Guía docente
230360 - DLCV - Aprendizaje Profundo para Visión por Computadora

Última modificación: 06/05/2019

Unidad responsable: Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona
Unidad que imparte: 739 - TSC - Departamento de Teoría de la Señal y Comunicaciones.

Titulación: **Curso:** 2019 **Créditos ECTS:** 2.5
Idiomas: Inglés

PROFESORADO

Profesorado responsable: Xavier Giró i Nieto

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

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:
CE1. Capacidad para aplicar métodos de la teoría de la información, la modulación adaptativa y codificación de canal, así como técnicas avanzadas de procesado digital de señal a los sistemas de comunicaciones y audiovisuales.

Transversales:
CT3. TRABAJO EN EQUIPO: Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad, y valorar de forma crítica los resultados de dicha gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

METODOLOGÍAS DOCENTES

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

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

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.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas aprendizaje autónomo	42,5	68.00
Horas grupo pequeño	4,0	6.40
Horas grupo grande	16,0	25.60

Dedicación total: 62.5 h

CONTENIDOS

1. Convolutional Neural Networks

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

Dedicación: 18h 30m
Grupo grande/Teoría: 4h
Grupo pequeño/Laboratorio: 3h
Actividades dirigidas: 4h
Aprendizaje autónomo: 7h 30m

2. Applications

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

Dedicación: 44h
Grupo grande/Teoría: 4h
Grupo mediano/Prácticas: 7h
Grupo pequeño/Laboratorio: 1h
Actividades dirigidas: 15h
Aprendizaje autónomo: 17h

ACTIVIDADES

Laboratory practical exercises

Descripción:
- 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)

Dedicación: 4h
Grupo pequeño/Laboratorio: 4h

Extended answer test (Final examination)

Dedicación: 1h
Grupo grande/Teoría: 1h

Final project presentations

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

Dedicación: 3h
Grupo grande/Teoría: 3h

SISTEMA DE CALIFICACIÓN

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

BIBLIOGRAFÍA

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

RECURSOS

Material audiovisual:
- Slides of the course and the bibliography referred within.

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