230376 - DLV - Deep Learning for Vision

Coordinating unit: 230 - ETSETB - Barcelona School of Telecommunications Engineering
Teaching unit: 739 - TSC - Department of Signal Theory and Communications
Academic year: 2019
Degree: MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Teaching unit Optional)
MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Teaching unit Optional)
ECTS credits: 3  
Teaching languages: English

Teaching staff
Coordinator: Xavier Giró i Nieto
Others: Giró Nieto, Xavier Sayrol Clols, Elisa Ruiz Hidalgo, Javier

Requirements
Deep Learning for Artificial Intelligence

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
- Group work
- Group work (distance)

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.
230376 - DLV - Deep Learning for Vision

---

**Study load**

<table>
<thead>
<tr>
<th>Total learning time: 75h</th>
<th>Hours large group: 12h</th>
<th>16.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hours small group: 12h</td>
<td>16.00%</td>
</tr>
<tr>
<td></td>
<td>Self study: 51h</td>
<td>68.00%</td>
</tr>
</tbody>
</table>

---

**Content**

1. **State of the Art in Computer Vision**

   **Learning time:** 12h
   - Theory classes: 12h

   **Description:**
   - Still images: Object detection, image and instance segmentation, saliency prediction, visual search.
   - 3D & volumes: 3D Analysis, 3D reconstruction, medical applications.

2. **Industry Talks**

   **Learning time:** 2h
   - Theory classes: 2h

   **Description:**
   Talks by industrial professionals who are applying deep learning to address their challenges.

---

**Planning of activities**

**Exam**

- **Hours:** 10h
  - Guided activities: 1h
  - Self study: 9h

**Description:**
Written exam to evaluate the learning over the lecture contents.

**Project**

- **Hours:** 51h
  - Theory classes: 2h
  - Guided activities: 12h
  - Self study: 37h

**Description:**
Development and training of a deep neural network that will solve a computer vision task.

**Descriptions of the assignments due and their relation to the assessment:**
- Oral presentation in class.
- Release of the source codes and trained models.
Qualification system

Exam: 50%
Project: 50%
Attendance: -10 % penalty for each not completed day

Bibliography

Basic:

Others resources:

Hyperlink
Fei-Fei Li, Andrej Karpathy, "CS231n: Convolutional neural networks for visual recognition". Stanford University, 2015
http://cs231n.stanford.edu/

Nom reCourse website.