

Aprendizaje Computacional para la Visión por Computador

Código: 43083
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
4314099 Visión por Computador / Computer Vision	OB	0	1

La metodología docente y la evaluación propuestas en la guía pueden experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Contacto

Nombre: Maria Vanrell Martorell

Correo electrónico: Maria.Vanrell@uab.cat

Equipo docente

Ramón Baldrich Caselles

Maria Vanrell Martorell

Fernando Luis Vilariño Freire

Dimosthenis Karatzas

Pau Rodriguez Lopez

Guillem Arias Bedmar

Lluís Gomez Bigorda

Uso de idiomas

Lengua vehicular mayoritaria: inglés (eng)

Prerequisitos

- Un grado en Ingeniería, Matemáticas, Física o similar.
- Competencias de programación en Python.

Objetivos y contextualización

Coordinador del módulo: Dr. Ramon Baldrich Caselles

El objetivo de este módulo es presentar las técnicas de aprendizaje automático para resolver problemas de visión por computador. El aprendizaje automático se ocupa del análisis automático de datos a gran escala. Hoy en día, conforma los conceptos básicos de muchos métodos de visión por computadora, especialmente aquellos relacionados con el reconocimiento o la clasificación de patrones visuales, donde los 'patrones' abarcan imágenes de objetos del mundo, escenas y secuencias de video de acciones humanas, por nombrar algunos.

Este módulo presenta los fundamentos y las técnicas más importantes para la clasificación de patrones visuales, centrándose principalmente en métodos supervisados. Además, se abordan temas relacionados como descriptores de imágenes y reducción de dimensionalidad. En la medida de lo posible, todas estas técnicas se prueban y evalúan en un proyecto práctico relacionado con la descripción de la escena a partir de imágenes, junto con las métricas y los procedimientos estándar para la evaluación del rendimiento, como las curvas de recuperación de precisión y la validación cruzada en k veces.

Los resultados del aprendizaje son:

(a) Distinga los principales tipos de técnicas de ML para la visión por computadora: supervisado versus no supervisado, generativo versus discriminativo, espacio de características originales versus kernelización de vectores de características.

(b) Conozca los puntos fuertes y débiles de los diferentes métodos, en parte aprendidos mientras resuelve un problema real de clasificación de patrones.

(c) Ser capaz de usar implementaciones de métodos existentes y construirlas desde cero.

El módulo profundiza en dos enfoques principales para introducir ML en el problema de clasificación de imágenes. Usando: a) descripción de imagen hecha a mano, b) descripción de imagen basada en datos. En el primer caso se usa Bag of Words, en el segundo el enfoque de aprendizaje profundo. El contenido de DL se desarrolla, proporcionando tanto bases teóricas de las diferentes partes de los modernos sistemas de Neural Networks como las mejores prácticas para aplicarlo en aplicaciones reales.

Competencias

- Asumir tareas de responsabilidad en la gestión de la información y el conocimiento.
- Comprender, analizar y sintetizar los conocimientos avanzados que existen en el área, así como proponer ideas innovadoras.
- Conceptualizar alternativas de soluciones complejas a problemas de visión y crear prototipos que demuestren la validez del sistema propuesto.
- Identificar los conceptos y aplicar las técnicas fundamentales más adecuadas para la solución de los problemas básicos de la visión por computador.
- Planificar, desarrollar, evaluar y gestionar soluciones a proyectos en los diferentes ámbitos de la visión por computador.
- Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
- Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
- Seleccionar las herramientas software y los conjuntos de entrenamiento más adecuados para desarrollar las soluciones a los problemas de visión por computador.
- Trabajar en equipos multidisciplinares.

Resultados de aprendizaje

1. Asumir tareas de responsabilidad en la gestión de la información y el conocimiento.
2. Comprender, analizar y sintetizar los conocimientos avanzados que existen en el área, así como proponer ideas innovadoras.
3. Identificar las mejores representaciones que se puedan definir para la resolución de problemas de aprendizaje computacional.
4. Identificar los algoritmos básicos de aprendizaje computacional y su aplicación.
5. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
6. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
7. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
8. Seleccionar técnicas de aprendizaje computacional y entrenarlas para que solucionen un proyecto concreto.
9. Trabajar en equipos multidisciplinares.

- Usar técnicas de aprendizaje computacional para planificar, desarrollar, evaluar y gestionar una solución a un problema concreto.

Contenido

- Introducción al aprendizaje computacional
- Entorno experimental
- Codificaciones: SVM y Random Forest
- Introducción a las Redes Neuronales
- Introducción al Aprendizaje Profundo (Deep Learning, DL)
- Redes Neuronales Convolucionales (CNN)
- Entrenamiento: preprocesado de datos, inicialización, optimización del gradiente
- Clasificación de imágenes
- Entender y visualización de CNNs
- Métodos eficientes para el Aprendizaje Profundo

Metodología

Sesiones supervisadas: *(Algunas d'aquestes sesiones podrien ser en línea síncronas)*

- Sesiones magistrales, donde los profesores explicarán contenidos generales de los diferentes temas. La mayoría serán necesarios para la resolución de problemas.

Sesiones dirigidas:

- Sesiones de proyecto, donde los objetivos y problemas de los proyectos se presentarán y discutirán. Los estudiantes deberán interactuar con el coordinador de proyecto sobre los problemas surgidos y las ideas aportadas para resolverlos. (Approx. 1 hora/semana)
- Sesiones de presentación, donde los estudiantes harán una presentación oral sobre como han solucionado el problema y sobre los resultados obtenidos.
- Sesión de exámen, donde los estudiantes son evaluados individualmente, demostrando la adquisición de los conocimientos desarrollados y la capacidad de resolución de problemas asociados.

Trabajo autónomo:

- Los estudiantes deberán estudiar y trabajar autonomamente con los materiales derivados de las clases magistrales y de las sesiones de proyecto.
- Los estudiantes trabajaran en grupo para resolver los problemas planteados en los proyectos con los siguientes entregables:
 - Código
 - Informe
 - Presentación oral

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Sesiones teóricas	20	0,8	4, 3, 5
Tipo: Supervisadas			

Sesiones de seguimiento de proyecto	8	0,32	1, 2, 4, 3, 7, 6, 8, 5, 9, 10
Tipo: Autónomas			
Trabajo autónomo	113	4,52	1, 2, 4, 3, 7, 6, 8, 5, 9, 10

Evaluación

La Nota final de este módulo se calcula con la siguiente fórmula:

$$\text{Nota final} = 0.4 \times \text{Examen} + 0.55 \times \text{Proyecto} + 0.05 \times \text{Asistencia}$$

donde,

Examen: es la nota obtenida en el examen final del módulo (tiene que ser ≥ 3).

Asistencia: es la nota derivada de la asistencia a las sesiones (mínimo 70%)

Proyecto: es la nota que pone el coordinador del proyecto de acuerdo con los resultados del seguimiento semanal del proyecto y de las entregas (tiene que ser ≥ 5). Todo de acuerdo con criterios específicos, tales como:

- Participación en las sesiones de discusión y en el trabajo en equipo (evaluación entre-miembros)
- Entrega de las partes obligatorias y opcionales del proyecto.
- Desarrollo del código (estilo, comentarios, etc.)
- Informe (justificación de las decisiones tomadas en el desarrollo del proyecto)
- Presentación (Presentación y demostración del proyecto desarrollado)

Sólo los estudiantes que han suspendido (nota final < 5.0) podran hacer el examen de recuperación.

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Asistencia a sesiones	0.05	0,5	0,02	1, 4, 3, 7, 5
Examen	0.4	2,5	0,1	1, 2, 3, 7, 8, 5, 10
Proyecto	0.55	6	0,24	1, 2, 4, 3, 7, 6, 8, 5, 9, 10

Bibliografía

Artículos en revistas:

1. Barber, D. "*Bayesian Reasoning and Machine Learning*". Cambridge University Press, 2012.
2. Yoshua Bengio. "*Learning Deep Architectures for AI*". Foundations and Trends in Machine Learning, Vol. 2, No. 1, 2009.
3. Christopher J. C. Burges. "*Dimension Reduction: A Guided Tour*". Foundations and Trends in Machine Learning, Vol. 2, No. 4, 2009.
4. Christoph H. Lampert. "*Kernel Methods in Computer Vision*". Foundations and Trends in Computer Graphics and Vision, Vol. 4, No. 3, 2008.
5. Tinne Tuytelaars and Krystian Mikolajczyk. "*Local Invariant Feature Detectors: A Survey*". Foundations and Trends in Computer Graphics and Vision, Vol. 3, No. 3, 2007.

Libros:

1. Ian Goodfellow, Yoshua Bengio and Aaron Courville. "*Deep Learning*". 2016. Cambridge, MA, USA: The MIT Press. ISBN: 978-0262035613
2. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, "*Foundations of Machine Learning*" MIT Press, 2012. <http://www.cs.nyu.edu/~mohri/mlbook/>
3. Z.H. Zhou. *Ensemble Methods: Foundations and Algorithms*. Chapman & Hall/CRC, 2012.

Informes técnicos:

1. Criminisi, A. and Shotton, J. and Konukoglu, E. "*Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning*". Technical report MSR-TR-2011-114. Microsoft Research, 2011.
http://research.microsoft.com/pubs/155552/decisionForests_MSR_TR_2011_114.pdf

Software

Entorno de programación en Python con especial atención a las librerías de visión por computador y Keras