

Titulació	Tipus	Curs
Computer Vision	OB	1

Professor/a de contacte

Nom : Maria Vanrell Martorell

Correu electrònic : maria.vanrell@uab.cat

Equip docent

Javier Ruiz Hidalgo

Ramon Morros Rubio

Gloria Haro Ortega

Montse Pardàs Feliu

Federico Sukno

Albert Clapés Sintes

Equip docent (extern a la UAB)

David Pujol

German Barquero

Artur Xarles

Idiomes dels grups

Podeu consultar aquesta informació al [final](#) del document.

Prerequisits

Un grau en Enginyeria, Matemàtiques, Física o similar.

Assignatura C3: Machine Learning for Computer Vision

Competències de programació en Python

Objectius

Coordinador del curs: Dr. Javier Ruiz / Dr. Albert Clapés

L'objectiu d'aquest curs és presentar els principals conceptes i tecnologies que són necessaris per a l'anàlisi

de vídeo. En primer lloc, es presenten les aplicacions d'anàlisi de seqüències d'imatges i els diferents tipus de dades on s'aplicaran aquestes tècniques. A més es proporciona una descripció general de les tècniques de processament de senyals i les arquitectures generals d'aprenentatge profund en què es basa l'anàlisi de vídeo. Es donaran exemples per a seqüències de vídeo monocàmera, multicàmera i amb càmeres de profunditat. S'estudiaran tant les bases teòriques com els algorismes més utilitzats. Per a cada matèria, es presentaran tècniques clàssiques juntament amb les tècniques d'aprenentatge profund de l'estat de l'art que condueixen a diferents enfocaments. Els temes principals seran la segmentació de vídeo, la sostracció i modelatge del fons, l'estimació de moviment, els algorismes de seguiment i l'anàlisi basat en models. També s'estudiaran tècniques de nivell superior, com el reconeixement de gestos o accions, la generació profunda de vídeos i l'aprenentatge profund multimodal.

Els estudiants treballaran en un projecte d'anàlisi de seqüències de vídeo. En una primera part, serà un sistema de monitorització de trànsit aplicat a ADAS (sistemes avançats d'assistència al conductor) on s'aplicaran algorismes i models de detecció d'objectes, segmentació, seguiment i estimació de flux-òptic per a vídeo. En una segona part, l'objectiu serà la detecció i reconeixement d'accions en vídeos.

Resultats d'aprenentatge

- CA03 (Definir tots els components que cooperen en un sistema complet d'anàlisi de seqüències d'imatges.) Definir tots els components que cooperen en un sistema complet d'anàlisi de seqüències d'imatges.
- CA06 (Aconseguir els objectius d'un projecte de visió fet en equip.) Aconseguir els objectius d'un projecte de visió fet en equip.
- KA06 (Identificar els problemes bàsics que s'han de solucionar en un problema de seqüències d'imatges d'escenes .) Identificar els problemes bàsics que s'han de solucionar en un problema de seqüències d'imatges d'escenes .
- KA14 (Proporcionar la millor modelització per solucionar problemes de segmentació de vídeos, d'estimació del moviment o del seguiment d'objectes.) Proporcionar la millor modelització per solucionar problemes de segmentació de vídeos, d'estimació del moviment o del seguiment d'objectes.
- SA05 (Resoldre un problema de reconeixement visual entrenant una arquitectura de xarxa neuronal profunda i avaluar-ne els resultats.) Resoldre un problema de reconeixement visual entrenant una arquitectura de xarxa neuronal profunda i avaluar-ne els resultats.
- SA11 (Definir els millors conjunts de dades per entrenar arquitectures de reconeixement visual.) Definir els millors conjunts de dades per entrenar arquitectures de reconeixement visual.
- SA15 (Preparar un informe que descrigui, justifiqui i il·lustri el desenvolupament d'un projecte de visió.) Preparar un informe que descrigui, justifiqui i il·lustri el desenvolupament d'un projecte de visió.
- SA17 (Preparar presentacions orals que permetin debatre els resultats del desenvolupament d'un projecte de visió.) Preparar presentacions orals que permetin debatre els resultats del desenvolupament d'un projecte de visió.

Continguts

1. Segmentació de vídeo
2. Estimació del moviment
3. Seguiment d'objectes
4. Xarxes Neuronals Recurrents
5. Atenció i Transformers per vídeo
6. Detecció i Reconeixement d'accions
7. Auto-supervisió i aprenentatge multi-modal per vídeo.
8. Adaptació-al-domini per vídeo
9. Detecció d'anomalies
10. Generació de vídeo

Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Sessions de seguiment de projectes	10	0,4	CA03, CA06, KA06, KA14, SA05, SA11, SA15, SA17
Sesions teòriques	35	1,4	CA03, CA06, KA06, KA14, SA05, SA11, SA15, SA17
Treball autònom	171	6,84	CA03, CA06, KA06, KA14, SA05, SA11, SA15, SA17

Sessions supervisades: *(Algunes d'aquestes sessions podrien ser en línia síncroniques)*

- Sessions de teoria, on els professors expliquen continguts generals sobre els diferents temes. Alguns d'ells s'empraran per resoldre els problemes.

Sessions dirigides:

- Sessions de projecte, on els problemes i els objectius dels projectes seran presentats i discutits, els estudiants interactuaran amb el coordinador del projecte sobre problemes i idees per resoldre el projecte (aprox. 1 hora/setmana).
- Sessió de presentació, on els estudiants fan una presentació oral sobre com han resolt el projecte i una demostració dels resultats.
- Sessió d'examen, on els estudiants són avaluats de forma individual sobre el coneixement obtingut i les habilitats per resoldre problemes.

Treball autònom:

- L'estudiant estudiarà de forma autònoma i treballarà els materials derivats de les sessions teòriques.
- L'estudiant treballarà en grups per a resoldre els problemes del projecte amb entregues de:
 - Codi
 - Informes
 - Presentació oral

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, perquè l'alumnat completi les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura.

Avaluació

Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Exàmen	0.4	2,5	0,1	CA03, KA06, KA14, SA05, SA11
Projecte	0.55	6	0,24	CA03, CA06, KA06, KA14, SA05, SA11, SA15, SA17
Assistència a sessions	0.05	0,5	0,02	CA03, CA06, KA06, KA14, SA05, SA11, SA15, SA17

La Nota final d'aquest mòdul es calcula amb la següent fórmula:

Nota final = 0.4 x Examen + 0.55 x Projecte + 0.05 x Assistència

on,

Exàmen: és la nota obtinguda en l'examen final del mòdul (ha de ser ≥ 3).

Assistència: és la nota derivada de l'assistència a les sessions (mínim 70%).

Projecte: és la nota que posa el coordinador del projecte d'acord amb els resultats del seguiment setmanal del

projecte i de les entregues (ha de ser ≥ 5). Tot d'acord amb criteris específics com ara:

- Participació en les sessions de discussió i en el treball en equip (avaluació entre-membres)
- Entrega de les parts obligatòries i opcionals del projecte.
- Desenvolupament del codi (estil, comentaris, etc.)
- Informe (justificació de les decisions preses en el desenvolupament del projecte)
- Presentació (Presentació i demostració del projecte desenvolupat)

Només els estudiants que han suspès (Nota final < 5.0) poden fer l'examen de recuperació.

Bibliografia

Articles:

1. M. Piccardi. "Background subtraction techniques: a review". Journal: IEEE Int. Conf. On Systems, Man and Cybernetics 2004 , v. 4, pp. 3099-3104, 2004.
2. A. Sobral, A. Vacavant, "A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos", Journal: Computer Vision and Image Understanding Vol. 122, pp. 4-21 · May 2014.
3. S. Baker, D. Scharstein, JP. Lewis, S. Roth, M. Black, R. Szeliski. "A database and evaluation methodology for optical flow". Journal: International Journal of Computer Vision, Vol. 92:1, pp. 1-31, 2011.
4. T. Cootes, G. Edwards, C. Taylor. "Active appearance models". Journal: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 6, pp. 681--685, 2001.
5. R. Poppe. "Vision-based Human motion analysis: an overview". Journal: Computer Vision and Image Understanding 108 (1-2): 4-18, 2007

Llibres:

1. "Sequential Monte Carlo methods in practice", A. Doucet, N. de Freitas and N. Gordon (Eds.), Springer, 2001.

Programari

Entorn de programació en Python amb especial atenció a les llibreries més usades de visió per computador i aprenentatge profund.

Grups i idiomes de l'assignatura

La informació proporcionada és provisional fins al 30 de novembre.
A partir d'aquesta data, podreu consultar l'idioma de cada grup a través d'aquest [enllaç](#). Per accedir a la informació, caldrà introduir el CODI de l'assignatura