

Guia docent

230726 - CVDL - Visió per Ordinador amb Aprenentatge Profund

Última modificació: 06/06/2025

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

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

Titulació: MÀSTER UNIVERSITARI EN ENGINYERIA DE TELECOMUNICACIÓ (Pla 2013). (Assignatura optativa).
MÀSTER UNIVERSITARI EN TECNOLOGIES AVANÇADES DE TELECOMUNICACIÓ (Pla 2019). (Assignatura optativa).

Curs: 2025

Crèdits ECTS: 5.0

Idiomes: Anglès

PROFESSORAT

Professorat responsable: JOSEP RAMON MORROS RUBIO

Altres:

CAPACITATS PRÈVIES

En el cas de que no hagueu cursat amb anterioritat DLAI no podreu matricular CVDL. Si creieu que teniu els coneixements necessaris per cursar CVDL haureu de contactar amb el coordinador/la coordinadora de CVDL per tal de que valori el teu background i pugui autoritzar una excepció al requisit. En el cas de què s'autoritzi l'excepció al requisit haureu de fer una instància que podreu sol·licitar:

- Abans de la matrícula, durant el període de sol·licituds d'excepció a la matrícula, per tal de que CVDL us aparegui al vostre potencial de matrícula i la pugueu matricular sempre que hi hagi places disponibles.

o bé

- Després de la matrícula, durant el període de modificació de matrícula, per incloure CVDL a la matrícula, sempre i quan restin places disponibles.

Tant les instàncies d'excepció a la matrícula com les de modificació de matrícula seran acceptades pel cap d'estudis del MATT sempre que el coordinador/la coordinadora li hagi comunicat l'autorització d'excepció al requisit de DLAI.

També heu de tenir els coneixements previs següents per seguir el curs:

- Processament d'imatge: píxels, espais de colors, histogrames, representació en el domini freqüencial
- Processament digital del senyal: filtres lineals, convolució
- Àlgebra vectorial i matricial

Les nocions de Python són útils, però es poden obtenir fàcilment durant el curs.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

1. 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.
2. Capacitat per a la integració de tecnologies i sistemes propis de la Enginyeria de Telecomunicació, amb caràcter generalista, i en contextos més amplis i multidisciplinaris com per exemple en bio-enginyeria, conversió fotovoltàica, nanotecnologia o telemedicina.

Transversals:

3. 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.

4. 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.

5. 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

- Classes presencials
- Treball pràctic
- Treball individual (no presencial)
- Exercicis
- Exàmens parcials

OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

Objectius d'aprenentatge de l'assignatura:

L'objectiu d'aquest curs és proporcionar una visió general dels conceptes i aplicacions de la visió per computador, utilitzant mètodes clàssics i d'aprenentatge profund. Introduïrem tècniques de baix nivell com ara l'extracció i la comparació de vectors de característiques, la detecció de contorns, models de càmera i de projecció i flux òptic; temes de nivell mitjà com la segmentació de vídeo i el seguiment de característiques; mètodes d'alt nivell com la detecció i seguiment d'objectes. A continuació, es mostraran exemples d'aplicació, com ara el reconeixement de cares i d'objectes.

Resultats d'aprenentatge de l'assignatura:

- Capacitat per comprendre i utilitzar tècniques d'anàlisi d'imatges i vídeos: extracció de característiques, segmentació de vídeo, Imatge estèreo, detecció d'objectes.
- Capacitat d'utilitzar algoritmes de visió informàtica per implementar aplicacions d'alt nivell.

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	39,0	31.20
Hores aprenentatge autònom	86,0	68.80

Dedicació total: 125 h

CONTINGUTS

(CAT) 1. Introduction

Descripció:

- Motivació, tipus de problemes en CV
- Formació d'imatges, percepció, sensors 3D

Dedicació: 7h

Grup gran/Teoria: 3h

Aprenentatge autònom: 4h

(CAT) 2. Image Structure

Descripció:

- Color, textura, filtrat i contorns
- Detecció i representació de punts característics i de blobs
- Modelat: RANSAC, transformada de Hough
- Mapes de prominència

Dedicació: 29h

Grup gran/Teoria: 9h

Activitats dirigides: 4h

Aprenentatge autònom: 16h

(CAT) 3. Stereo and 3D applications

Descripció:

- Geometria de càmera única, calibració de la càmera
- Geometria epipolar, homografia
- Estimació de posi de la càmera i registre de sensors mitjançant l'aprenentatge profund

Dedicació: 30h

Grup gran/Teoria: 9h

Activitats dirigides: 4h

Aprenentatge autònom: 17h

(CAT) 4. Video tracking

Descripció:

- Flux òptic: Lucas-Kanade, Shi-Tomasi, mètodes d'aprenentatge profund
- Seguiment bayesià: Kalman, Filtres de partícules
- Mètodes de seguiment amb aprenentatge profund

Dedicació: 29h

Grup gran/Teoria: 9h

Activitats dirigides: 4h

Aprenentatge autònom: 16h

(CAT) 5. Detection and recognition

Descripció:

- Introducció al reconeixement visual. Repàs de 'machine learning'. Aprenentatge profund i xarxes neuronals revolucionàries
- Classificació d'imatge: Model de 'bag of words'. Classificació de la imatge mitjançant CNNs
- Detecció d'objectes: Finestres lliscants i funcions locals. Detecció d'objectes mitjançant CNNs
- Segmentació d'objectes: Segmentació semàntica. Segmentació d'instància

Dedicació: 30h

Grup gran/Teoria: 9h

Activitats dirigides: 4h

Aprenentatge autònom: 17h

ACTIVITATS

(CAT) EXERCICIS

Descripció:

- Detecció de contorns i modelat de formes: mètodes Canny, Hough, Ransac y DL
- Cercar correspondències entre imatges: Harris, SIFT
- Estimació de la matriu fonamental
- Aplicació de l'homografia: creació de panorama
- Detecció i reconeixement d'objectes

Dedicació: 6h

Aprenentatge autònom: 6h

(CAT) EXAMEN PARCIAL

Descripció:

Examen parcial

Dedicació: 2h

Grup gran/Teoria: 2h

(CAT) EXAMEN PARCIAL

Descripció:

Examen parcial

Dedicació: 2h

Grup gran/Teoria: 2h

SISTEMA DE QUALIFICACIÓ

Primer parcial: 40%

Segon parcial: 40%

Laboratori/Exercicis: 20%



BIBLIOGRAFIA

Bàsica:

- 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 línia]. 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.

Complementària:

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

RECURSOS

Altres recursos:

Google Colab