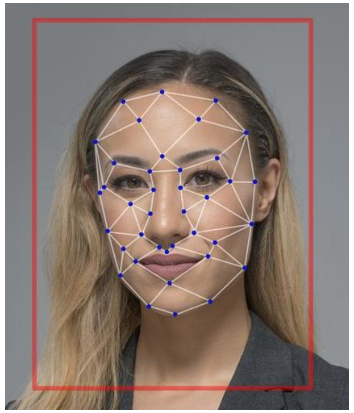
Autonomous Vehicles

Applications

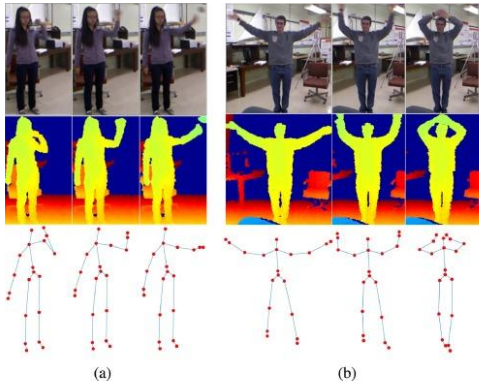


Object recognition

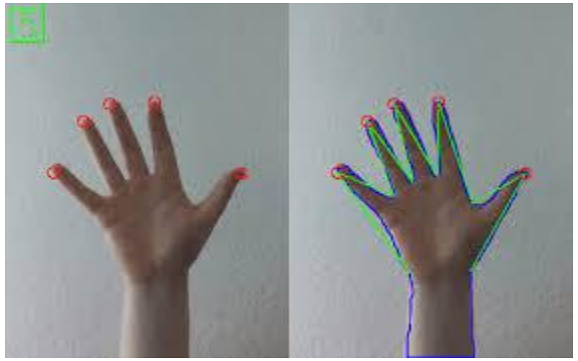
Face detection



Face recognition



Action recognition



Hand Gesture Recognition

Expression Recognition

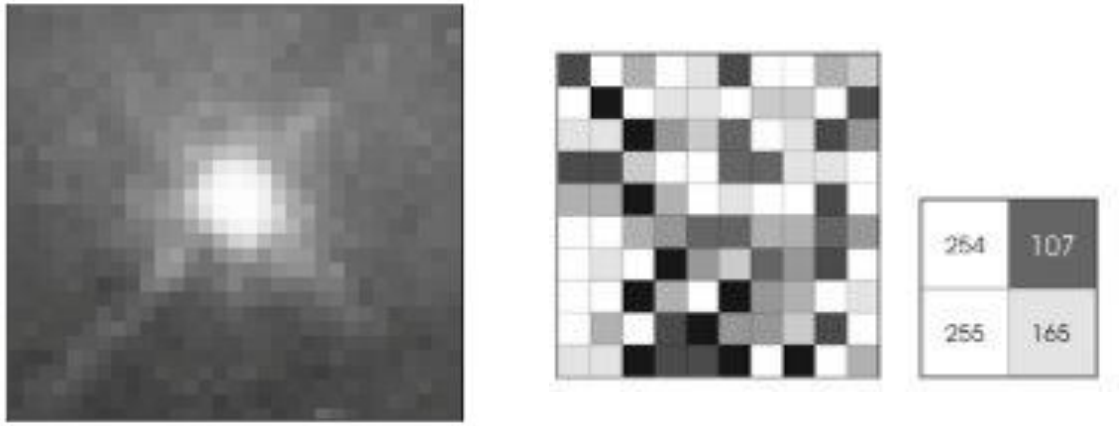
Basics on Computer vision

Computer Vision is more than just learning!

- In Computer Vision, we care how to obtain the visual data (sensor design, active vision), how to represent the visual data, and others

What is image?

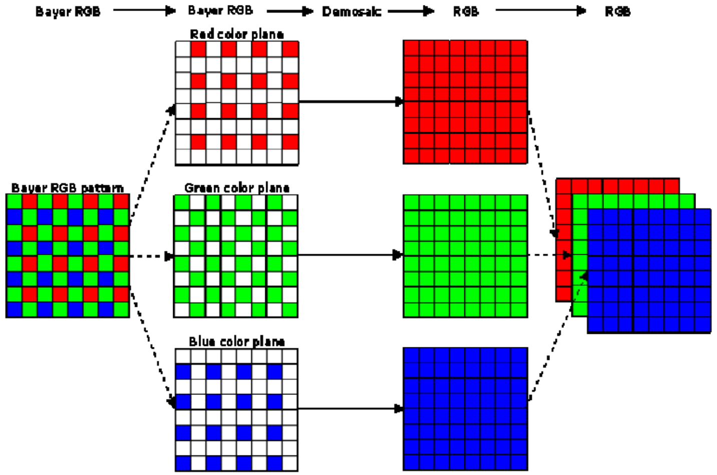
- An image is an array, or a matrix, of square pixels (picture elements) arranged in columns and rows
- For a grayscale image (Black and White), each pixel typically consists of 8 bits (1 byte)
- 8 bits represents $2^8 = 256$ tonal levels (0-255)
- Each pixel has an assigned intensity that ranges from 0 to 255



What is image?

- A “true colour” image has 24 bits (3 bytes e.g., RGB)
- Each pixel is represented has 24 bits to encode color information = 8x 8x 8 bits = 256 x 256 x256 colours = ~16 million colours

Consumer camera sensor

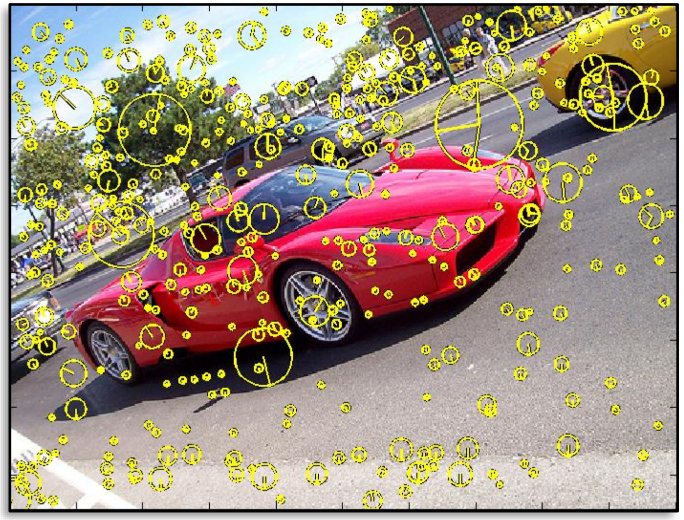


Local Visual Features

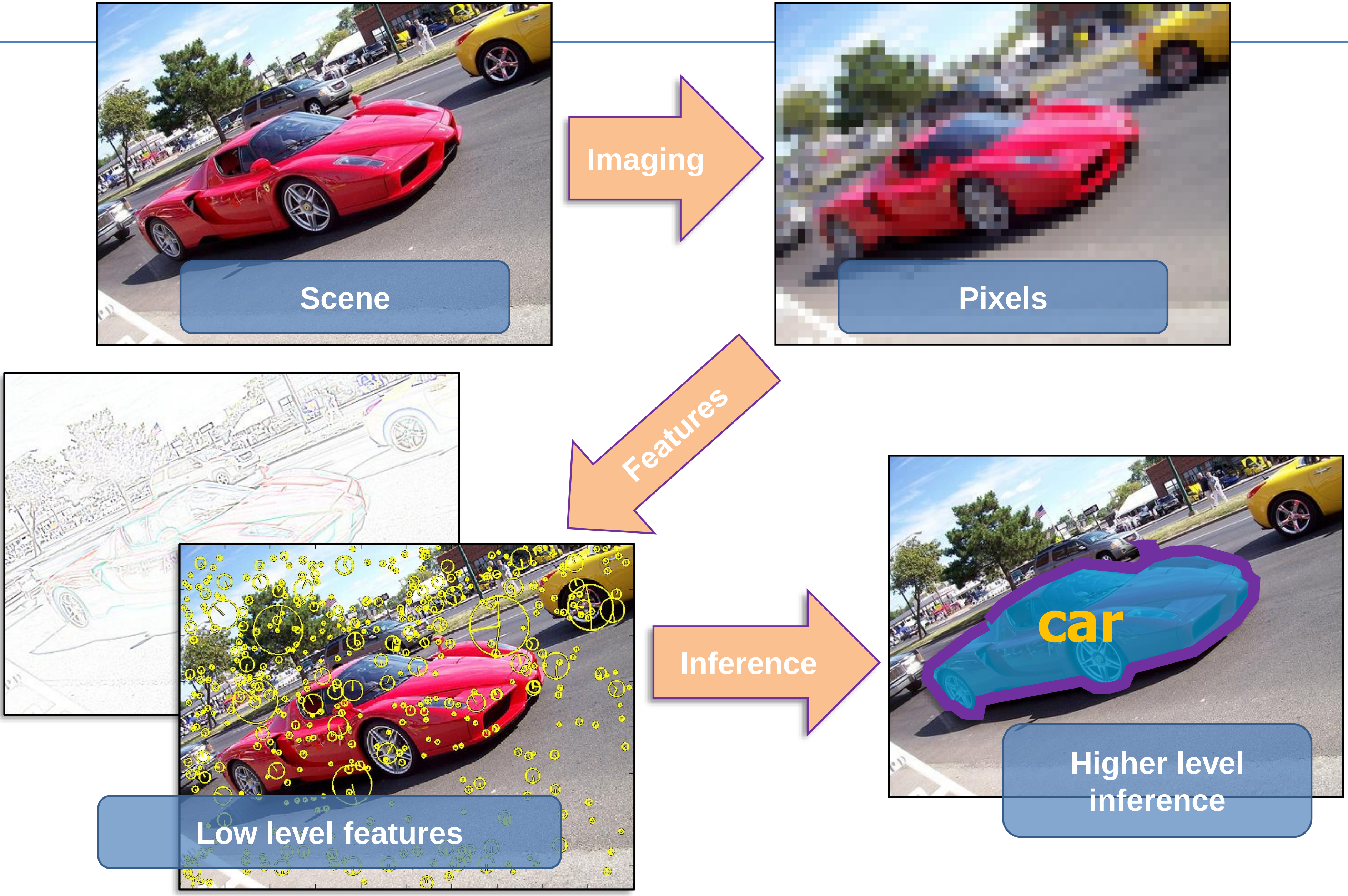
Images are high dimensional!

(640 width) *(480 height) = (307200 pixels)

Compute image statistics in a region (e.g., estimate the distribution of image gradient orientations)



From Pixels to Visual features



Scene

Imaging

Pixels

Features

Low level features

Inference

car

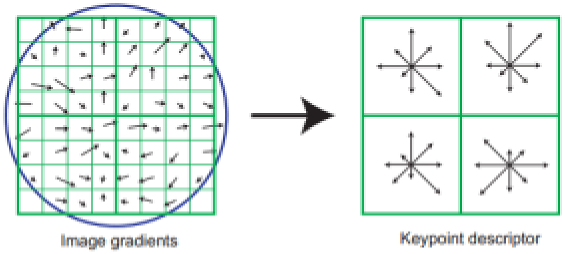
Higher level inference

Some famous features

- SIFT: Scale Invariant Feature Transform
- HOG: **Histogram of oriented gradients**
- LBP: Local binary pattern

SIFT

- **scale-invariant feature transform** (**SIFT**) is a feature detection algorithm in computer vision to detect and describe local features in images
- Published by David G. Lowe in 1999
- Invariant to scaling, rotation and translation
- Transforms an image into a large collection of local feature vectors (local descriptors calledSIFT keys)



Object recognition

HOG

- The **histogram of oriented gradients (HOG)** is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image
- Intuitively it tries to capture the shape of structures in the region by capturing information about gradients.

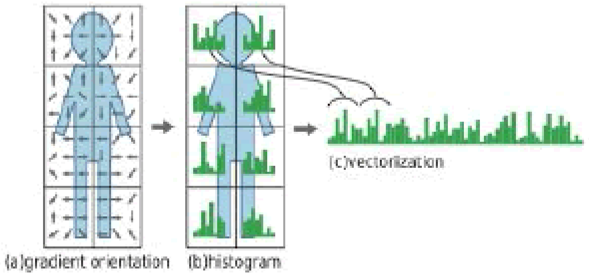
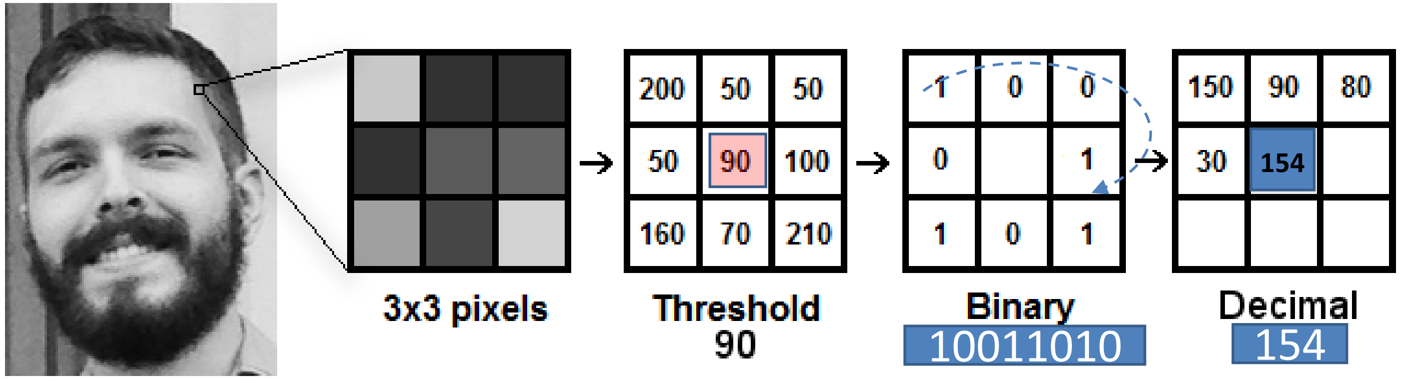


Figure 3.5 (a) (b) Overview of HOG calculation

Person following with a mobile robot

LBP

Local Binary Pattern



200	50	50
50	90	100
160	70	210

1	0	0
0		1
1	0	1

150	90	80
30	154	

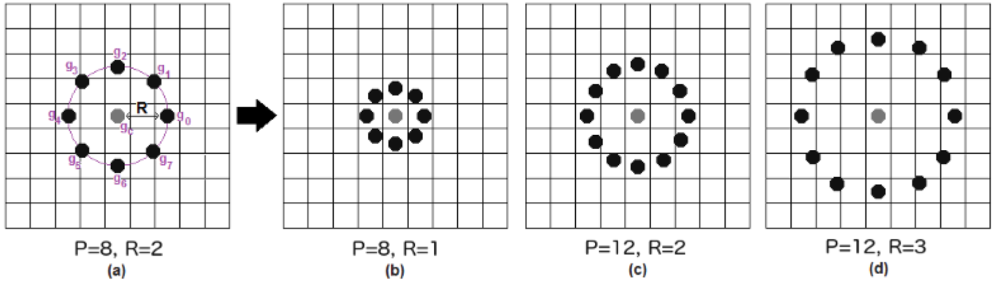
LBP Result

Facial image in grayscale

The intensity of each pixel (0~255)

$$10011010 = (1 \times 2^7) + (0 \times 2^6) + (0 \times 2^5) + (1 \times 2^4) + (1 \times 2^3) + (0 \times 2^2) + (1 \times 2^1) + (0 \times 2^0) = 154$$

Different number of radius and neighbors



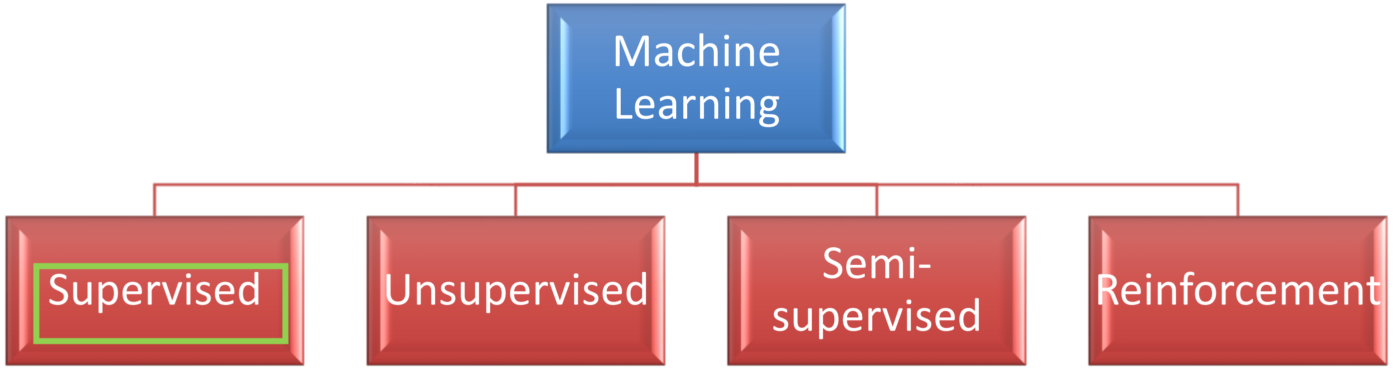
Programming tools

- OpenCV is a great performing computer vision tool and it works well with C++ as well as Python
- OpenCV is prebuilt with all the necessary techniques and algorithms to perform several

- Matlab is a great tool for creating image processing applications and is widely used in research

Basics on Machine Learning

Types of machine learning

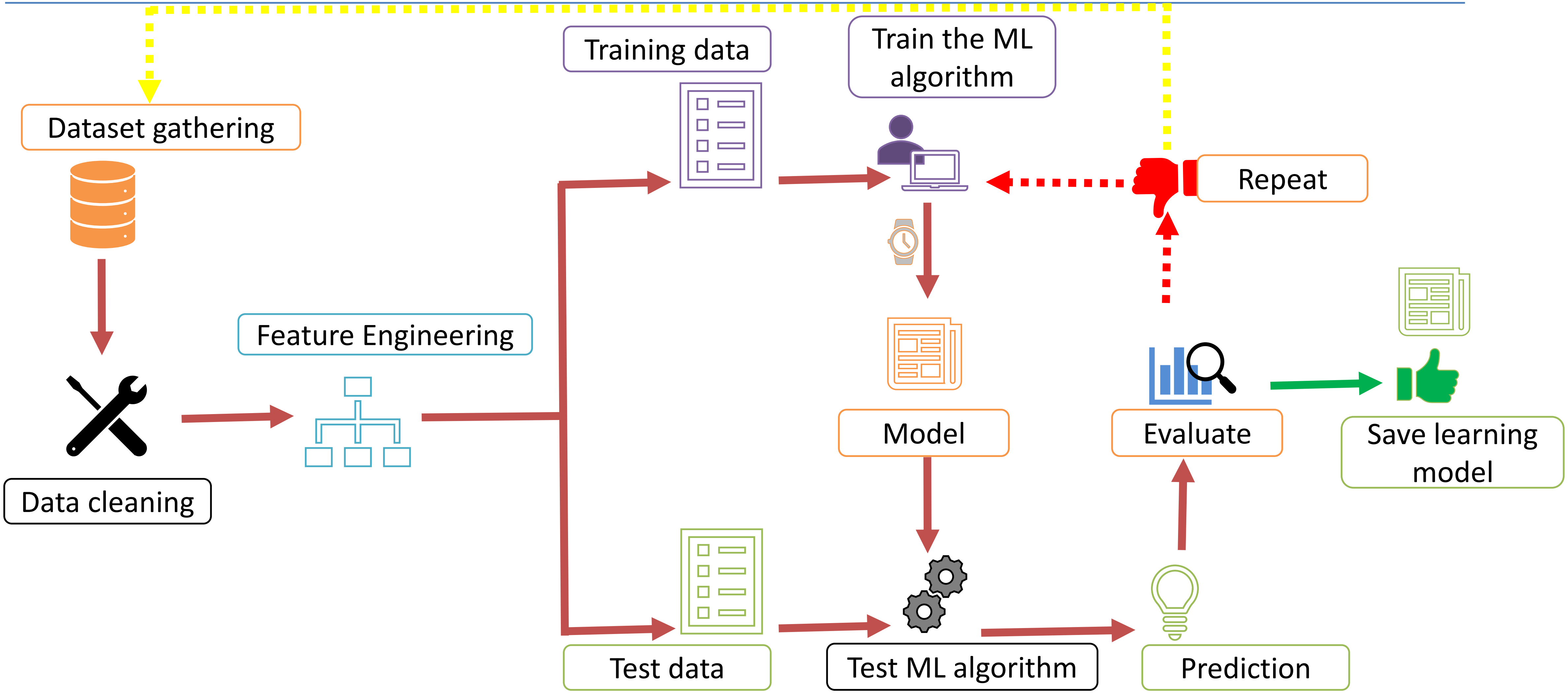


Task driven
(regression/classification)

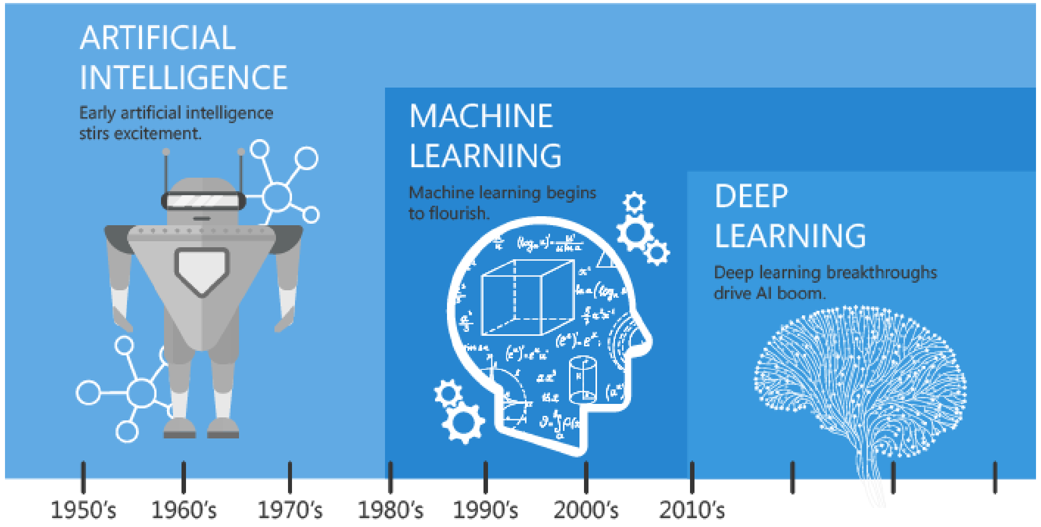
Data driven
(Clustering)

Algorithm learns to react
to environment

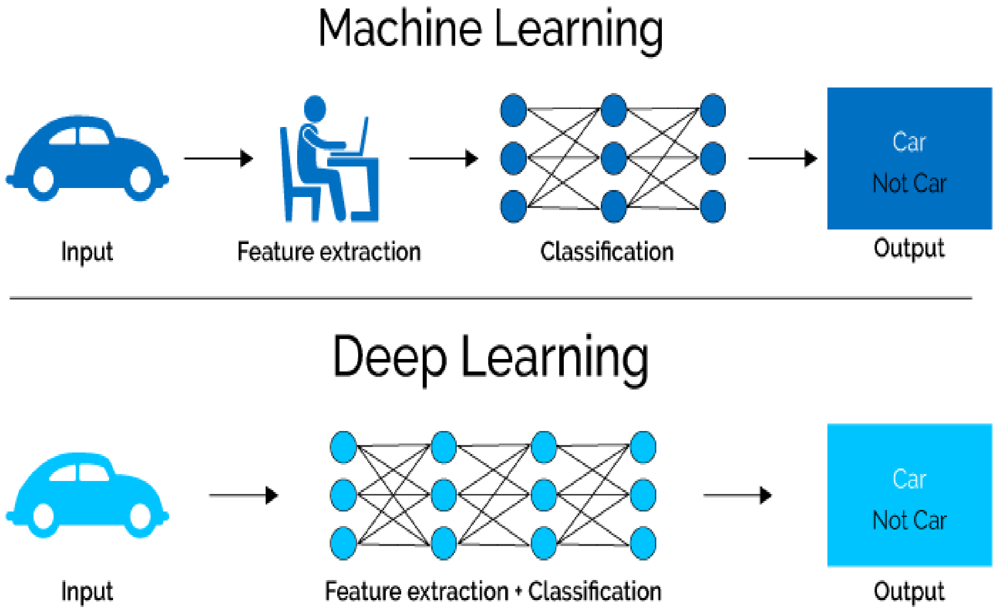
Machine learning process



Machine Learning VS Deep Learning


ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
MACHINE LEARNING
Machine learning begins to flourish.
DEEP LEARNING
Deep learning breakthroughs drive AI boom.
1950's
1960's
1970's
1980's
1990's
2000's
2010's

Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.



Questions

References

- https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b
- https://docs.opencv.org/3.4.3/da/d60/tutorial_face_main.html
- https://www.pyimagesearch.com/2015/12/07/local-binary-patterns-with-python-opencv/
- https://www.superdatascience.com/opencv-face-detection/

Object recognition on Raspberry

- Tutorial on: https://www.framboise314.fr/i-a-realisez-un-systeme-de-reconnaissance-dobjets-avec-raspberry-pi/
- Use neural networks

Face recognition project

- Face recognition
- https://github.com/informramiz/opencv-face-recognition-python/blob/master/README.md
- Digit recognition
- https://www.learnopencv.com/handwritten-digits-classification-an-opencv-c-python-tutorial/
- Face recognition: HOG+KNN
- https://github.com/emersonlaurentino/hog-face-recognition/blob/master/knn.py
- https://github.com/irfanhanif/FaceRecognition-HOG-PCA-SVM/blob/master/experiment.py
- https://github.com/vj18/HOG-SVM-Classifier/blob/master/makeSvmModel.py ok
- https://github.com/gsg213/Face-Recognition-using-LBP/blob/master/take_data_set.py