Course guide
270732 - OR - Object Recognition

Last modified: 02/02/2024

Unit in charge: Barcelona School of Informatics
Teaching unit: 1004 - UB - (ENG)Universitat de Barcelona.

Degree: MASTER'S DEGREE IN ARTIFICIAL INTELLIGENCE (Syllabus 2017). (Optional subject).

Academic year: 2023 **ECTS Credits:** 4.0 **Languages:** English

LECTURER

Coordinating lecturer: SERGIO ESCALERA GUERRERO

Others:

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CEA13. Capability to understand advanced techniques of Modeling , Reasoning and Problem Solving, and to know how to design, implement and apply these techniques in the development of intelligent applications, services or systems.
CEA14. Capability to understand the advanced techniques of Vision, Perception and Robotics, and to know how to design, implement and apply these techniques in the development of intelligent applications, services or systems.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA4. Capability to understand the basic operation principles of Computational Intelligence main techniques, and to know how to use in the environment of an intelligent system or service.
CEA6. Capability to understand the basic operation principles of Computational Vision main techniques, and to know how to use in the environment of an intelligent system or service.
CEA8. Capability to research in new techniques, methodologies, architectures, services or systems in the area of ??Artificial Intelligence.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CEP6. Capability to assimilate and integrate the changing economic, social and technological environment to the objectives and procedures of informatic work in intelligent systems.
CEP8. Capability to respect the surrounding environment and design and develop sustainable intelligent systems.

Generical:
CG2. Capability to lead, plan and supervise multidisciplinary teams.

Transversal:
CT5. APPROPIATE ATTITUDE TOWARDS WORK: Capability to be motivated for professional development, to meet new challenges and for continuous improvement. Capability to work in situations with lack of information.

Basic:
CB7. Ability to integrate knowledges and handle the complexity of making judgments based on information which, being incomplete or limited, includes considerations on social and ethical responsibilities linked to the application of their knowledge and judgments.

TEACHING METHODOLOGY

T - Each week it will be a 1.5h theoretical topic exposition class.
P - Each week it will be a 1h practical session.
The rest of the course are devoted to autonomous lectures, programming, and studying.

LEARNING OBJECTIVES OF THE SUBJECT

1.Introduction to object and human recognition
Multi-modal object recognition
Multi-part object recognition
Multi-scale object recognition
Multi-view object recognition
Multi-class object recognition
Multi-label object recognition
Multi-ple data: deep-learning for large scale object recognition
Object Recognition in context: scene understanding and grammars
Human Pose Recovery
Human Behavior Analysis

STUDY LOAD

Type	Hours	Percentage
Hours small group	8,0	8.00
Self study	64,0	64.00
Hours large group	12,0	12.00
Hours medium group	12,0	12.00
Guided activities	4,0	4.00

Total learning time: 100 h

CONTENTS

Introduction to object and human recognition

Convolutional neural networks

Recurrent Neural Networks in Vision

Object detection and segmentation

Human pose estimation

Human Behavior

Transformers / self-attention in Vision

Graph Neural Networks in Vision

ACTIVITIES

Paper presentation

Specific objectives:
1

Related competencies :
CG2. Capability to lead, plan and supervise multidisciplinary teams.
CEA6. Capability to understand the basic operation principles of Computational Vision main techniques, and to know how to use in the environment of an intelligent system or service.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CEP6. Capability to assimilate and integrate the changing economic, social and technological environment to the objectives and procedures of informatic work in intelligent systems.
CEA13. Capability to understand advanced techniques of Modeling , Reasoning and Problem Solving, and to know how to design, implement and apply these techniques in the development of intelligent applications, services or systems.
CEA14. Capability to understand the advanced techniques of Vision, Perception and Robotics, and to know how to design, implement and apply these techniques in the development of intelligent applications, services or systems.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA4. Capability to understand the basic operation principles of Computational Intelligence main techniques, and to know how to use in the environment of an intelligent system or service.
CEA8. Capability to research in new techniques, methodologies, architectures, services or systems in the area of ??Artificial Intelligence.
CEP8. Capability to respect the surrounding environment and design and develop sustainable intelligent systems.
CT5. APPROPIATE ATTITUDE TOWARDS WORK: Capability to be motivated for professional development, to meet new challenges and for continuous improvement. Capability to work in situations with lack of information.
CB7. Ability to integrate knowledges and handle the complexity of making judgments based on information which, being incomplete or limited, includes considerations on social and ethical responsibilities linked to the application of their knowledge and judgments.

Full-or-part-time: 6h 30m
Guided activities: 1h 30m
Self study: 5h

Paper presentation 2

Full-or-part-time: 6h 30m
Guided activities: 1h 30m
Self study: 5h

Exam

Full-or-part-time: 33h
Guided activities: 3h
Self study: 30h

Laboratory 1

Full-or-part-time: 7h
Guided activities: 3h
Self study: 4h

Laboratory 2

Full-or-part-time: 7h
Guided activities: 3h
Self study: 4h

Laboratory 3

Full-or-part-time: 7h
Guided activities: 3h
Self study: 4h

Laboratory 4

Full-or-part-time: 7h
Guided activities: 3h
Self study: 4h

Theoretical class

Full-or-part-time: 22h 30m
Theory classes: 22h 30m

Practical sessions

Full-or-part-time: 15h
Laboratory classes: 15h

GRADING SYSTEM

The course will follow a continuous evaluation consisting in four practical reports (PR) and two in-class presentations (PS). At the end of the course a test exam will be performed (TS). The final score (FS) will be computed as follows:
FS = 0.5 * PR + 0.3 * PS + 0.2 * TS
A minimum score of 3 over 10 points is required for each part PR, PS, and TS in order to compute the final score FS.

BIBLIOGRAPHY

Basic:
- Forsyth, D.A.; Ponce, J. Computer vision: a modern approach. 2nd ed. Boston, Mass.: Pearson Education, 2012. ISBN 0273764144.
- Szeliski, R. Computer vision: algorithms and applications. London: Springer, 2011. ISBN 9781848829350.

Complementary:

- Escalera, S. "Human behavior analysis from depth maps". Articulated motion and deformable objects: 7th International Workshop, AMDO 2012: proceedings [on line]. pp. 282-292 [Consultation: 17/03/2020]. Available on: https://link.springer.com/book/10.1007/978-3-642-31567-1.- Felzenszwalb, P.F.; Girshick, R.B.; McAllester, D.; Ramanan, D. "Object detection with discriminatively trained part-based models". IEEE Transactions on Pattern Analysis and Machine Intelligence [on line]. vol. 32, no. 9, pp. 1627-1645, Sept. 2010 [Consultation: 17/03/2020]. Available on: https://ieeexplore.ieee.org/servlet/opac?punumber=34.- Gatta, C.; Puertas, E.; Pujol, O. "Multi-scale stacked sequential learning". Pattern recognition [on line]. Vol. 44, Issues 10¿11, Oct.¿Nov. 2011, pp. 2414-2426 [Consultation: 17/03/2020]. Available on: https://www.sciencedirect.com/science/journal/00313203.- Hartley, R.; Zisserman, A. Multiple view geometry in computer vision. 2nd ed. Cambridge: Cambridge University Press, 2003. ISBN 0521540518.
- Escalera, S.; Pujol, O.; Radeva, P. "On the decoding process in ternary error-correcting output codes". IEEE Transactions on Pattern Analysis and Machine Intelligence [on line]. 2010, vol. 32, issue 1, pp. 120-134 [Consultation: 19/03/2020]. Available on: https://ieeexplore.ieee.org/servlet/opac?punumber=34.- Escalera, S.; Tax, D.M.J.; Pujol, O.; Radeva, P.; Duin, R.P.W. "Subclass problem-dependent design for error-correcting output codes". IEEE Transactions in Pattern Analysis and Machine Intelligence [on line]. 2008, vol. 30, issue 6, pp. 1041-1054 [Consultation: 19/03/2020]. Available on: https://ieeexplore.ieee.org/servlet/opac?punumber=34.- Madjarov, G.; Kocev, D.; Gjorgjevikj, D.; DÅ¾eroski, S. "An extensive experimental comparison of methods for multi-label learning". Pattern Recognition [on line]. Vol. 45, Issue 9, Sept.2012, pp. 3084-3104 [Consultation: 19/03/2020]. Available on: https://www.sciencedirect.com/science/journal/00313203.- Clocksin, W.F.; Fitzgibbon, A.W.; Torr, P.H.S. (eds.). Proceedings of the British Machine Vision Conference (BMVA). Oxford, UK: British Machine Vision Association, 2005.
- Torralba, A.; Fergus, R.; Freeman, W.T. "80 million tiny images: a large data set for nonparametric object and scene recognition". IEEE Transactions on Pattern Analysis and Machine Intelligence [on line]. Vol. 30, Issue 11, Nov. 2008, pp. 1958-1970 [Consultation: 19/03/2020]. Available on: https://ieeexplore.ieee.org/servlet/opac?punumber=34.- Oliva, A.; Torralba, A. "The role of context in object recognition". Trends in Cognitive Sciences [on line]. Vol. 11, Issue 12, Dec. 2007, pp. 520-527 [Consultation: 19/03/2020]. Available on: https://www.sciencedirect.com/science/journal/13646613.- IEEE 11th International Conference on Computer Vision, 14-21 Oct. 2007. Rio de Janeiro, Brazil: IEEE Computer Society, 2007.
- Zhu, S.-C.; Mumford, D. A stochastic grammar of images. Hanover, MA: Now Publishers, 2007. ISBN 9781601980601.
- Conference on Computer Vision and Pattern Recognition (CVPR), 20-25 June 2011. Providence, RI, USA: IEEE, 2011.
- Starner, T.; Weaver, J.; Pentland, A. "Real-time American sign language recognition using desk and wearable computer based video". IEEE transactions on pattern analysis and machine intelligence [on line]. 1998, vol. 20, issue 12, pp. 1371-1375 [Consultation: 19/03/2020]. Available on: https://ieeexplore.ieee.org/servlet/opac?punumber=34.- Proceedings of the IEEE.
- Burkhardt, H.; Neumann, B. (eds). European Conference on Computer Vision: ECCV 1998: Computer Vision. Berlin: Springer, 1998.
- ICML '05: Proceedings of the 22nd international conference on machine learning. Bonn, Germany: International Machine Learning Society, 2005.
- Goodfellow, I.; Courville, A.; Bengio, Y. Deep learning [on line]. Cambridge, Massachusetts: The MIT Press, 2016 [Consultation: 07/03/2024]. Available on: https://www.deeplearningbook.org/. ISBN 9780262035613.
- LeCun, Y.; Bengio, Y.; Hinton, G. "Deep learning". Nature [on line]. 521, pp. 436¿444(2015) [Consultation: 19/03/2020]. Available on: https://www-nature-com.recursos.biblioteca.upc.edu/articles/nature14539.