EPFL

CS-503 **Visual intelligence : machines and minds**
Zamir Amir

Cursus	Sem.	Type
Civil & Environmental Engineering		Obl.
Computer science	MA2, MA4	Opt.
Data Science	MA2, MA4	Opt.
Digital Humanities	MA2, MA4	Opt.
Minor in Imaging	E	Opt.
Neuro-X minor	E	Opt.
Neuro-X	MA2, MA4	Opt.
SC master EPFL	MA2, MA4	Opt.

Language of teaching	English
Credits	6
Session	Summer
Semester	Spring
Exam	During the semester
Workload	180h
Weeks	14
Hours	**4 weekly**
Lecture	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 -- with a primary focus on embodied intelligence and vision 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 inspiration around the missing capabilities and how to approach them, we will turn to visual perception in biological organisms.

The course includes lectures and projects. There will be a heavy emphasis on the *projects* and *hands-on experience*. The course project will be around designing, implementing, and testing a solution to a (preferably open) problem pertinent to visual perception. The students are encouraged to work in groups, self-propose a project that excites them, 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 their course projects. In particular, the lectures will discuss the following:

1. A recap of basic computer vision concepts: classification, detection, grouping, image transformations, optical flow, 3D from X, etc., and recent successful neural network architectures, such as Transformers.
2. Psychology/physiology of the visual system.
3. Perception-action loop: active perception and embodied vision.

The course interests masters/PhD students interested in research in computer vision, machine learning, and perceptual robotics, as well as senior undergraduate students interested in understanding state-of-the-art computer vision.

Keywords

Computer vision, Machine learning, Embodied intelligence, Robotics, Cognition, Neural networks, AI.

Learning Prerequisites

Required courses

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

• Deep Learning (EE-559) or Artificial Neural Networks (CS-456) or equivalent course on the basics of deep learning.

Recommended courses

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

Important concepts to start the course

• Deep learning and machine learning.
• Python programming.
• 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, demonstrate continuous progress throughout the semester, and provide a final written report and presentation.

Assessment methods

• Project (70%) [distributed over the project proposal, milestone reports, final report and presentation]

• Homeworks (30%)

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.
• Animal Eyes, Michael Land and Dan-Eric Nilsson, 2012.

Ressources en bibliothèque

• Animal Eyes / Land
• An immense world / Yong (added)
• Vision Science / Palmer
• Ecological Approach to Visual Perception / Gibson
• Computer Vision / Szeliski

Notes/Handbook

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

Moodle Link

• https://go.epfl.ch/CS-503