AI: COMPUTER VISION

Grado en Computación e Inteligencia Artificial / Bachelor in Computer Science and Artificial Intelligence BCSAI SEP-2025 AICV-CSAI.3.M.A

Area Computer Science
Number of sessions: 30
Academic year: 25-26
Degree course: THIRD
Number of credits: 6.0
Semester: 2º
Category: COMPULSORY
Language: English

Professor: **RUBÉN SÁNCHEZ GARCÍA**

E-mail: rsanchezgarcia@faculty.ie.edu

Dr. Ruben Sanchez-Garcia holds a BSc in Biotechnology, another in Computer Science, and an MSc in Biophysics. He completed his PhD in 2020 at the Spanish National Center of Biotechnology (CNB-CSIC) and the Autonomous University of Madrid (UAM) under the supervision of Professor Jose Maria Carazo and Dr. Joan Segura. Following his doctoral studies, he joined the Department of Statistics at the University of Oxford as a postdoctoral researcher until the 2024/25 academic year, when he joined IE University. Simultaneously, since 2022, he was part of the Sustaining Innovation Postdoctoral Researcher program at Astex Pharmaceuticals in Cambridge, UK.

Dr. Sanchez-Garcia's research focuses on developing machine learning algorithms for structural biology, with a particular emphasis on deep learning methods for cryogenic electron microscopy and structure-based drug discovery. He has co-authored more than 30 peer-reviewed publications and developed several well-known algorithms, including BIPSPI and DeepEMhancer.

Office Hours

Office hours will be on request. Please contact at:

rsanchezgarcia@faculty.ie.edu

SUBJECT DESCRIPTION

Vision plays an essential role in human perception, and therefore understanding and automating human tasks that rely in vision is a relevant scientific challenge today, with countless innovative applications in industrial automation, autonomous vehicles, precision agriculture, e-health, etc.

Computer Vision is a key engineering field that adresses this challenge by attempting to obtain a high-level understanding of the world around us, captured in digital images and videos.

This course offers an introduction to the most relevant topics in Computer Vision and the engineering aspects related. It integrates contents related to traditional Computer Vision techniques as well as modern approaches relying on Machine Learning.

LEARNING OBJECTIVES

The objective of this course is to provide a solid basis that will help students build sophisticated applications and manage complex projects related to Computer Vision.

Students will obtain a solid understanding of the essence of digital images and videos, from formation and codification to processing and storage. The course will also help students understand the most relevant problems in Computer Vision today and learn the best techniques known today to address them.

The following skills will be developed during this course:

- Ability to correctly pose, discuss and solve most common Computer Vision problems.
- Fluency with the development of Computer Vision applications in Python.
- Teamwork skills: communication, coordination and leadership.

TEACHING METHODOLOGY

In this course the student will acquire knowledge mainly through the lectures and will put it in practice individually through coding exercises and an individual project.

Learning Activity	**Weighting**	**Estimated time a student should dedicate to prepare for and participate in**
Lectures	43.3 %	65.0 hours
Exercises in class, Asynchronous sessions, Field Work	40.0 %	60.0 hours
Group work	10.0 %	15.0 hours
Individual studying	6.7 %	10.0 hours
TOTAL	100.0 %	150.0 hours

AI POLICY

In today's world, generative artificial intelligence (GenAI) is changing how we work, study and, in general, how we get things done. However, in the context of this course, the use of GenAI is not permitted, unless it is otherwise stated by the instructor. The use of GenAI tools would jeopardize the students' ability to acquire fundamental knowledge or skills of this course.

If a student is found to have used AI-generated content for any form of assessment, it will be considered academic misconduct, and the student might fail the respective assignment or the course.

PROGRAM

SESSION 1 (LIVE IN-PERSON)

Course presentation

- Definition and relationship with other disciplines
- Main concepts
- Historical background
- State of the art

SESSION 2 (LIVE IN-PERSON)

Computer Vision Engineering (I)

- Basic illumination techniques
- Illumination sources
- Camera elements
- Fisheye lenses and Omnidirectional cameras
- Optics fundamentals
- Sensors
- Examples of applications

SESSION 3 (LIVE IN-PERSON)

Computer Vision Engineering (II)

- Color perception
- Color spaces
- Image compression
- First steps with OpenCV

SESSION 4 (LIVE IN-PERSON)

Camera configuration and streaming

- Acquisition methods
- Focus
- Iris
- Exposure, White balance, Gamma
- Setting up an RSTP server

SESSION 5 (LIVE IN-PERSON)

Introduction to projective geometry

- Projective space Pn
- Projective space P2
- Projective space P3

SESSION 6 (LIVE IN-PERSON)

Image transformations

- Scaling
- Translation
- Rotation
- Affine
- Perspective

SESSION 7 (LIVE IN-PERSON)

Image acquisition model

- Lenses models: pinhole, thin and thick
- Projection models
- Intrinsic and extrinsic parameters
- Coordinate systems involved

SESSION 8 (LIVE IN-PERSON)

Three-dimensional vision I

- Human 3D vision
- Stereoscopic vision
- Structured light projection
- Structured light
- Time of flight

Three-dimensional vision II

- Depth from Focus
- Shape given shadows and texture
- Optical flow
- Applications in motion detection and tracking

SESSION 9 (LIVE IN-PERSON)

Multiple view geometry

- Single camera model
- Dual-camera model
- Homography
- Epipolar geometry
- Structure from motion

SESSION 10 (LIVE IN-PERSON)

Multiple view geometry

- Single camera model
- Dual-camera model
- Homography
- Epipolar geometry

- Structure from motion

SESSION 11 (LIVE IN-PERSON)

Depth cameras

- Stereoscopic camera configuration and capture
- Structured light camera configuration and capture

SESSION 12 (LIVE IN-PERSON)

Camera calibration I

- Introduction
- Tools
- Zhang's method
- Calibration process issues

SESSION 13 (LIVE IN-PERSON)

Camera calibration II

- Pattern image acquisition
- Running the calibration process
- Checking reprojection error

SESSION 14 (LIVE IN-PERSON)

Mid-term exam

SESSION 15 (LIVE IN-PERSON)

Pixelwise operations

- Pixelwise operations
- Pixel transformations
- Histogram equalization
- Binarization

SESSION 16 (LIVE IN-PERSON)

Local operations

- Introduction
- Separable filters
- Linear filtering examples
- Bandpass and orientable filters
- Other local operators

SESSION 17 (LIVE IN-PERSON)

Global operations

- Fourier transform
- Bidimensional Fourier transform
- DFT interpretation
- Other global transformations

SESSION 18 (LIVE IN-PERSON)

Salient feature detection

- Introduction
- Definitions
- Surface curvature
- Curvature-based method
- Harris and Stephens method
- Gradient-based methods

SESSION 19 (LIVE IN-PERSON)

Intro to object recognition

- Introduction
- Edge detection-based techniques
- Image thresholding
- Region-based algorithms

SESSION 20 (LIVE IN-PERSON)

Image matching techniques

- Template matching
- Feature matching + homography

SESSION 21 (LIVE IN-PERSON)

Convolutional neural networks

- Introduction
- Applications
- Functional perspective

SESSION 22 (LIVE IN-PERSON)

Convolutional neural networks II

- Deep learning frameworks
- Implementing a CNN from scratch

SESSION 23 (LIVE IN-PERSON)

Convolutional neural networks III

Advanced architectures

- AlexNet

- VGG
- GoogLeNet
- ResNet
- Unet

SESSION 24 (LIVE IN-PERSON)

Recurrent Neural Networks

- Introduction
- Gated Recurrent Units (GRU) networks
- Long Short Term Memory (LSTM) networks
- Image captioning with attention

SESSION 25 (LIVE IN-PERSON)

Detection and segmentation using Deep Learning

- Semantic segmentation
- Classification + Localization
- Object detectionInstance segmentation

SESSION 26 (LIVE IN-PERSON)

Detection and segmentation using Deep Learning II

Training YoLo for object detection

SESSION 27 (LIVE IN-PERSON)

Generative models I

- Unsupervised Learning
- Variational Autoencoders (VAE)
- PixelRNN and PixelCNN

SESSION 28 (LIVE IN-PERSON)

Generative models II

- Generative Adversarial Networks (GANs)
- Diffusion models

SESSION 29 (LIVE IN-PERSON)

Transformer architecture

- Introduction to Self-Attention and Transformers
- Vision Transformers (ViT): Architecture and Workflow
- Hybrid Models: Combining CNNs and Transformers
- Applications in Computer Vision: Classification, Detection, and Generation

SESSION 30 (LIVE IN-PERSON)

IE UNIVERSITY

Final exam

EVALUATION CRITERIA

The course will be graded based on intermediate tests in which students will have to solve Computer Vision problems and communicate their progress through intermediate deliverables plus ellaborate a final project.

criteria	percentage	Learning Objectives	Comments
Final Exam	35 %		Minimum passing grade of 3.5
Individual work	15 %		
Class Participation	5 %		
Intermediate tests	30 %		
Group Work	15 %		

RE-SIT / RE-TAKE POLICY

Each student has four chances to pass any given course distributed over two consecutive academic years: ordinary call exams and extraordinary call exams (re-sits) in June/July.

Students who do not comply with the 80% attendance rule during the semester will fail both calls for this Academic Year (ordinary and extraordinary) and have to re-take the course (i.e., re-enroll) in the next Academic Year.

Evaluation criteria:

- Students failing the course in the ordinary call (during the semester) will have to re-sit the exam in June / July (except those not complying with the attendance rule, who will not have that opportunity and must directly re-enroll in the course on the next Academic Year).
- The extraordinary call exams in June / July (re-sits) require your physical presence at the campus you are enrolled in (Segovia or Madrid). There is no possibility to change the date, location or format of any exam, under any circumstances. Dates and location of the June / July re-sit exams will be posted in advance. Please take this into consideration when planning your summer.
- The June / July re-sit exam will consist of a comprehensive exam. Your final grade for the course will depend on the performance in this exam only; continuous evaluation over the semester will not be taken into consideration. Students will have to achieve the minimum passing grade of 5 and can obtain a maximum grade of 8.0 (out of 10.0) – i.e., “notable” in the in the re-sit exam.
- Retakers: Students who failed the subject on a previous Academic Year and are now re-enrolled as re-takers in a course will be needed to check the syllabus of the assigned professor, as well as contact the professor individually, regarding the specific evaluation criteria for them as retakers in the course during that semester (ordinary call of that Academic Year).

  The maximum grade that may be obtained in the retake exam (3rd call) is 10.0.

After ordinary and extraordinary call exams are graded by the professor, you will have a possibility to attend a review session for that exam and course grade. Please be available to attend the session in order to clarify any concerns you might have regarding your exam. Your professor will inform you about the time and place of the review session. Any grade appeals require that the student attended the review session prior to appealing.

Students failing more than 18 ECTS credits in the academic year after the June-July re-sits will be asked to leave the Program. Please, make sure to prepare yourself well for the exams in order to pass your failed subjects.

In case you decide to skip the opportunity to re-sit for an exam during the June / July extraordinary call, you will need to enroll in that course again for the next Academic Year as a re-taker and pay the corresponding extra cost. As you know, students have a total of four allowed calls to pass a given subject or course, in order to remain in the program.

BIBLIOGRAPHY

Recommended

- Hartley, R.I. and Zisserman, A.. *Multiple View Geometry.* ISBN 9780521540513 (Digital)
- Forsyth & Ponce. *Computer Vision: A modern approach.* ISBN 0130851981 (Digital)
- Livingstone, M.. *Vision and Art: The Biology of Seeing.* ISBN 9780810904064 (Digital)
- Richard Szeliski. *Computer Vision Algorithms and Applications.* ISBN 9783030343743 (Printed)
- Christopher M Bishop. *Pattern recognition and machine learning.* ISBN 9780387310732 (Printed)
- Max Drummy et al. *Dive into deep learning: tools for engagement.* ISBN 9781544361376 (Printed)

BEHAVIOR RULES

Please, check the University's Code of Conduct here. The Program Director may provide further indications.

The use of mobile phones in classroom is strictly prohibited.

No eating, drinking, smoking or vaping is allowed inside the classroom.

Students arriving 5 minutes after the start time of the lecture might not be admitted.

Please let your instructor know if you must leave the classroom before the end of the lecture. Students might not allowed to leave the classroom once the lecture has started, unless explicitly authorised by the instructor.

ATTENDANCE POLICY

Please, check the University's Attendance Policy here. The Program Director may provide further indications.

ETHICAL POLICY

Please, check the University's Ethics Code here. The Program Director may provide further indications.