

Guia docent

230360 - DLCV - Aprenentatge Profund per a Visió per Computador

Última modificació: 06/05/2019

Unitat responsable: Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona

Unitat que imparteix: 739 - TSC - Departament de Teoria del Senyal i Comunicacions.

Titulació: Curs: 2019

Crèdits ECTS: 2.5

Idiomes: Anglès

PROFESSORAT

Professorat responsable: Xavier Giró i Nieto

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

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE1. Capacitat per aplicar mètodes de la teoria de la informació, la modulació adaptativa i codificació de canal, així com tècniques avançades de processat digital del senyal als sistemes de comunicacions i audiovisuals.

Transversals:

CT3. TREBALL EN EQUIP: Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o duent a terme tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, tot assumint compromisos considerant els recursos disponibles.

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

METODOLOGIES DOCENTS

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

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

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.



HORES TOTS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	16,0	25.60
Hores aprenentatge autònom	42,5	68.00
Hores grup petit	4,0	6.40

Dedicació total: 62.5 h

CONTINGUTS

1. Convolutional Neural Networks

Descripció:

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

Dedicació: 18h 30m

Grup gran/Teoria: 4h

Grup petit/Laboratori: 3h

Activitats dirigides: 4h

Aprenentatge autònom: 7h 30m

2. Applications

Descripció:

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

Dedicació: 44h

Grup gran/Teoria: 4h

Grup mitjà/Pràctiques: 7h

Grup petit/Laboratori: 1h

Activitats dirigides: 15h

Aprenentatge autònom: 17h



ACTIVITATS

Laboratory practical exercises

Descripció:

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

Dedicació: 4h

Grup petit/Laboratori: 4h

Extended answer test (Final examination)

Dedicació: 1h

Grup gran/Teoria: 1h

Final project presentations

Descripció:

Oral presentation of a solved Project (30 minutes)

Dedicació: 3h

Grup gran/Teoria: 3h

SISTEMA DE QUALIFICACIÓ

Final examination: 30%

Final project: 30%

Laboratory assessments: 30%

Communication skills: 10%

BIBLIOGRAFIA

Bàsica:

- Goodfellow, Ian; Bengio, Y.; Courville, A. Deep Learning [en línia]. 2016 [Consulta: 22/02/2016]. Disponible a: <http://www.deeplearningbook.org/>.

RECURSOS

Material audiovisual:

- Slides of the course and the bibliography referred within.

Enllaç web:

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