Vision and Cognitive Systems

SCQ1097939 - LM CS,DS,CYB,PD

Course “wrap up”

Prof. Lamberto Ballan

University of Padova, A.Y. 2023/24 January 9, 2024

DIPARTIMENTO MATEMATICA
Department of Mathematics “Tullio Levi-Civita”

Who we are

Instructor

Prof. Lamberto Ballan
<lamberto.ballan@unipd.it>

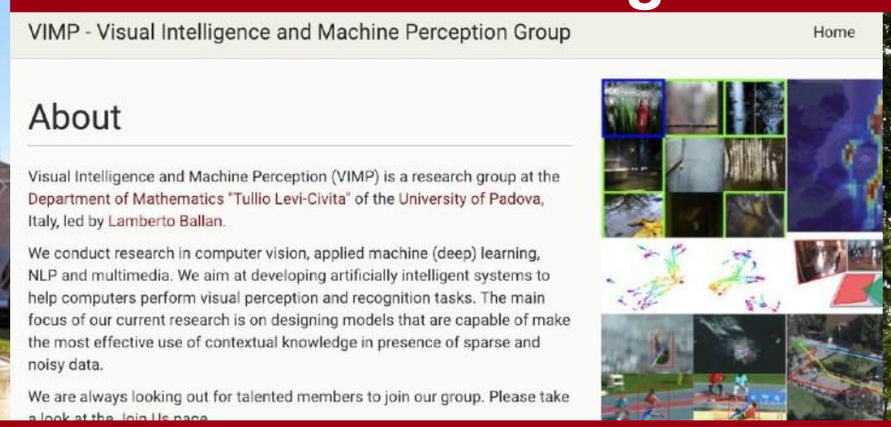
Teaching Assistant

Elena Izzo
<elena.izzo@phd.unipd.it>

Teaching Assistant

Filippo Ziliotto
<filippo.ziliotto@phd.unipd.it>

Visual Intelligence & Machine Perception (VIMP) *group*



http://vimp.math.unipd.it

VCS: Syllabus

- Introduction
  AI revolution; from human cognition to machine intelligence and cognitive systems
- Brief Introduction to Machine (Deep) Learning
  Classification; supervised learning; training and testing; bias
- (Early) Computer vision and image processing
  Image formation and filtering; feature detectors and descriptors
- Visual Recognition and Representation Learning in Vision
  Bag-of-Features; spatial pyramids and pooling; CNNs
- Beyond Image Classification and Supervised Learning
  R-CNN and segmentation; image captioning; sequential data in vision tasks; predictive vision; self-supervised learning

Our timeline

1966: Marvin Minsky assigns computer vision as an undergrad summer project

1970s: interpretation of synthetic worlds and carefully selected images

1980s: shift towards geometry and increased mathematical rigor

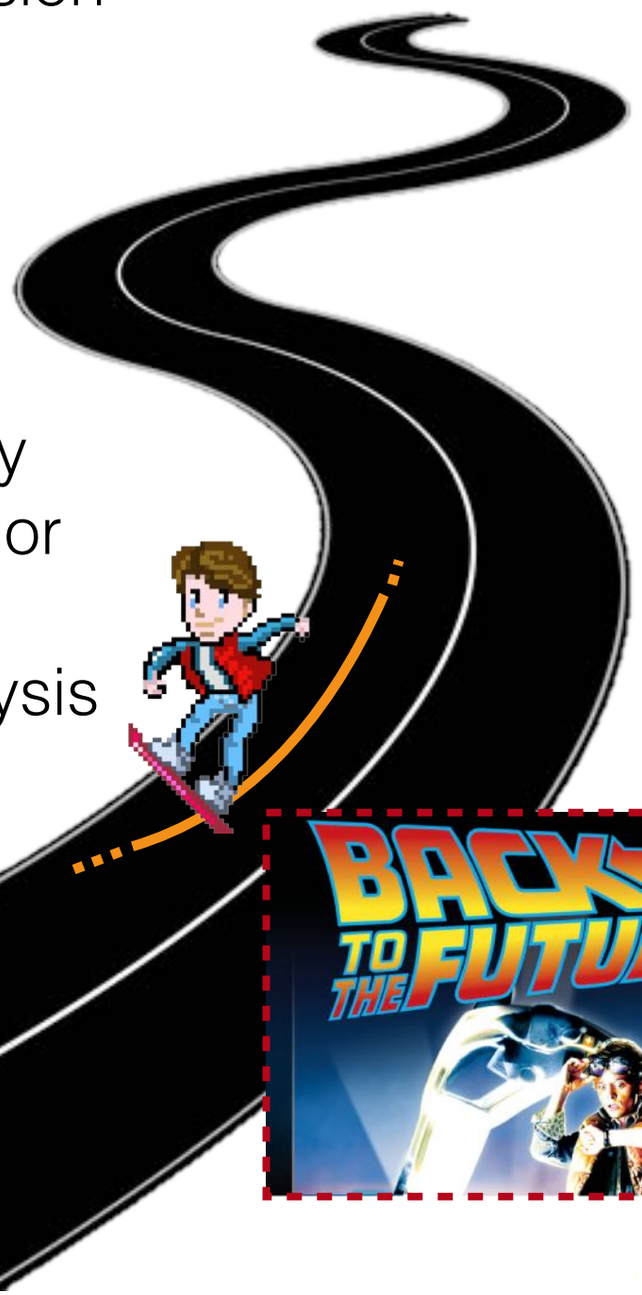
1990s: face recognition, statistical analysis

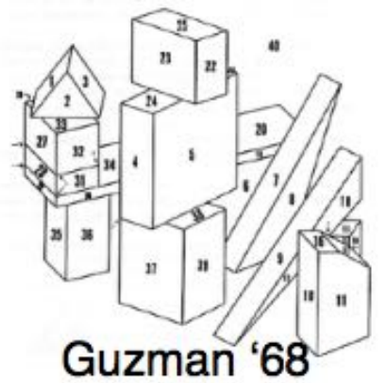
2000s: object recognition, categorization, annotated datasets available

2010s: large-scale visual recognition, visual intelligence

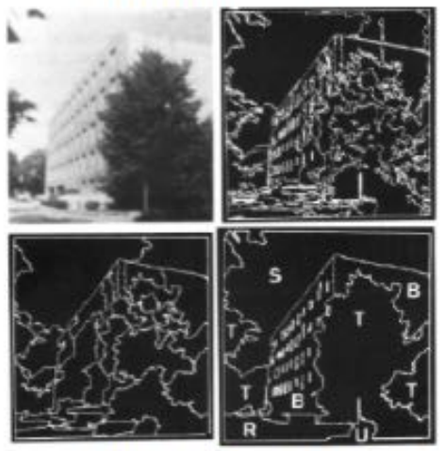
2020s: ???

(big/"foundation" models, …)





Guzman '68



Ohta Kanade '78

Turk and Pentland '91

Adapted from slide by O. Russakovsky

Human and Artificial Intelligence

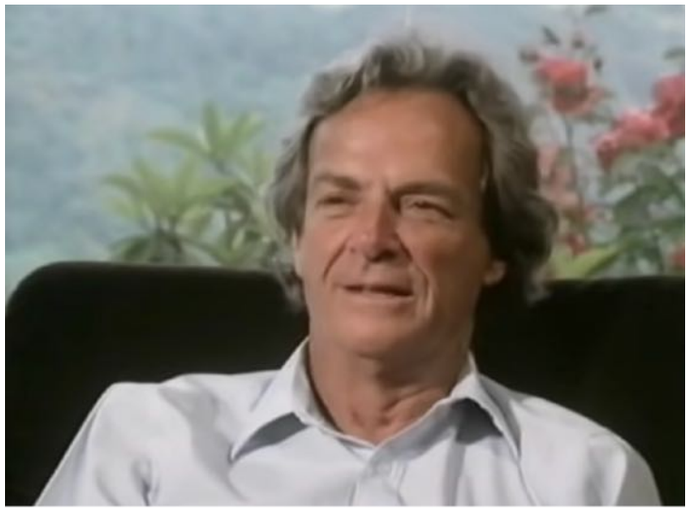
Audience Question:

(Re-recorded by me because the question is poorly audible)

Do you think there will ever be
a machine that will think like human beings
and be more intelligent than human beings?

Richard Feynman - September 26th, 1985

Human and Artificial Intelligence



First, intelligence has to be defined.

If you were to ask me: “are they better chess player than any human can possibly can be?”, yes, I’ll get you; someday! They are better chess players than most human being right now.

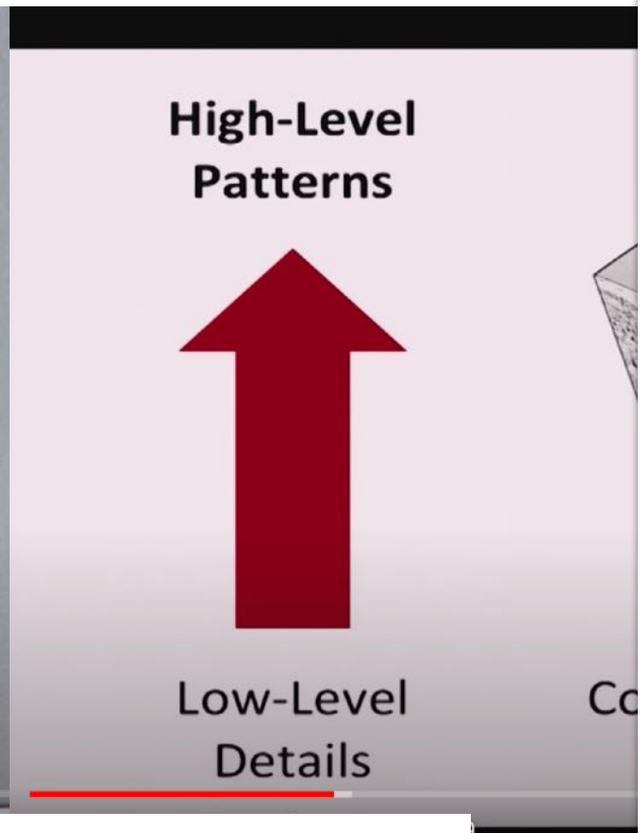
Kasparov *vs.* IBM Deep Blue (May 1997): 2.5-3.5

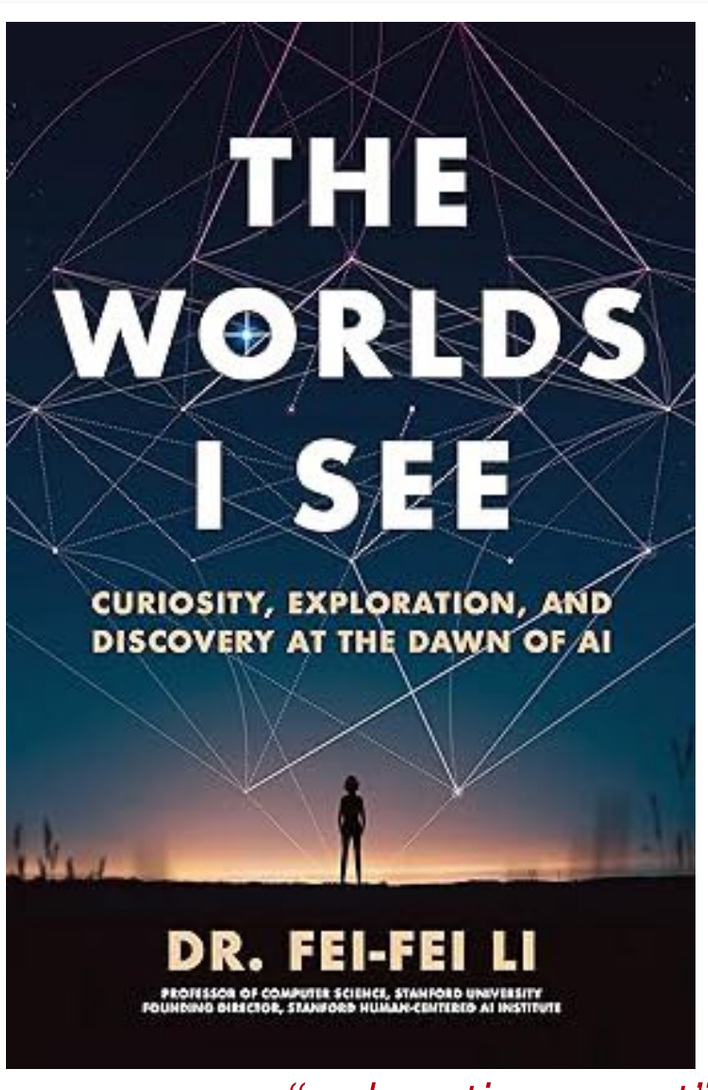
One of the things, by the way, that we always do, is that we want the machine to be better than anybody.

Richard Feynman - September 26th, 1985

Course “wrap up” / food for thought

- We will watch (and discuss) this talk:
  - “Human-centered AI: a Case for Cognitively Inspired Machine Intelligence”, Fei-Fei Li (Stanford)





“advertisement”

https://www.youtube.com/watch?v=JIBvfxg2iJ0

Course schedule

- Our last lectures:

L12	W11	Monday, 11 December 2023	12:30	Object Detection & Segmentation
MT	W11	Tuesday, 12 December 2023	12:30	Written exam ("compitino")
Lab7	W12	Monday, 18 December 2023	12:30	Lab7 - CNN
OA1	W12	Tuesday, 19 December 2023	12:30	Open discussion about MLP/CNN solutions, projects
L13	W13	Monday, 8 January 2024	12:30	Sequential data in Vision, Predictive Vision
L14	W13	Tuesday, 9 January 2024	12:30	Course "wrap up", "Cognitive Inspired AI"
L15	W14	Monday, 15 January 2024	12:30	Vision Transformers and SSL
L16	W14	Tuesday, 16 January 2024	12:30	Embodied AI and visual navigation

Visual Intelligence & Machine Perception

Visual Intelligence & Machine Perception (VIMP) *group*

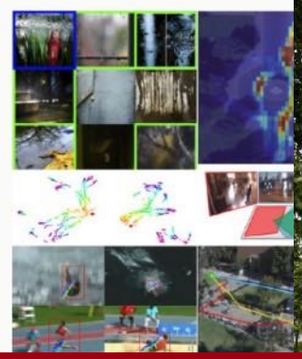
VIMP - Visual Intelligence and Machine Perception Group Home

About

Visual Intelligence and Machine Perception (VIMP) is a research group at the Department of Mathematics "Tullio Levi-Civita" of the University of Padova, Italy, led by Lamberto Ballan.

We conduct research in computer vision, applied machine (deep) learning, NLP and multimedia. We aim at developing artificially intelligent systems to help computers perform visual perception and recognition tasks. The main focus of our current research is on designing models that are capable of make the most effective use of contextual knowledge in presence of sparse and noisy data.

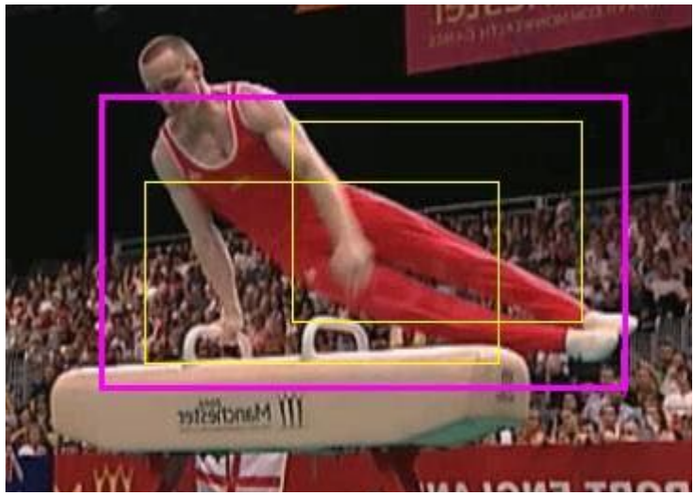
We are always looking out for talented members to join our group. Please take a look at the Join Us page

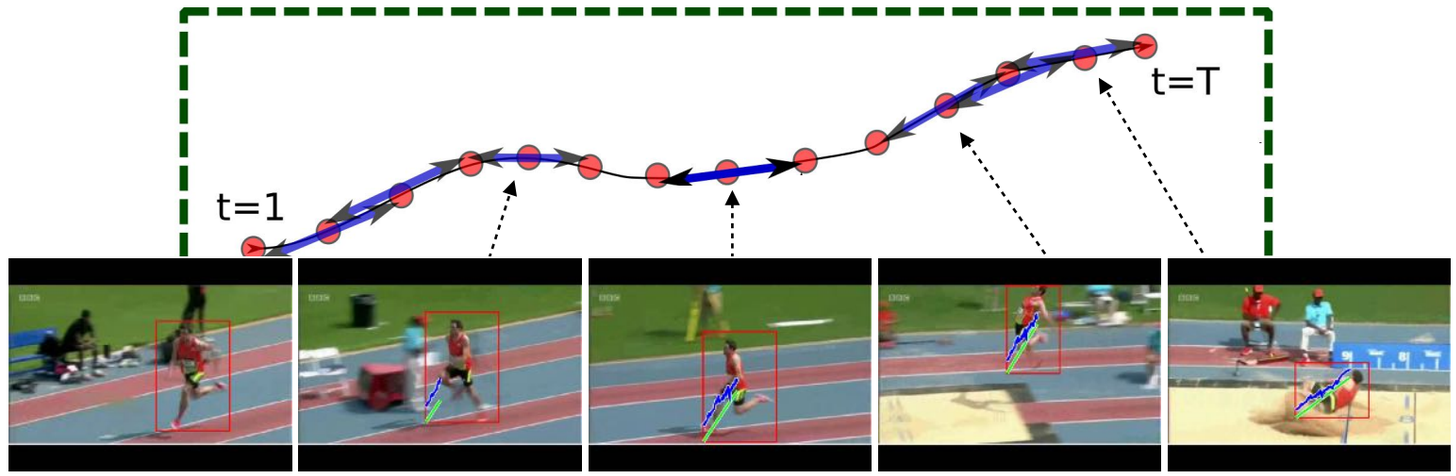


http://vimp.math.unipd.it

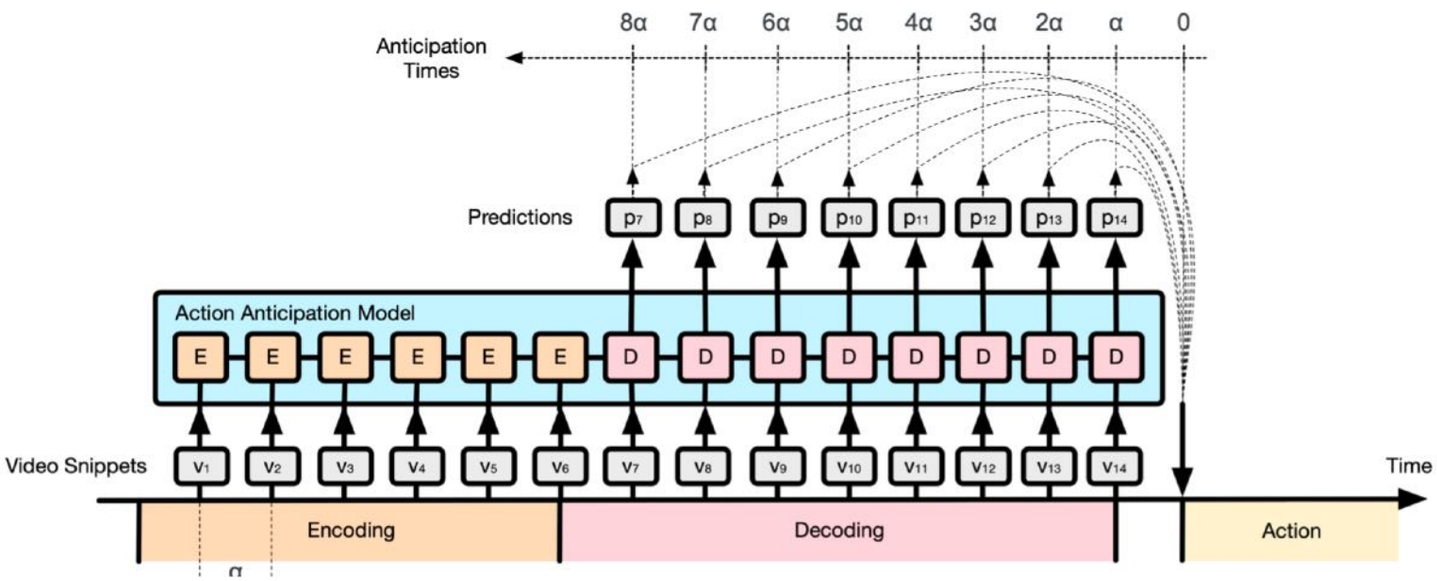
Action detection and anticipation

- Predicting action progress/completion in videos





- Action anticipation and early action detection



EPIC KITCHENS

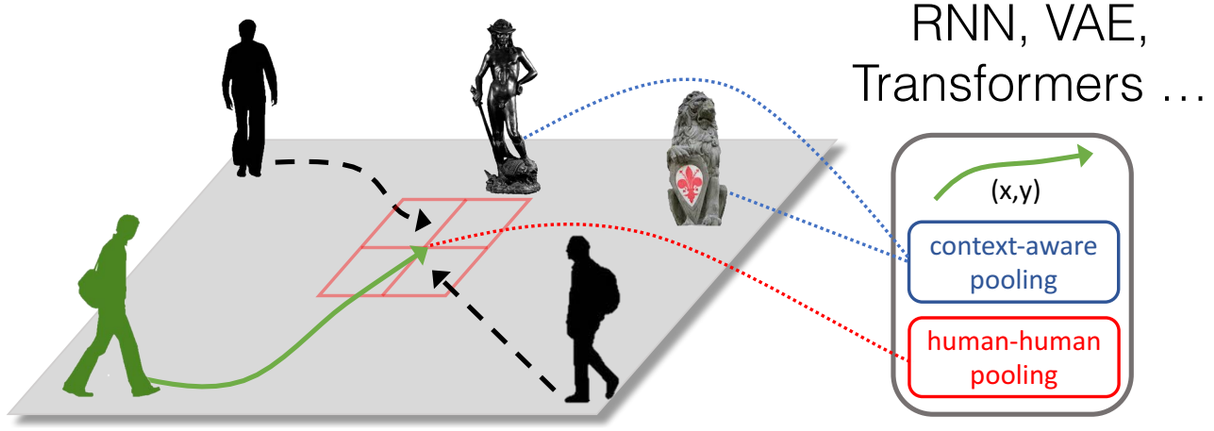
55 hours

2513 actions

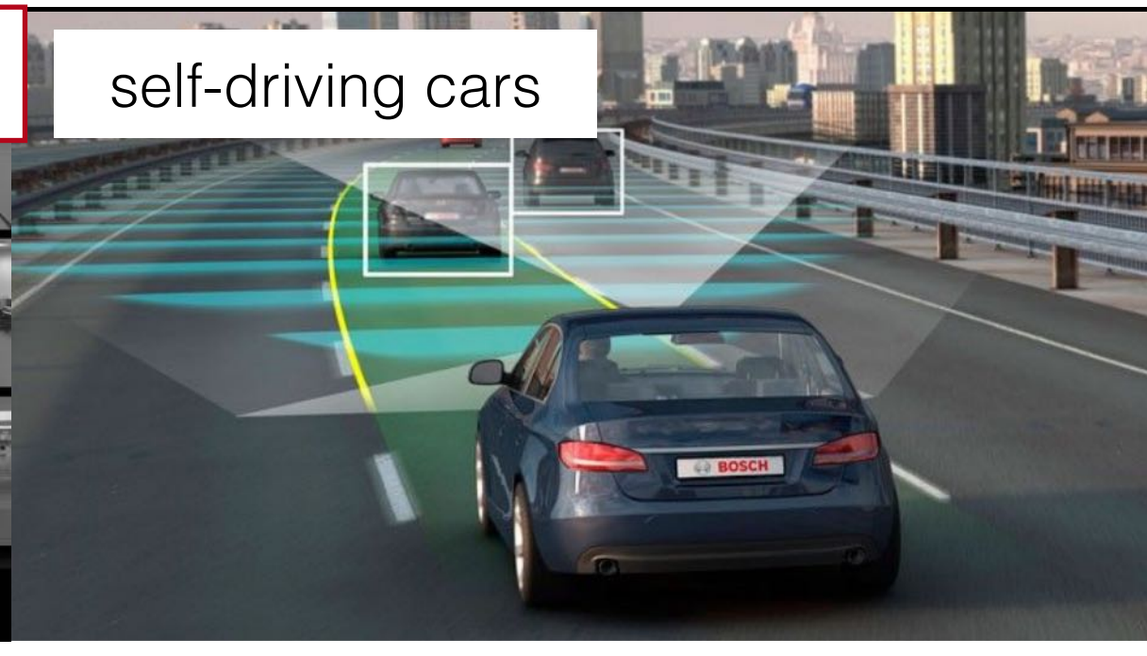
(125 verbs, 352 nouns)

Trajectory prediction, tracking

- Context-aware trajectory prediction in the wild

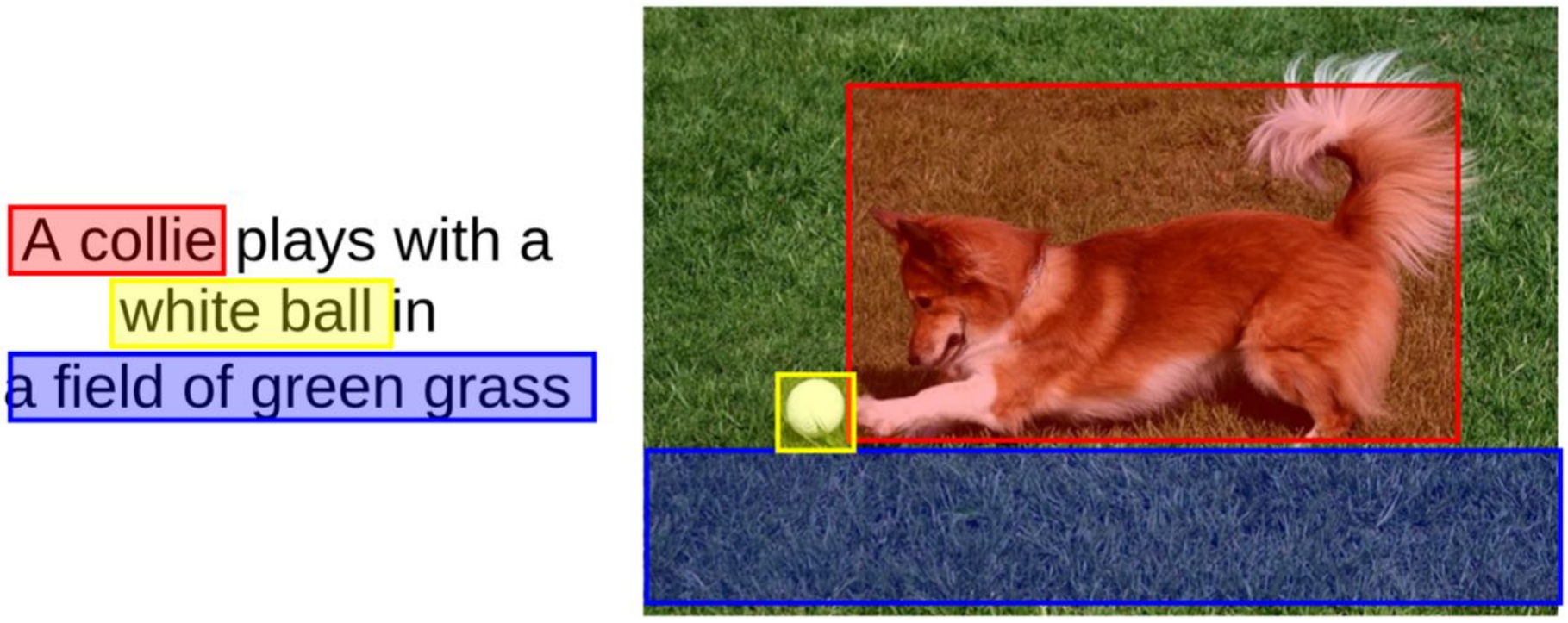


Applications in various domains…

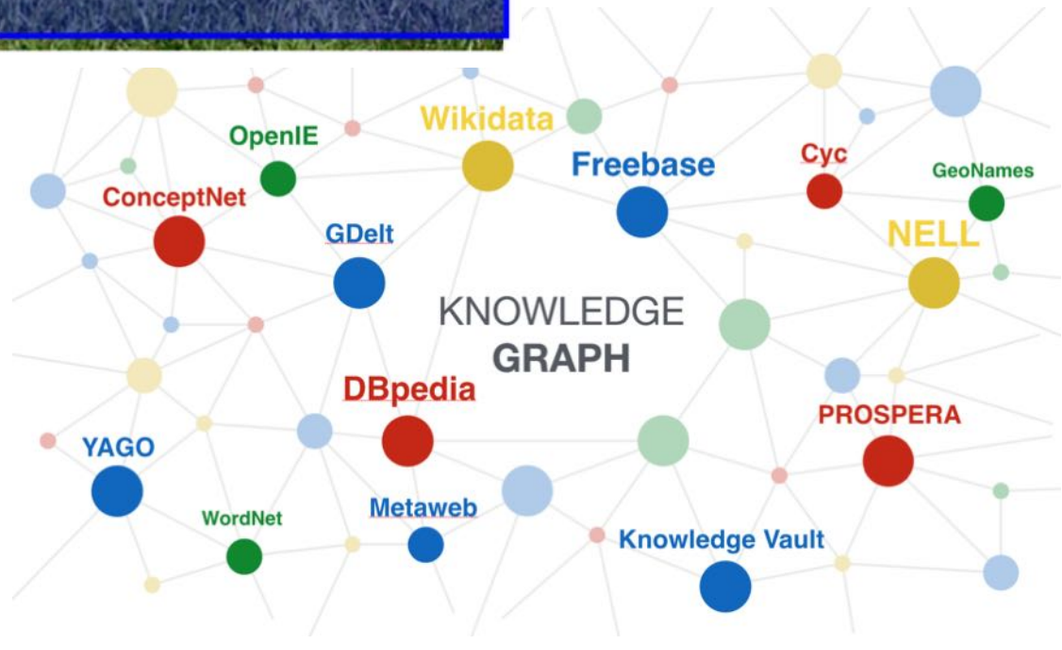


Vision & language, multimedia

- Visual-textual grounding and knowledge acquisition

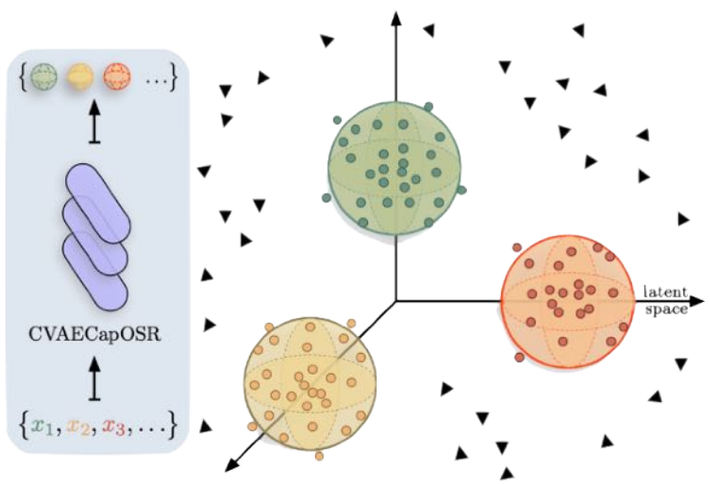


Visual-Textual-Knowledge
Entity Linking

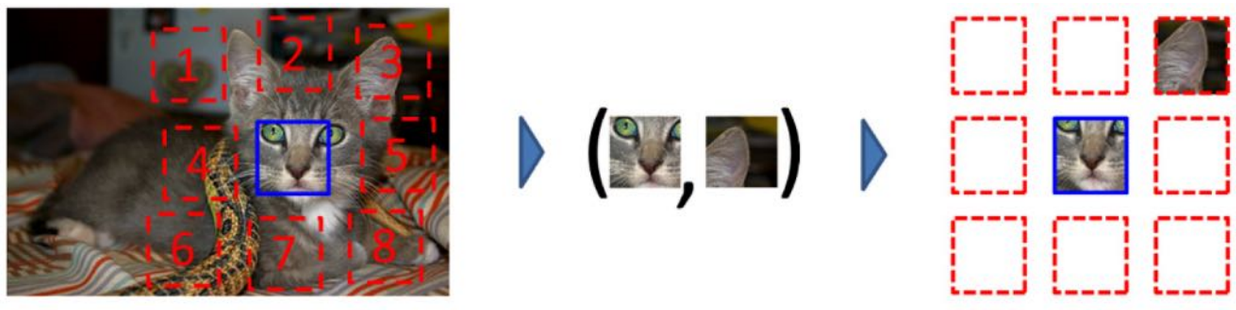


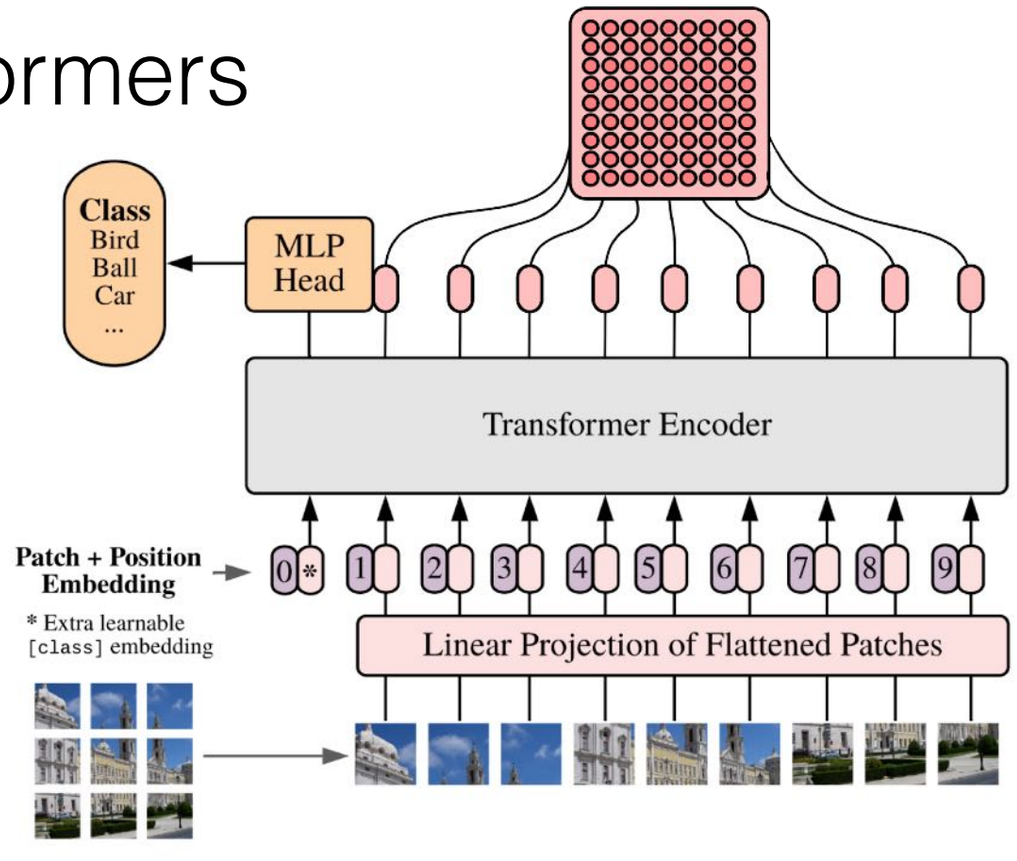
Incremental learning, self-supervision

- Open-world (open-set) recognition



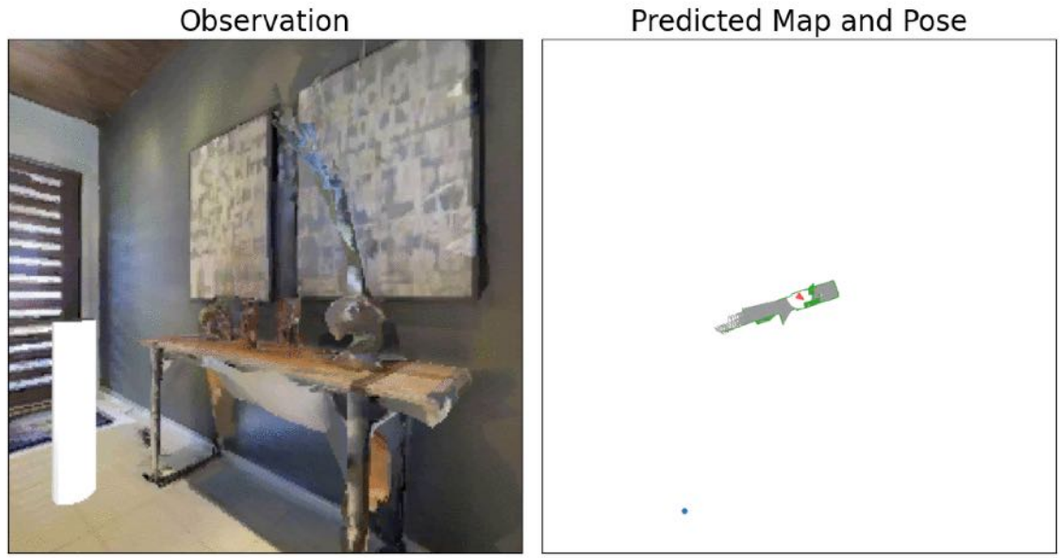
- Self-supervised vision transformers





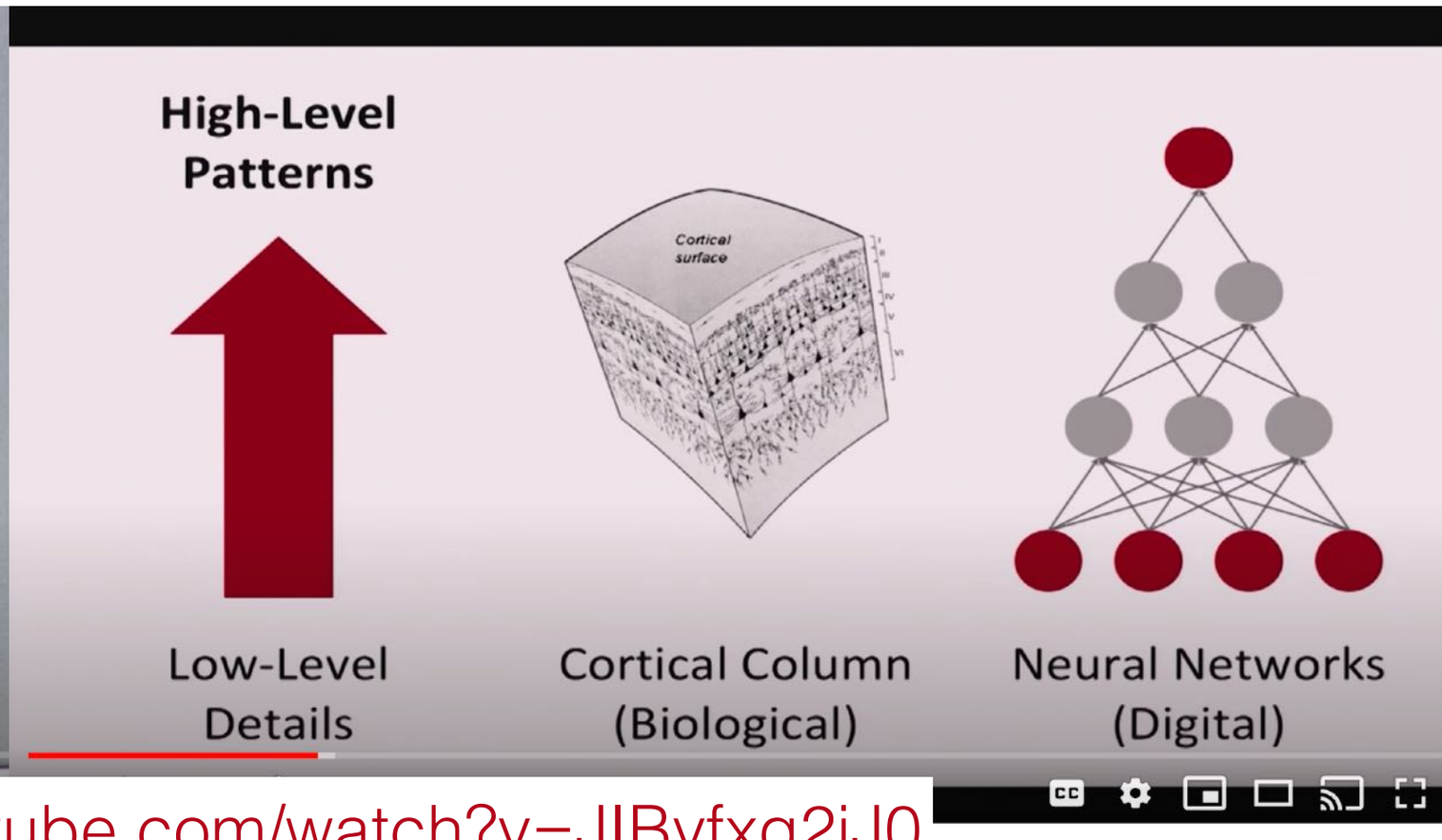
Embodied-AI & visual navigation

- Point Goal Navigation, Object Goal navigation, …



Course “wrap up” / food for thought

- We will watch (and discuss) this talk:
  - “Human-centered AI: a Case for Cognitively Inspired Machine Intelligence”, Fei-Fei Li (Stanford)



https://www.youtube.com/watch?v=JIBvfxg2iJ0

Contact

- **Office:** Torre Archimede, room 6CD3
- **Office hours** (ricevimento)**:** Friday 9:00-11:00

✉ lamberto.ballan@unipd.it

🏠 http://www.lambertoballan.net

🏠 http://vimp.math.unipd.it

@ twitter.com/lambertoballan