Course guide
230726 - CVDL - Computer Vision with Deep Learning

Last modified: 06/06/2025

Unit in charge: Barcelona School of Telecommunications Engineering
Teaching unit: 739 - TSC - Department of Signal Theory and Communications.

Degree: MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Optional subject).
MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Optional subject).

Academic year: 2025 **ECTS Credits:** 5.0 **Languages:** English

LECTURER

Coordinating lecturer: JOSEP RAMON MORROS RUBIO

Others:

PRIOR SKILLS

If you have not previously taken DLAI, you will not be able to enroll in CVDL. If you believe that you have the necessary knowledge to take CVDL, you must contact the CVDL coordinator so that he or she can assess your background and authorize an exception to the requirement. If the exception to the requirement is authorized, you must make an application that you can request:
- Before enrollment, during the enrollment exception application period, so that CVDL appears in your enrollment potential, and you can enroll whenever there are available places.
or
- After enrollment, during the enrollment modification period, to include CVDL in the enrollment, as long as there are still available places
Both enrollment exception applications and enrollment modification applications will be accepted by the MATT Head of Studies provided that the coordinator has communicated the authorization of the exception to the DLAI requirement.

You should also have the following previous knowledge to take the course:
- Image processing: pixels, color spaces, histograms, frequency domain representation
- Digital signal processing: linear filters, convolution
- Vector and matrix algebra

Notions of python are useful, but these are easily obtained during the course.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
1. Ability to apply information theory methods, adaptive modulation and channel coding, as well as advanced techniques of digital signal processing to communication and audiovisual systems.
2. Ability to integrate Telecommunication Engineering technologies and systems, as a generalist, and in broader and multidisciplinary contexts, such as bioengineering, photovoltaic conversion, nanotechnology and telemedicine.

Transversal:
3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

TEACHING METHODOLOGY

- Lectures
- Practical work
- Individual work (distance)
- Exercises
- Mid and final term exams

LEARNING OBJECTIVES OF THE SUBJECT

Learning objectives of the subject:

The aim of this course is to provide an overview of concepts and applications of computer vision, with both classic and Deep Learning methods. We will introduce low level techniques such as feature extraction and matching, edge detection, cameras and projection models and optical flow; mid-level topics such as video segmentation and feature tracking; high level methods such as object tracking. Then, examples of application will be shown, such as face and object recognition.

Learning results of the subject:

- Ability to understand and use techniques for image and video analysis: feature extraction, video segmentation, stereo, object detection.
- Ability to use computer vision algorithms to implement high-level applications.

STUDY LOAD

Type	Hours	Percentage
Self study	86,0	68.80
Hours large group	39,0	31.20

Total learning time: 125 h

CONTENTS

1. Introduction

Description:
- Motivation, types of problems in CV
- Image formation, perception, 3D sensors

Full-or-part-time: 7h
Theory classes: 3h
Self study : 4h

2. Image Structure

Description:
- Color, texture, filtering, and contours
- Detection and representation of interesting points and 'blobs'
- Modeling: RANSAC, Hough transform
- Saliency maps

Full-or-part-time: 29h
Theory classes: 9h
Guided activities: 4h
Self study : 16h

3. Stereo and 3D applications

Description:
- Single-camera geometry, camera calibration
- Epipolar geometry, homography
- Camera pose estimation and sensor registration using deep learning

Full-or-part-time: 30h
Theory classes: 9h
Guided activities: 4h
Self study : 17h

4. Video tracking

Description:
- Optical flow: Lucas-Kanade, Shi-Tomasi, Deep Learning methods
- Bayesian tracking: Kalman, Particle filters
- Deep Learning tracking methods

Full-or-part-time: 29h
Theory classes: 9h
Guided activities: 4h
Self study : 16h

5. Detection and recognition

Description:
- Introduction to visual recognition. Review of machine learning Deep learning and convolutional neural networks
- Image classification: Bag of words model. Image classification using CNNs
- Object detection: Sliding windows and local features. Object detection using CNNs
- Object segmentation: Semantic segmentation. Instance segmentation

Full-or-part-time: 30h
Theory classes: 9h
Guided activities: 4h
Self study : 17h

ACTIVITIES

EXERCISES

Description:
- Detecting contours and modeling shapes: Canny, Hough, Ransac, DL
- Finding correspondences between images: Harris, SIFT
- Fundamental matrix estimation
- Application of homography: panorama creation
- Object detection & recognition

Full-or-part-time: 6h
Self study: 6h

EXTENDED ANSWER TEST

Description:
Mid-term examination.

Full-or-part-time: 2h
Theory classes: 2h

EXTENDED ANSWER TEST

Description:
Second term examination

Full-or-part-time: 2h
Theory classes: 2h

GRADING SYSTEM

First-term examination: 40%
Second term examination: 40%
Laboratory/Exercises assessments: 20%

BIBLIOGRAPHY

Basic:
- Szeliski, R. Computer vision: algorithms and applications. 2nd ed. Cham, Switzerland: Springer, 2022. ISBN 9783030343712.
- Forsyth, D.A.; Ponce, J. Computer vision: a modern approach [on line]. 2nd ed. Boston, Mass.: Pearson Education, 2012 [Consultation: 09/09/2020]. Available on: https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=5173504. ISBN 9780273764144.

Complementary:
- Hartley, R.; Zisserman, A. Multiple view geometry in computer vision. 2nd ed. Cambridge: Cambridge University Press, 2003. ISBN 0521540518.
- Wang, Y.; Ostermann, J.; Zhang, Y.-Q. Video processing and communications. Upper Saddle River: Prentice Hall, 2002. ISBN 9788131733646.
- Hanjalic, A. Content-based analysis of digital video [on line]. Boston: Kluwer Academic, 2004 [Consultation: 29/07/2013]. Available on: http://link.springer.com/book/10.1007/b106003/page/1. ISBN 978-1402081149.

RESOURCES

Other resources:
Google Colab