uc3m | Universidad Carlos III de Madrid

Speech, audio, image, and video processing applications

Academic Year: (2022 / 2023)   **Review date: 12-05-2022**

Department assigned to the subject: Department of Signal and Communications Theory

Coordinating teacher: GALLARDO ANTOLIN, ASCENSION

Type: Electives ECTS Credits : 3.0

Year : 2 Semester : 1

OBJECTIVES

OBJECTIVES

Similarly to other Master's elective courses, the student will acquire a greater specialization in different areas of Telecommunications technologies. Particularly, this course covers the following Signal Processing skills:

1.- General/Cross-curricular objectives
1.1. General basic knowledge
1.2. Analysis and synthesis abilities
1.3. Capability of applying the knowledge they have acquired
1.4. Problem-solving skills
1.5. Capability of integration of knowledge

2.- Specific learning objectives
(Knowledge-related objectives)
2.1. Mathematical basis of signal processing
2.2. General knowledge on potential speech/audio/image/video processing applications
2.3. Basic subsystems of speech/audio/image/video processing applications

(Instrumental learning objectives)
2.5. Use of speech/audio/image/video processing software
2.6. Mastering of basic processing tools
2.7. Solving speech/audio/image/video processing problems by using several basic tools

(Attitudinal learning objectives)
2.8. Individual- and team-work
2.9. Decision-making
2.10. Analysis and problem-solving capabilities

DESCRIPTION OF CONTENTS: PROGRAMME

The goal of this subject is to provide the student with an introduction to recent signal processing techniques with application to speech, audio, image and video. To that end, a Project-Based Learning Approach is followed. The emphasis is put on lab exercises, so that the student can be assessed according to her work on a mini project.

1.- Course presentation

2.- Introduction to Deep Learning
      2.1. Neural Networks
      2.2. Deep Neural Networks (DNNs)

3.- Fundamentals and Techniques of Image Processing
      3.1. Digital Representation of the Image. Color spaces.
      3.2. Point Operations. Filtering.
    3.3. Convolutional Neural Networks (CNNs)

4.- Fundamentals and Techniques of Speech and Audio Processing
      4.1. Digital Representation of Speech and Audio Signals. Spectrogram.
    4.2. DNNs for Speech and Audio Processing.

4.3. Recurrent Neural Networks (RNNs)

5.- Fundamentals and Techniques of Video Processing
   5.1. Digital Representation of Video Signals.
   5.2. DNNs for Video Processing.

6.- Case Studies and Applications of Speech, Audio, Image and Video Processing

LEARNING ACTIVITIES AND METHODOLOGY

Two teaching activities are proposed: theoretical classes with examples and lab exercises.

THEORETICAL CLASSES WITH EXAMPLES (2 ECTS)

The theoretical class will be given in the blackboard, with slides or by any other means to illustrate the concepts of the lectures. In these classes the explanation will be completed with examples (AF1, MD1).

In these sessions the student will acquire the basic concepts of the course. It is important to highlight that these classes require the initiative and the personal and group involvement of the students (there will be concepts that the students should develop by themselves) (AF3, MD3).

LABORATORY EXERCISES (1 ECTS)

Some basic selected concepts learnt during the course are applied in the lab. The students should participate actively in the exercise implementation.

There will be two types of lab exercises:
-        Guided lab exercises: getting used to speech, audio, image and video processing using Python (AF2, AF4, MD4).
-        Final Project: image or speech/audio processing problem to be solved in groups (AF5, AF6, AF7, MD2, MD5).

ASSESSMENT SYSTEM

SE2.- Individually or in group works performed during the course.

In the ordinary call, the final grade of the subject will be based on the following submissions:

- 3 tests about guided lab exercises (30%).
- Final Project (70%).

In the extraordinary call, the final course grade is established by the delivery of the lab exercises and the final project.

% end-of-term-examination:	0
% of continuous assessment (assigments, laboratory, practicals…):	100

BASIC BIBLIOGRAPHY

 - Aurélien Géron Hands-On Machine Learning with Scikit-Learn and TensorFlow, O'Reilly Media, 2017

 - Francois Chollet Deep Learning with Python, Manning Publications, 2017

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

 - Pradeepta Mishra PyTorch Recipes A Problem-Solution Approach, Berkeley, CA : Apress : Imprint: Apress, 2019

ADDITIONAL BIBLIOGRAPHY

 - Ben Gold (Author), Nelson Morgan (Author), Dan Ellis (Author) Speech and Audio Signal Processing: Processing and Perception of Speech and Music., Wiley, 2011
 - Christopher M. Bishop Pattern Recognition and Machine Learning, Springer, 2006

 - Forsyth & Ponce Computer Vision: A Modern Approach, Pearson, 2012

 - Gonzalez and Woods Digital Image Processing 4th Edition, Pearson, 2018

 - Wilhelm Burger and Mark J. Burge Principles of Digital Image Processing: Core Techniques, Springer-Verlag, 2009