COMPUTER ENGINEERING (LM55)

(Lecce - Università degli Studi)

Teaching COMPUTER VISION

GenCod A005792

Owner professor Cosimo DISTANTE

Teaching in italian COMPUTER VISION

Teaching COMPUTER VISION

SSD code ING-INF/03

Reference course COMPUTER ENGINEERING

Course type Laurea Magistrale

Credits 9.0

Teaching hours Front activity hours: 81.0

For enrolled in 2021/2022

Taught in 2021/2022

Course year 1

Language ENGLISH

Curriculum PERCORSO COMUNE

Location Lecce

Semester Second Semester

Exam type Oral

Assessment Final grade

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 agricolture, 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 os 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
Color perception, color spaces and processing
Image Filtering
LAB Introduction to Python and Operations with images
interpolation, optimization, image pyramids and blending
Machine learning
LAB (pytorch basics? Dataloaders, ML?, T-sne?)
loss functions, optimization with stochastic gradient descent
backpropagation and neural networks, computational graphs and gradient estimation
Convolutional Neural Network, CNN activation functions, data preprocessing, weight normalization, batch normalization, monitoring the learning process, hyperparameter optimization, Regularization (Dropout, drop connect, fractional pooling, cotout, mixup)
CNN architectures (Alexnet, VGG, GoogleNet, ResNET, DenseNet, SENet, EfficientNet), Siamese Architectures (applications to face verification, people and vehicle re-identification)
LAB CNN
Recurrent neural networks, Attention mechanisms
Object detection and segmentation
LAB - object detection - segmentation
Generative Models
edges, feature matching Ransac and alignment
Optical flow, 3D, Depth perception and stereo
SLAM/SfM
Camera Calibration - distortion models and compensations - linear methods for camera parameters. Calibration with a checkerboard
LAB camera calib
3D shapes

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 2nd Edition 2022; we will use an online copy (fill the form) at this link.
We will be using Piazza for all course notes, homework and final project.
A copy and link will be provided in website.
A textbook for Deep Learning with Pytorch script can be accessed at this link
Deep Learning, MIT Press book, Ian Goodfellow and Yoshua Bengio and Aaron Courville