

Guía docente

230726 - CVDL - Visión por Computador con Aprendizaje Profundo

Última modificación: 24/03/2026

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: MÁSTER UNIVERSITARIO EN INGENIERÍA DE TELECOMUNICACIÓN (Plan 2013). (Asignatura optativa).
MÁSTER UNIVERSITARIO EN TECNOLOGÍAS AVANZADAS DE TELECOMUNICACIÓN (Plan 2019). (Asignatura optativa).

Curso: 2026

Créditos ECTS: 5.0

Idiomas: Inglés

PROFESORADO

Profesorado responsable:

Otros:

CAPACIDADES PREVIAS

En caso de que no haya cursado con anterioridad DLAI no podrá matricular CVDL. Si cree que tiene los conocimientos necesarios para cursar CVDL deberá contactar con el coordinador/a coordinadora de CVDL para que valore tu background y pueda autorizar una excepción al requisito. En el caso de que se autorice la excepción al requisito deberá hacer una instancia que podrá solicitar:

- Antes de la matrícula, durante el período de solicitudes de excepción a la matrícula, para que CVDL le aparezca en su potencial de matrícula y la pueda matricular siempre que haya plazas disponibles.

o bien

- Después de la matrícula, durante el período de modificación de matrícula, para incluir CVDL en la matrícula, siempre y cuando queden plazas disponibles

Tanto las instancias de excepción en la matrícula como las de modificación de matrícula serán aceptadas por el jefe de estudios del MATT siempre que el coordinador/a le haya comunicado la autorización de excepción al requisito de DLAI.

También son necesarios los siguientes conocimientos previos para seguir el curso:

- Procesado de imagen: píxeles, espacios de color, histogramas, representación del dominio de frecuencia
- Procesado de señal digital: filtros lineales, convolución
- Álgebra vectorial y matricial

Las nociones de Python son útiles, pero se obtienen fácilmente durante el curso.

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

1. Capacidad para aplicar métodos de la teoría de la información, la modulación adaptativa y codificación de canal, así como técnicas avanzadas de procesamiento digital de señal a los sistemas de comunicaciones y audiovisuales.
2. Capacidad para la integración de tecnologías y sistemas propios de la Ingeniería de Telecomunicación, con carácter generalista, y en contextos más amplios y multidisciplinares como por ejemplo en bioingeniería, conversión fotovoltaica, nanotecnología, telemedicina.

Transversales:

3. 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.

4. 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.

5. 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

- Clases presenciales
- Trabajo practico
- Trabajo individual (no presencial)
- Ejercicios
- Exámenes parciales

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

Objetivos de aprendizaje de la asignatura:

El objetivo de este curso es proporcionar una visión general de los conceptos y aplicaciones de la visión por computador, utilizando métodos clásicos y de aprendizaje profundo. Introduciremos técnicas de bajo nivel, como extracción y comparación de características, detección de contornos, modelos de cámara y de proyección y flujo óptico; temas de nivel medio como segmentación de video y seguimiento de características; métodos de alto nivel como el seguimiento de objetos. Luego, se mostrarán ejemplos de aplicación, como el reconocimiento de caras y objetos.

Resultados de aprendizaje de la asignatura:

- Capacidad para comprender y utilizar técnicas para el análisis de imágenes y videos: extracción de características, segmentación de video, imagen estéreo, detección de objetos.
- Capacidad para utilizar algoritmos de visión por computador para implementar aplicaciones de alto nivel.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTE

Tipo	Horas	Porcentaje
Horas grupo grande	39,0	31.20
Horas aprendizaje autónomo	86,0	68.80

Dedicación total: 125 h

CONTENIDOS

(CAST) 1. Introduction

Descripción:

- Motivación, tipos de problemas en CV
- Formación de imágenes, percepción, sensores 3D.

Dedicación: 7h

Grupo grande/Teoría: 3h

Aprendizaje autónomo: 4h

(CAST) 2. Image Structure

Descripción:

- Color, textura, filtrado y contornos.
- Detección y representación de puntos característicos y de 'blobs'
- Modelado: RANSAC, transformada de Hough
- Mapas de prominencia

Dedicación: 29h

Grupo grande/Teoría: 9h

Actividades dirigidas: 4h

Aprendizaje autónomo: 16h

(CAST) 3. Stereo and 3D applications

Descripción:

- Geometría de cámara única, calibración de cámara
- Geometría epipolar, homografía
- Estimación de pose de cámara y registro de sensor usando aprendizaje profundo

Dedicación: 30h

Grupo grande/Teoría: 9h

Actividades dirigidas: 4h

Aprendizaje autónomo: 17h

(CAST) 4. Video tracking

Descripción:

- Flujo óptico: Lucas-Kanade, Shi-Tomasi, métodos de aprendizaje profundo
- Seguimiento bayesiano: Kalman, filtros de partículas
- Métodos de seguimiento de aprendizaje profundo

Dedicación: 29h

Grupo grande/Teoría: 9h

Actividades dirigidas: 4h

Aprendizaje autónomo: 16h

(CAST) 5. Detection and recognition

Descripción:

- Introducción al reconocimiento visual. Repaso de 'Machine learning' Aprendizaje profundo y redes neuronales convolucionales
- Clasificación de imagen: Modelo de 'bag of words'. Clasificación de imágenes usando CNN
- Detección de objetos: Ventanas deslizantes y características locales. Detección de objetos utilizando CNN
- Segmentación de objetos: Segmentación semántica. Segmentación de instancias

Dedicación: 30h

Grupo grande/Teoría: 9h

Actividades dirigidas: 4h

Aprendizaje autónomo: 17h

ACTIVIDADES

(CAST) EJERCICIOS

Descripción:

- Detección de contornos modelado de formas: métodos Canny, Hough, Ransac y DL
- Encontrar correspondencias entre imágenes: Harris, SIFT
- Estimación de la matriz fundamental
- Aplicación de la homografía: creación de panorama.
- Detección y reconocimiento de objetos.

Dedicación: 6h

Aprendizaje autónomo: 6h

(CAST) EXAMEN PARCIAL

Descripción:

Examen parcial

Dedicación: 2h

Grupo grande/Teoría: 2h

(CAST) EXAMEN PARCIAL

Descripción:

Examen parcial

Dedicación: 2h

Grupo grande/Teoría: 2h

SISTEMA DE CALIFICACIÓN

Cualquier acto de fraude académico, plagio o uso o mera tenencia al alcance de medios no autorizados en cualquier actividad de evaluación comportará la calificación de cero (0) en la prueba o entrega afectada. Además, de acuerdo con la normativa de la Universidad, la posible derivación de los hechos para la apertura de un expediente disciplinario implicará que la asignatura quede en el estado provisional de "pendiente de evaluación" hasta la resolución del expediente. La gestión de estas incidencias se lleva a cabo de acuerdo con el [Marco de actuación para la integridad académica en la evaluación de la UPC](#).

Primer parcial: 40%
Segundo parcial: 40%
Laboratorio/Ejercicios: 20%

BIBLIOGRAFÍA

Básica:

- 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 [en línea]. 2nd ed. Boston, Mass.: Pearson Education, 2012 [Consulta: 09/09/2020]. Disponible a: <https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=5173504>. ISBN 9780273764144.

Complementaria:

- 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 [en línea]. Boston: Kluwer Academic, 2004 [Consulta: 29/07/2013]. Disponible a: <http://link.springer.com/book/10.1007/b106003/page/1>. ISBN 978-1402081149.

RECURSOS

Otros recursos:

Google Colab