Course guide
230360 - DLCV - Deep Learning for Computer Vision

Last modified: 06/05/2019

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

Degree: **Academic year:** 2019 **ECTS Credits:** 2.5
Languages: English

LECTURER

Coordinating lecturer: Xavier Giró i Nieto

Others: Xavier Giró i Nieto, Elisa Sayrol, Amaia Salvador, Kevin McGuinness and Eva Mohedano

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE1. 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.

Transversal:
CT3. 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.

CT4. 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.

CT5. 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
- Application classes
- Group work
- Group work (distance)
- Short answer test (Test)

LEARNING OBJECTIVES OF THE SUBJECT

The aim of this course is to train students in methods of deep learning for computer vision. Convolutional neural networks (convnets) will be presented and analyzed in detail to understand the potential of these state of the art tools in visual pattern recognition. Engineering tips and scalability issues will be addressed to solve tasks such as image classification, object detection or automatic textual captioning. Hands-on sessions will provide development skills so that attendees can solve a selected task in an open scientific benchmark and, if successful, submit their results.

STUDY LOAD

Type	Hours	Percentage
Hours small group	4,0	6.40
Hours large group	16,0	25.60
Self study	42,5	68.00

Total learning time: 62.5 h

CONTENTS

1. Convolutional Neural Networks

Description:
- Architecture: Forward and recurrent networks.
- Backpropagation
- Layer Visualization.
- Memory and computational requirements.
- Best practices.
- Fine-tunning

Full-or-part-time: 18h 30m
Theory classes: 4h
Laboratory classes: 3h
Guided activities: 4h
Self study : 7h 30m

2. Applications

Description:
- Image retrieval and classification
- Face and object detection/recognition.
- Semantic segmentation
- Saliency prediction
- Image captioning
- Multimodal fusion

Full-or-part-time: 44h
Theory classes: 4h
Practical classes: 7h
Laboratory classes: 1h
Guided activities: 15h
Self study : 17h

ACTIVITIES

Laboratory practical exercises

Description:
- Training of a convnet for character recognition. (1 hour)
- Visualization and ablation of convnet layers. (1 hour)
- Fine-tunning a convnet for transfer learning. (1 hour)
- Local image analysis. (1 hour)

Full-or-part-time: 4h
Laboratory classes: 4h

Extended answer test (Final examination)

Full-or-part-time: 1h
Theory classes: 1h

Final project presentations

Description:
Oral presentation of a solved Project (30 minutes)

Full-or-part-time: 3h
Theory classes: 3h

GRADING SYSTEM

Final examination: 30%
Final project: 30%
Laboratory assessments: 30%
Communication skills: 10%

BIBLIOGRAPHY

Basic:
- Goodfellow, Ian; Bengio, Y.; Courville, A. Deep Learning [on line]. 2016 [Consultation: 22/02/2016]. Available on: http://www.deeplearningbook.org/.

RESOURCES

Audiovisual material:
- Slides of the course and the bibliography referred within.

Hyperlink:
- Fei-Fei Li, Andrej Karpathy, "CS231n: Convolutional Neural Networks for Visual Recognition". Stanford University 2015. http://cs231n.stanford.edu/