

Computational vision

Introduction

INSTRUCTORS: Francesca Odone and Matteo Moro

TA: Jacopo Dapueto

ENROLL TO THE AULAWEB PAGE (2024-2025)

Attendance modalities

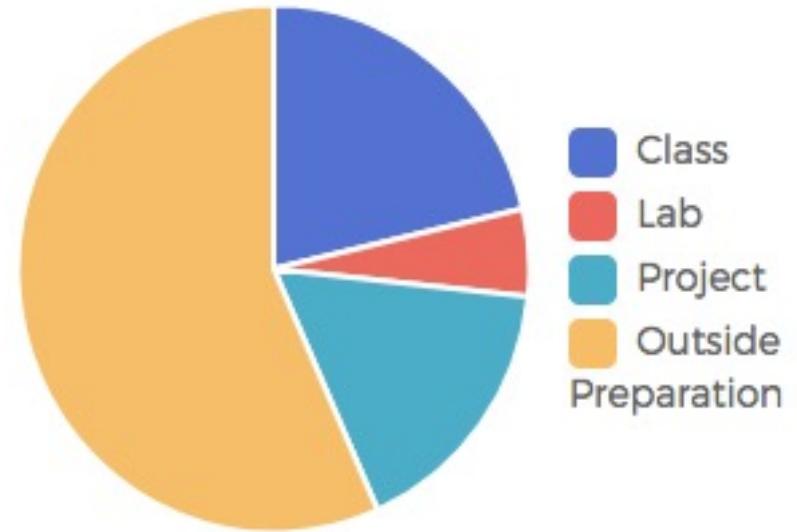
Course style

~24 hours theory

16 hours hands-on activities

1 CFU Project

Effort Breakdown



Material

Classes will often be held at the BLACKBOARD, especially in the first part of the course

We will post material (slides, papers, links) on Aulaweb

Reference book

R. Szeliski “Computer Vision: Algorithms and Applications, 2nd ed.”

2022

<https://szeliski.org/Book/>



READ Chapter 1

Exam modalities

EVALUATION MODALITY

Submit homeworks and mid-term assignment

- 50% project (individual or in pair)
- 50% oral covering all the program, included hands on activities

Details on the project/extension of hands-on

- We will evaluate the code and a short technical report
- There will be NO presentation, but you may present a demo at the oral; we might ask questions about the project during the oral

Details on the theory assessment

- The theory assessment will **cover the whole program --- both theory and lab classes/material**
- You will be asked broad questions, but to pass the exam you must provide technically sound answers (*intuitions are not enough to pass the exam*)
- At the end of the semester (May) we will organize one or several mock oral sessions, on 1 pre-assigned topic per person, not mandatory, dedicated in particular to students with no/little experience with oral exams. More details will be provided later in the semester

What is computer vision

Computer vision as an AI task

Intelligence and Perception

In the evolution the perception of surrounding environment played a crucial role



Vision is an “innate” ability of humans and, as such, it is often given for granted

In this photo there are some boats...
Tell me how many do you see

All of them!



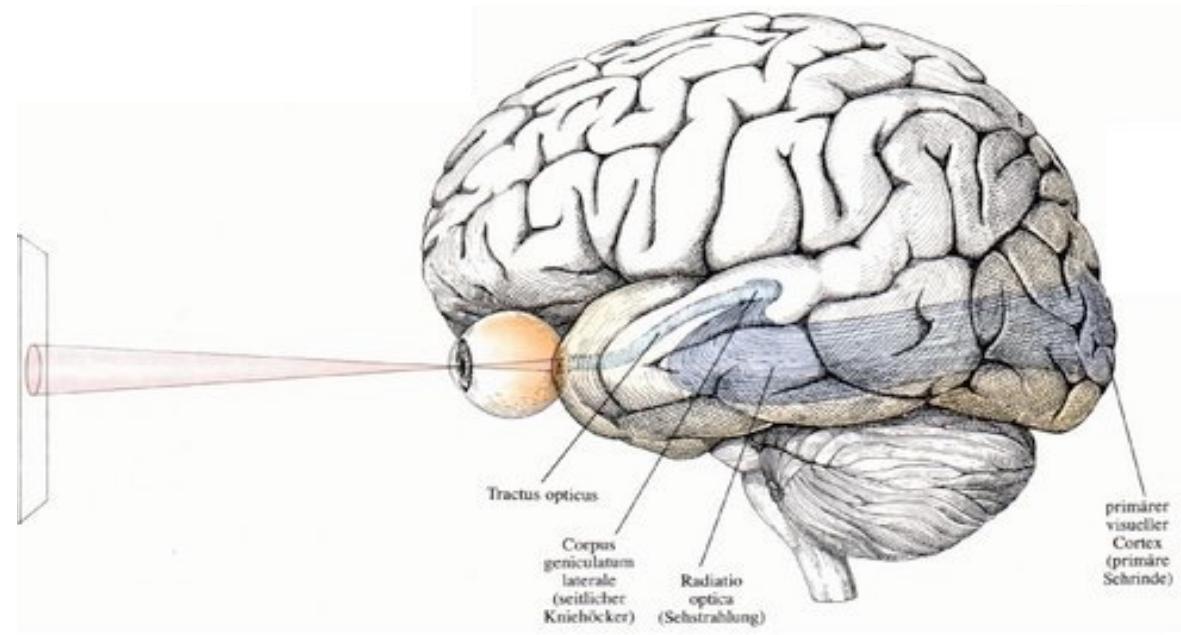
Computer vision as an AI task

vision
perception
intelligence

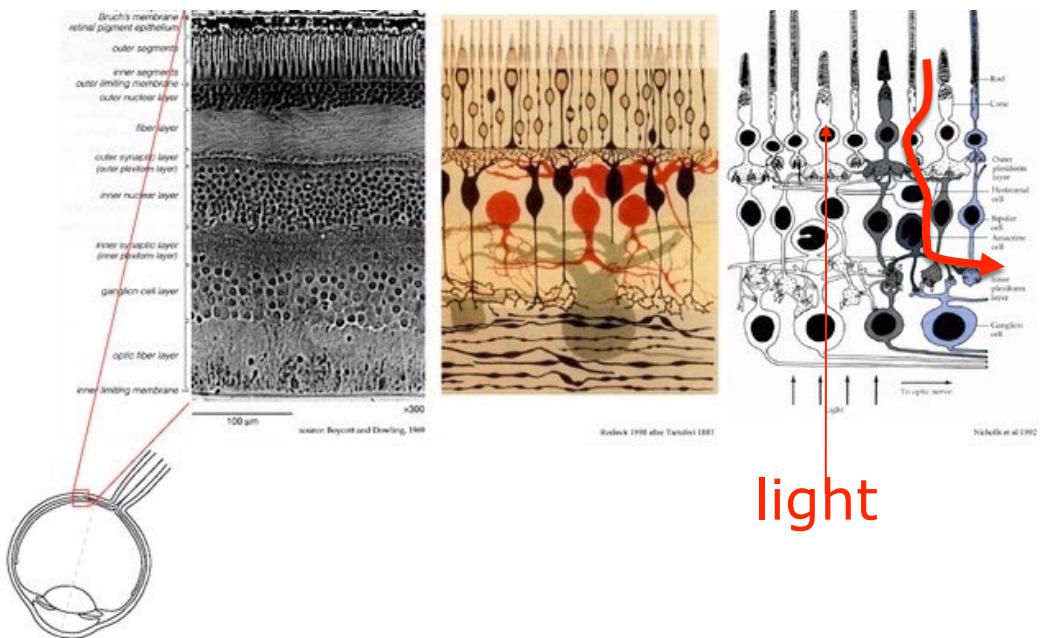
“It seems a human!”



Human vision



Hubel 1985

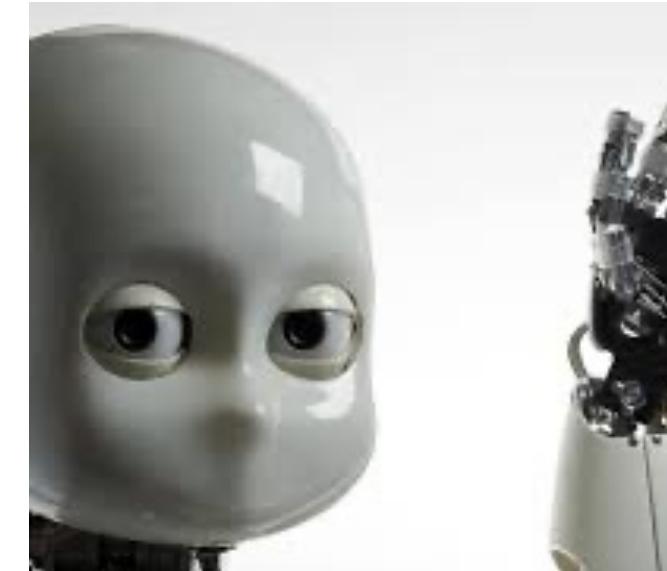


Neural
signal

light

Computer vision

A camera can play the role of the computer “eye”



- The acquired *images* are **meaningless**, until they are processed by computer vision algorithms
- Similarly, the image formed on the retina needs to be interpreted by the brain

98	103	102	110	118	118	119	119	118	118	109	88
98	105	101	110	118	118	119	118	116	113	105	84
92	98	96	109	116	121	130	130	142	141	151	145
95	98	98	104	110	112	124	127	148	147	157	159
95	98	98	104	110	112	124	127	148	147	157	159
103	104	107	111	116	121	128	128	137	135	146	169
101	106	106	110	116	119	128	128	134	133	145	166
99	109	106	118	127	131	143	145	154	153	155	168
102	110	110	121	131	136	148	148	157	157	160	169
102	110	111	124	136	140	153	154	164	165	167	174
105	113	112	124	130	135	147	147	159	159	167	175
104	113	112	125	134	137	144	147	161	161	169	177
102	110	108	122	131	131	140	140	149	150	157	168
103	109	109	121	128	131	139	140	149	148	156	167
101	106	103	116	127	133	144	143	148	148	149	159
84	94	91	103	113	118	132	134	145	146	146	149
85	92	91	103	114	119	134	135	146	145	146	149
70	82	81	91	97	100	112	115	131	130	139	142
70	82	81	91	97	100	114	115	131	132	139	142
77	76	76	82	89	89	100	101	115	113	127	135
111	85	84	79	81	81	90	90	102	100	111	125
107	86	88	79	79	79	88	88	100	101	110	126

Computer Vision

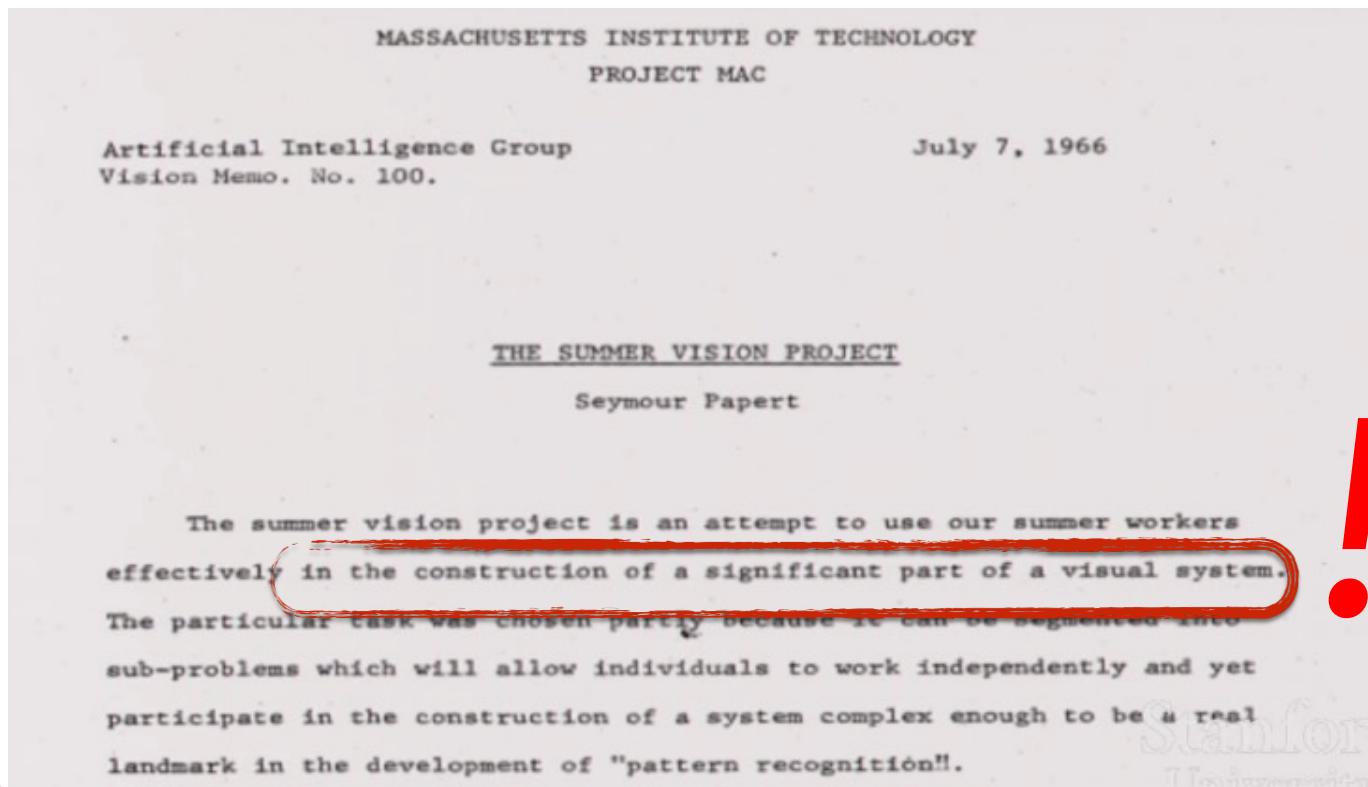
Definition

Computer vision is about extracting a **description** of the world starting from images or image sequences

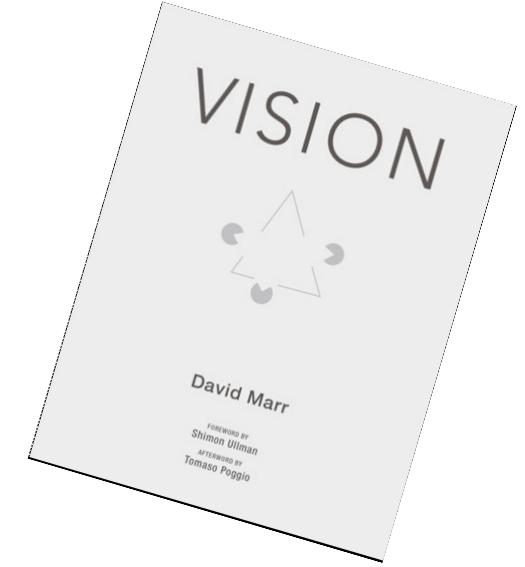
“Description” is a vague term, including geometry (shapes, distances,...), dynamics (motion, actions, ...), semantics (categories, properties, ..)

It all started

In 1966 Marvin Minsky, one of the fathers of AI, underestimates the visual perception complexity and assigns a project to a summer intern requiring to “solve the vision of a computer”



The birth of a new discipline



VISION by David Marr “A computational investigation into the human representation and processing of visual information”

It defines the building blocks at the basis of modern computer vision

Interesting projects in classical CV

Image based rendering The Matrix (1997)

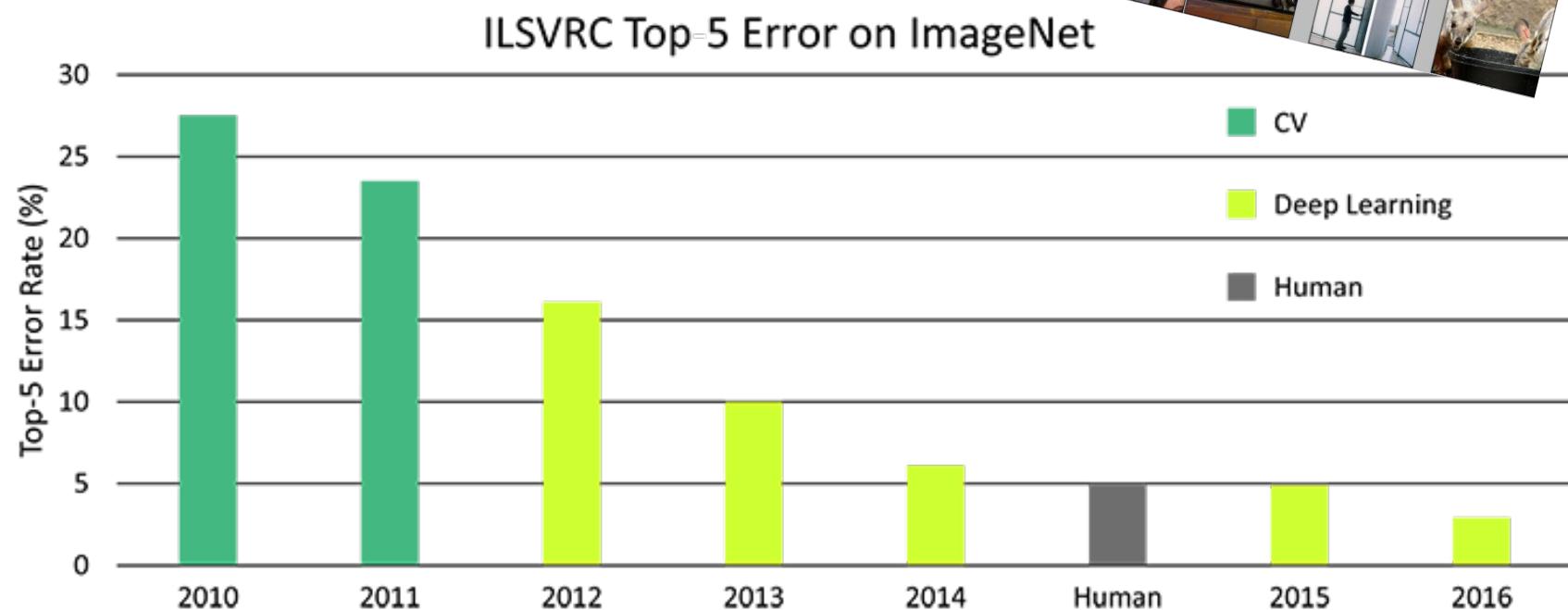
Have a look at Pauldevebec.com 1997

Image stitching / panoramas



Machine learning and Deep Learning in CV

Larger amount of data needed



COOL STUFF

Deep dream



“Initially it was invented to help scientists and engineers to see what a deep neural network is seeing when it is looking in a given image. Later the algorithm has become a new form of psychedelic and abstract art.”

Deepdreamgenerator.com

Body Pose estimation



Deeplabcut.com

Computer Vision application domains

OCR

Quality control

Video-surveillance

HMI / VR / AR

Medical

Sports

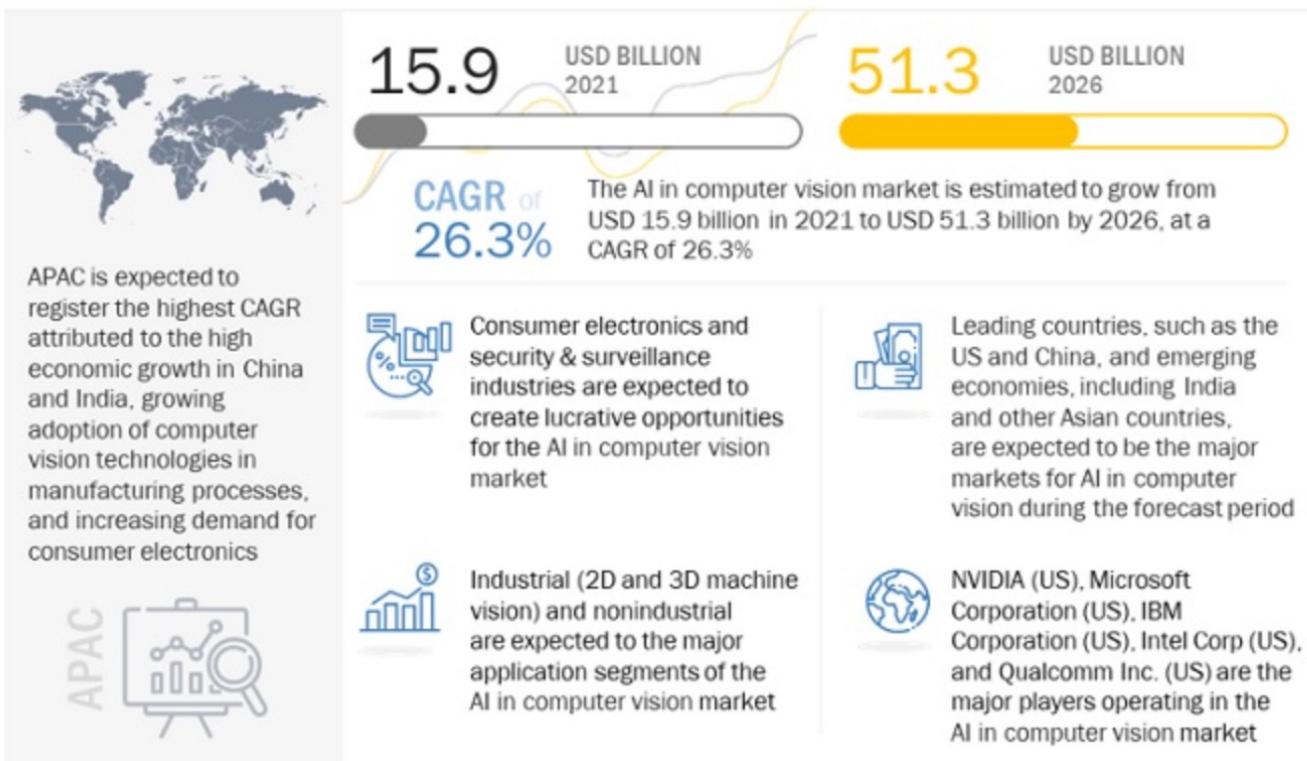
Smartphones and apps

Movies and media

Driving assistance

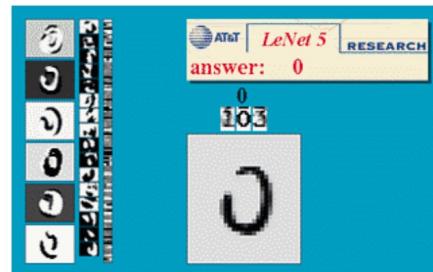
.....

Attractive Opportunities in AI in Computer Vision Market



<https://www.cs.ubc.ca/~lowe/vision.html>

Computer Vision application domains



(a)



(b)



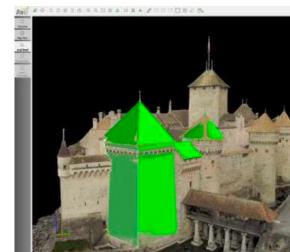
(c)



(d)



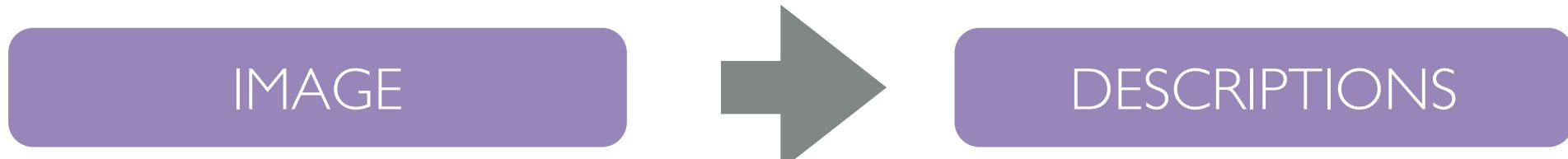
(e)



(f)

Figure 1.4 Some industrial applications of computer vision: (a) optical character recognition (OCR), <http://yann.lecun.com/exdb/lenet>; (b) mechanical inspection, <http://www.cognitens.com>; (c) warehouse picking, <https://covariant.ai>; (d) medical imaging, <http://www.clarontech.com>; (e) self-driving cars, (Montemerlo, Becker et al. 2008) © 2008 Wiley; (f) drone-based photogrammetry, <https://www.pix4d.com/blog/mapping-chillon-castle-with-drone>.

Computer Vision and Our MSc



extracting descriptions of the world from images

descriptions of what kind? (either *qualitative* or *quantitative*)

- *low level*: image features, image descriptors [DSIP]
- *geometric*: shape and position of object or relative distances in the real 3D world [CG + AR]
- *semantic*: what objects do I see? [CV+DL]
- *dynamic*: scene changes, objects velocities, actions, ... [(a little) in CV, a lot of possibilities as a master thesis]

Course Agenda (more details on Aulaweb)

- Images, filters and features refresh Image analysis
- Image similarity: scale invariant interest points, descriptors and matching Dynamics
- Motion analysis and optical flow Depth and geometry
- 3D vision Dynamics
- Unsupervised segmentation Depth and geometry

MID-TERM ASSIGNMENT

- Neural image representations and DL applications (overview) Semantics
- Instance detection and segmentation Semantics
- Semantic segmentation Semantics
- Humans in the image: Pose estimation and action recognition Semantics

UniGe

