

COMPUTER ENGINEERING (LM55)

(Lecce - Università degli Studi)

Teaching COMPUTER VISION

GenCod A005792

Owner professor Cosimo DISTANTE

Teaching in italian COMPUTER VISION **Course year** 1

Teaching COMPUTER VISION

Language INGLESE

SSD code ING-INF/03

Curriculum PERCORSO COMUNE

Reference course COMPUTER ENGINEERING

Course type Laurea Magistrale

Location Lecce

Credits 9.0

Semester Primo-Semestre

Teaching hours Ore-Attività-frontale: 81.0

Exam type Orale

For enrolled in 2019/2020

Assessment Voto-Finale

Taught in 2019/2020

Course timetable

<https://easyroom.unisalento.it/Orario>

BRIEF COURSE DESCRIPTION

Computer Vision today is everywhere in our society and images have become pervasive, with applications in several sectors; just to mention some in: apps, drones, healthcare and precision medicine, precision agriculture, searching, understanding, control in robotics and self-driving cars. The course introduces the basics of image formation, reconstruction and inferring motion models, as well as camera calibration theory and practice.

Recent developments in neural networks (Deep Learning) have considerably boosted the performance of the visual recognition systems in tasks such as: classification, localisation, detection, segmentation etc. Students will learn the building blocks of a general convolutional neural network, the way how it is trained and optimized, how to prepare a dataset and how to measure the final performance.

REQUIREMENTS

No prior experience with computer vision is assumed, although previous knowledge of visual computing or signal processing will be helpful. The following skills are necessary for this class:

- Math: Linear algebra, vector calculus, and probability. Linear algebra is the most important.
- Data structures: Students will write code that represents images as feature and geometric constructions.
- Programming: A good working knowledge. All lecture code and project starter code will be Python, and Pytorch for Deep Learning, but student familiar with other frameworks such as tensorflow is ok.

COURSE AIMS

Upon completion of this course, students will:

1. Be familiar with both the theoretical and practical aspects of computing with images;
2. Have described the foundation of image formation, measurement, and analysis;
3. Have implemented common methods for robust image matching and alignment;
4. Understand the geometric relationships between 2D images and the 3D world;
5. Have gained exposure to object and scene recognition and categorization from images;
6. Grasp the principles of state-of-the-art deep neural networks; and
7. Developed the practical skills necessary to build computer vision applications.

TEACHING METHODOLOGY

Teaching is based on theoretical and practical lectures. The student will write in python algorithms taught in class

ASSESSMENT TYPE

Oral session. The student will explain the developed project and shall answer two or more questions regarding theoretical aspects of the studied topics

ASSESSMENT SESSIONS

The student must develop a project by choosing a practical simple application with some algorithms done during the course. The choice is at total disposal of the student, as well as the fact of developing it in group or solo. In group setting the students must proof their own activities developed in the common project application.

The final examination is based on oral assessment of the topics covered during lectures.

OTHER USEFUL INFORMATION

For the LAB practice, students may use for the deep learning development the Google Colab or Cloud Platform.

FULL SYLLABUS

Introduction to Computer Vision
Image Formation
2D and 3D geometric primitives - Projections
image enhancement
LAB Introduction to Python and Operations with images
Color perception, color spaces and processing
Image Filtering
image pyramids and blending
Local feature detector
LAB SIFT with MatLab Find image rotation & scale SURF Object Detection
Image Alignment I- warping, homography estimation direct linear transform
Image Alignment II- robust motion estimation with Ransac - perspective n point problem.
Registration examples: face recognition, medical imaging
Camera Calibration - distortion models and compensations - linear methods for camera parameters. Calibration with a checkerboard
LAB Mosaicking with SURF Face Detection and Tracking (Nose Skin) Face Detection and Tracking with Kanade - Lucas - Tomasi feature tracker
Motion Analysis and background modelling, application to intelligent videovigilance
Multiview geometry - Epipolar geometry, position error estimation, stereo rig, Essential matrix estimation, rectification, Reconstruction, correspondence problem, weak calibration and ransac estimation of fundamental matrix
LAB - Camera calibration
Image Classification - Key nearest neighbor, linear classifiers
Image Classification - loss functions, optimization with stochastic gradient descent
LAB - Stereo calibration and reconstruction
Image Classification - backpropagation and neural networks, computational graphs and gradient estimation
Image Classification - Convolutional Neural Network architecture
Image Classification - CNN activation functions, data preprocessing, weight normalization, batch normalization, monitoring the learning process, hyperparameter optimization, Regularization (Dropout, drop connect, fractional pooling, cotout, mixup)
Image Classification - CNN activation functions, data preprocessing, weight normalization, batch normalization, monitoring the learning process, hyperparameter optimization, Regularization (Dropout, drop connect, fractional pooling, cotout, mixup)
Image Classification - Object detection and Image segmentation with Convolutional neural networks. Introduction to auto-encoders
LAB - Introduction to Pytorch framework
LAB - Deep learning applications to object detection (Yolo and Faster R-CNN)
LAB - Deep Learning application to segmentation with mark R-CNN

REFERENCE TEXT BOOKS

There is no requirement to buy a book. The goal of the course is to be self contained, but sections from the following textbooks will be suggested for more formalization and information.
The primary course text will be Rick Szeliski's draft [Computer Vision: Algorithms and Applications](#); we will use an online copy of the June 19th draft. A copy and link will be provided in website. The secondary text is Forsyth and Ponce, [Computer Vision: A Modern Approach \(new Edition coming out in 2020\)](#)
Deep Learning, MIT Press book, Ian Goodfellow and Yoshua Bengio and Aaron Courville