

CS-503

Visual intelligence : machines and minds

Zamir Amir

Cursus	Sem.	Type
Computer science	MA1, MA3	Opt.
Data Science	MA1, MA3	Opt.
SC master EPFL	MA1, MA3	Opt.

Language of teaching	English
Credits	5
Session	Winter
Semester	Fall
Exam	During the semester
Workload	150h
Weeks	14
Hours	4 weekly
Courses	2 weekly
Exercises	2 weekly
Number of positions	

Summary

The course will discuss classic material as well as recent advances in computer vision and machine learning relevant to processing visual data. The primary focus of the course will be on embodied intelligence and perception for active agents.

Content

Visual perception is the capability of inferring the properties of the external world merely from the light reflected off the objects therein. This is done beautifully well by simple (e.g. mosquitoes) or complex (e.g. humans) biological organisms. They can see and understand the complex environment around them and act accordingly; all done in an efficient and astonishingly robust way. Computer vision is the discipline of replicating this capability for machines. The progress in computer vision has brought about successful applications, such as face detection/recognition or handwriting recognition. However, a large gap to sophisticated perceptual capabilities, such as those exhibited by animals, remains.

The goal of this course is to discuss what is possible in computer vision today, and what is not. We will overview the basic concepts in computer vision and recent advances in machine learning relevant to processing visual data and active perception. For inspirations around what the missing capabilities are and how to approach them, we will turn to visual perception in biological organisms.

The course has a heavy emphasis on projects and hands-on experience. The course project will be around designing, implementing, and testing a solution to an open problem pertinent to visual perception. The students are encouraged to work in groups, self-propose a project that makes them excited, and go for ambitious yet feasible projects. The course staff will provide support throughout the semester with the projects. In the lectures, the students will learn about the principles of computer vision, the current limits, and the visual perception in humans and animals, which will help them with formulating and executing their course projects. In particular, the lectures will discuss:

1) A recap of basic computer vision concepts: classification, detection, segmentation, transformations, optical flow, 3D from X, etc, 2), What/why/how of visual representations. Supervised, self-supervised, unsupervised learning of representations. 3), Psychology of the visual system. 4), Physiology of the visual system. 5), Perception-action loop: active perception and embodied intelligence.

The course is of interest to MS/PhD students interested in research in computer vision, machine learning, and perceptual robotics as well as senior undergraduate students interested in gaining an advanced understanding of SOTA computer vision.

Keywords

Computer vision, machine learning, cognition, embodied intelligence, robotics, neural networks, AI.

Learning Prerequisites**Required courses**

Introduction to Machine Learning (CS-233) or Machine Learning (CS-433) or equivalent course on the basics of machine learning and deep learning.

Recommended courses

Computer vision (CS-442) or equivalent undergraduate course on the basics of computer vision.

Important concepts to start the course

- Python programming.
- Basics of deep learning and machine learning.
- Basics of probability and statistics.

Learning Outcomes

By the end of the course, the student must be able to:

- Define the basic concepts in computer vision, such as detection, segmentation, 3D from X, covered in the lectures.
- Explain the range of theories in psychology around visual perception, covered in the lectures.
- Design and implement computer vision/machine learning algorithms to address problems with real- world complexity.
- Design and implement proper evaluation pipelines for computer vision/machine learning algorithms to assess their performance in the real-world.
- Assess the limits and performance pitfalls of a given computer vision/machine learning algorithm, especially when facing real-world complexity

Transversal skills

- Write a scientific or technical report.
- Make an oral presentation.
- Assess progress against the plan, and adapt the plan as appropriate.
- Demonstrate the capacity for critical thinking

Teaching methods

Lectures. Lab sessions. Project Tutoring. Course Project.

Expected student activities

- In regard to the lectured material, the students are expected to study the provided reading material, actively participate in the class, engage in the discussions, and answer homework questions. In regard to the course project, the students are expected to formulate and implement an in-depth project and demonstrate continuous progress throughout the semester.

Assessment methods

- Project (70%) [Project proposal, Project checkpoint reports, Final project report and presentation]
- Homeworks (20%)
- Class attendance and engagement (10%)

Supervision

Office hours	Yes
Assistants	Yes
Forum	Yes

Resources**Bibliography**

- Vision Science: Photons to Phenomenology, Steven Palmer, 1999.
- The Ecological Approach to Visual Perception, Jame Gibson, 1979.
- Computer Vision: Algorithms and Applications, Richard Szeliski, 2020

Ressources en bibliothèque

- [Vision science : photons to phenomenology / Palmer](#)
- [The Ecological Approach to Visual Perception /Gibson](#)
- [Computer Vision: Algorithms and Applications / Szeliski \(2022? ; online drafts\)](#)

Notes/Handbook

The reference reading of different lectures will be from different books (main ones listed above) and occasionally from papers. Resources will be provided in class. Full-text books are not mandatory.