UAB

Universitat Autònoma de Barcelona

2019/2020

Visual Recognition

Code: 43088
ECTS Credits: 6

Degree	Type	Year	Semester
4314099 Computer Vision	OB	0	2

Contact

Name: Joan Serrat Gual

Email: Joan.Serrat@uab.cat

Teachers

Jose Luis Gomez Zurita

Carles Ventura Royo

Luis Herranz Arribas

Lluis Gomez Bigorda

External teachers

Adriana Romero

Michal Drozdzal

Use of Languages

Principal working language: english (eng)

Prerequisites

- Degree in Engineering, Maths, Physics or similar.
- Module 2 "Machine learning for computer vision"

Objectives and Contextualisation

In Computer Vision, visual recognition corresponds to the task of explaining the content of an image in terms of "What is it?" "where is this?". The answer to these questions is usally a class label corresponding to the object or object types in the image, a tight bounding box containing the object in question, or, at a finer level, the region (pixels) that is its outline. These tasks are called, respectively, image classification, object detection and semantic segmentation. A question is "give me objects like this one", that requires learning a similary metric between images, even in the case come from different modalities, like sketches and photographs, through the so called encoder-decoder architectures. VR module covers neural network architectures addressing these four types of tasks. And, as a practical complement, methods to implement them.

Specifically, in this module we give to the student an overview of the latest methods based on deep learning techniques to solve visual recognition problems. The final aim is the understanding of complex scenes to build feasible systems for automatic image understanding able to answer the complex question of what objects and where are these objects in a complex scene.

Having addressed the task of classification in module M2, the students will learn a large family of successful architectures of deep convolutional networks that have been proved to solve the visual tasks of detection and

segmentation and recognition. In addition to these two visual tasks, this module addresses also advanced topics in deep learning such as architectures for image generation (GANs and VAEs) plus encoder-decoder architectures for multimodal applications.

Competences

- Accept responsibilities for information and knowledge management.
- Choose the most suitable software tools and training sets for developing solutions to problems in computer vision.
- Conceptualise alternatives to complex solutions for vision problems and create prototypes to show the validity of the system proposed.
- Continue the learning process, to a large extent autonomously.
- Identify concepts and apply the most appropriate fundamental techniques for solving basic problems in computer vision.
- Plan, develop, evaluate and manage solutions for projects in the different areas of computer vision.
- Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
- Understand, analyse and synthesise advanced knowledge in the area, and put forward innovative ideas.
- Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.
- Work in multidisciplinary teams.

Learning Outcomes

1. Accept responsibilities for information and knowledge management.
2. Choose the learnt techniques and train them to resolve a particular visual recognition project.
3. Continue the learning process, to a large extent autonomously.
4. Identify the basic problems to be solved in object and scene recognition, along with the specific algorithms.
5. Identify the best representations that can be defined for solving problems of both object and scene visual recognition.
6. Plan, develop, evaluate and manage a solution to a particular visual recognition problem.
7. Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
8. Understand, analyse and synthesise advanced knowledge in the area, and put forward innovative ideas.
9. Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.
10. Work in multidisciplinary teams.

Content

1. Object detection
2. Semantic and instance segmentation
3. Reinforcement learning
4. Architectures for image generation
     generative adversarial networks (GAN)
     variational autoencoders (VAE)
6. Metric learning
7. Encoder-decoder architectures for multimodal applications
8. Conditional GANs

Methodology

The learning methdology is based on the lectures and exercises but mainly on the project, which is developed during the whole module. It consists in solving some tasks for scene understanding applied to autonomous

driving. The goal is to learn the basic concepts and techniques to build deep neural networks to detect, segment and recognize specific objects, focusing on images recorded by an on-board vehicle camera for autonomous driving.

The learning objectives are using different a deep learning (DL) programming frameworks (at present, PyTorch) and basic DL methods such as feed forward networks (MLP) and Convolutional Neural Networks (CNN). It includes the understanding of standard networks for detection (RCNN, Fast RCNN, Faster RCNN, YOLO) and segmentation (FCN, SegNet, UNET). The students will learn through a project based methodology using modern collaborative tools at all stages of the project development.

Students will acquire the skills for the tasks of designing, training, tuning and evaluating neural networks to solve the problem of automatic image understanding.

All this is done through three formats:

1. Supervised sessions: lectures where the instructors will explain general contents about the different topics. They will be used to solve the project and/or proposed exercises.
2. Directed sessions:
    1. Project Sessions, where the problems and goals of the projects will be presented and discussed, students will interact with the project coordinator about problems and ideas on solving the project. Additionally, the students give oral presentations about how have they solved the project and report results (approx once per week)
    2. Exam Session, where the students are evaluated individually. Knowledge achievements and problem-solving skills
4. Autonomous work :
    study and work with the materials derived from the lectures, plus solving some small practical exercises to better understand theoretical lectores that arenot directly involved in the project solution
    work in groups to solve the problems of the projects with deliverables: code, reports,oral presentations, exercises

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
lectures	20	0.8	4, 5, 2
Type: Supervised			
project	8	0.32	1, 8, 4, 5, 7, 3, 2, 9, 10
Type: Autonomous			
homework, exercises	112	4.48	1, 8, 4, 5, 6, 7, 3, 2, 9, 10

Assessment

The final marks for this module will be computed with the following formula:

Final Mark= 0.4 x Exam+ 0.55 x Project + 0.05 x Attendance

where

Exam:is the mark obtained in the Module Exam. This mark can be increased by getting extra points given by delivered exercises in specific lectures.

Attendance: is the mark derived from the control of attendance at lectures (minimum 70%)

Project:is the mark provided by the project coordinator based on the weekly follow-up of the project and deliverables. All accordingly with specific criteria such as:

- Participation in discussion sessions and in team work (inter-member evaluations)

- Delivery of mandatory and optional exercises.

- Code development (style, comments, etc.)

- Report (justification of the decisions in your project development)

- Presentation (Talk and demonstrations on your project)

The Exam mark can be increased by extra points gained from delivered exercises proposed in certain lectures, but only if Exam Mark is greater than 3.

Only those students that fail (FinalMark< 5.0) can do a retake exam.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
exam	0.4	3	0.12	1, 8, 4, 5, 7, 9
project	0.55	7	0.28	1, 8, 4, 5, 6, 7, 3, 2, 9, 10

Bibliography

Generic references :

1. Deep Learning. Ian Goodfellow, Yoshua Bengio, Aaron Courville. MIT Press, 2016.

2. Neural networks and deep learning. Michael Nielsen. http://neuralnetworksanddeeplearning.com

Most of the content is related to the state of the art in the different topics, so there are no books published but survey and research papers specific of each one that will be selected by the lecturers.