UNIVERSITÀ DEGLI STUDI DI GENOVA

ISTITUTO ITALIANO DI TECNOLOGIA

Università degli Studi di Genova – Istituto Italiano di Tecnologia

Corso di Dottorato "Sciences & Technologies for Electronics & Telecommunication"

Curriculum "Computer Vision, Pattern Recognition and Machine Learning"

Anno Accademico 2017-2018
Ciclo XXXIII

7 positions available with scholarship

## Research Themes

## 1. Computer vision for behavioral analysis and activity recognition

Tutors: Vittorio Murino

Description: Study and development of techniques and systems for the analysis of behaviours, actions, expressions/emotions, and social signals in general, referred to both single persons and groups. In this context, methods for tracking, recognition, and classification of persons and objects starting from images and/or sequences acquired from cameras distributed in the environment in several sparse locations, and from other types of sensors (e.g., microphones) will be considered. The main goal is to exploit hints and findings coming from social sciences to capture and model human behaviour. Computer vision and machine learning constitute the focus of this research, and particular attention will be specifically devoted to deep learning models. Requirements: degree in robotics, bioengineering, computer science, computer engineering, or related disciplines, attitude for problem solving, c++ programming. A background on machine learning is an asset.

## 2. Computer vision for the prediction of human intentions and activities

Tutors: Vittorio Murino

Description: We want to go further the analysis of human movements and explicit behaviours to design methods and algorithms for the prediction of human intentions. Experiments are already running consisting in performing simple human-object interactions with a goal which is not explicitly stated nor viewed, and the task is to classify the actual intention subsuming that action from 2D and 3D data. We also plan to carry out other visual experiments considering more complex human interactions involving many subjects, while predicting the final outcome of the observed interplay. Computer vision and machine learning, and more specifically convolutional neural networks (CNN) and Long-Short Term Memory (LSTM) networks, will be most investigated models.

## 3. Part-based human body modeling for Socially-Aware Computer Vision

Tutors: Vittorio Murino, Alessio del Bue

Description: To recognize and interpret human nonverbal behavior it is fundamental to try to identify the subjects involved, especially in the wild, that is in real situations. To this end, part-based human body modeling is a mandatory task aimed at extracting from images the different components of the human body, like head, torso, arms, legs, etc., so as to estimate posture, gesture and gaze, all social cues widely known as useful hints to classify behavior and recognize situations. Further, real time tracking of body parts is equally important to increase such recognition performances, possibly adding prediction functionalities to these algorithms. Computer vision and machine learning constitute the focus of this research, and particular attention will be specifically devoted to deep learning models. Applications in the contest of surveillance and security as well as in the biomedical field (e.g., monitoring of elderly people) are envisioned.

## 4. Crowd behavioral analysis and event recognition

Tutors: Vittorio Murino

Description: Study and development of techniques and systems for the analysis of behaviours, events, social signals in general, referred to a large mass of people (crowd). The analysis and modelling of behaviour of groups and crowd seen as single entities will be considered. There is evidence that large groups of people and crowd are characterised by a collective behaviour which may emerge in different situations and can lead to interesting outcome from the point of view of the surveillance applications, and may help to detect and predict coming events. Machine learning as well as computer vision constitute the focus of this research, starting from early work in human body modelling/tracking to novel social force models able to grasp the complex dynamics of the human flow. Particular attention will also be specifically devoted to deep learning models.

## 5. Re-identification

Tutors: Vittorio Murino

Description: Study and development of biometric techniques for scene analysis and understanding. The research will mainly focus on person characterization, with possible focus on the usage of soft biometrics cues (3D, attributes) and in challenging conditions (e.g., crowd). The idea is to recover the identity of persons as viewed in different times and places, also considering face/body attributes, the so-called re-identification problem. Not only optical cameras will be used, but other information derived from different sensors may also be utilized (e.g., range). Moreover, the use of a pan-tilt-zoom (PTZ) camera able to identify specific features of a single person or groups, or addressing non-cooperative face recognition at distance, could be subjects of investigation. The robustness to environmental (real) conditions and the non-cooperation of the subjects are the main features to which the developed techniques will have to cope with to deploy this technology in real scenarios. Computer vision and machine learning constitute the focus of this research, and particular attention will also be specifically devoted to deep learning models.

## 6. Time-lapse Computer Vision for long-term 3D learning

Tutor: Alessio Del Bue

Description This topic is related to the research and implementation of smart systems that leverage the big data information coming from 24h/7d video streams in order to automatically obtain computational models of the environment (3D surface, photometric and object/human motion models). Extracting such information from a single view is a complex and ill-posed task, but potentially disruptive because of its applicability to any video camera installed outdoor and indoor. We will use recent advancements in photometry modelling, geometrical reasoning and deep learning to untangle the complexity of the problem and to provide a working system that can be generalised to several scenarios. In particular, we will leverage photometric cues (gradual change of lighting), motion patterns of dynamic objects (from people detections and motion flows) and common knowledge about scene context in order to infer the dynamic 3D structure of the scene. The implemented system will be tested directly on the camera network installed in IIT and it will use the HPC servers running state of the art matrix factorization and deep learning algorithms.

## 7. 3D scene understanding and spatial reasoning for autonomous systems

Tutor: Alessio Del Bue

Description This Computer Vision related theme will study novel geometrical approaches and deep network architectures to make reliable hypothesis about the composition and its 3D geometry using images only. The goal is to implement a system that can provide high performing solutions to the Visual Question and Answering (VQA) problem – a visual Turing test to assess the capabilities of machines to obtain a human-like description of the visual world. VQA is a novel research topic that might have soon a tremendous impact in robotics, human-computer interaction, autonomous driving and search/retrieval systems. The research theme will in particular focus on studying new models for geometrical 3D reasoning in order to infer the scene layout, attributes (colour, material) and existence of any objects in a single image or video sequence. The system, being prompted by a user (voice or text), will be able to answer and possibly learn about its mistake using strategies based on reinforcement learning or similar approaches. This study will be made in close collaboration with the Robotic labs in IIT to enable the current robotic platforms with cutting-edge scene understanding capabilities.

## 8. Biomedical imaging

Tutor: Diego Sona, Vittorio Murino

Description The wide adoption of biomedical sensors (e.g., MRI, TAC, SPECT, MEG, EEG/EMG, Electron and Fluorescent Microscopy, etc.) in various clinical and biological investigations is fostering an increasing interest in advanced tools supporting the expert in the analysis and interpretation of the produced big amounts of data. In this perspective, this theme will address all lines of research related to the development of computer aided detection (CAD) systems, ranging from image processing and analysis, segmentation, object detection up to automatic determination of diseases' biomarkers, etc. Particular attention will be devoted to structural data (3D) and functional data (multivariate time-series and videos). The development of such CAD tools will also require the design of novel computer vision and pattern recognition techniques for advanced and automatic analysis of biomedical data. In this line, particular attention will be given to deep learning models, specifically to convolutional neural networks (CNN) and Long-Short Term Memory (LSTM) networks.

## 9. Connectomics

Tutor: Diego Sona, Vittorio Murino

Description The brain is a complex interconnected system that can be investigated at different granularities, from the macro-scale level (i.e., functional interactions between brain areas) using non-invasive techniques (e.g., MRI; MEG, EEG, etc.) down to the meso-scale level (functional interactions between neurons in a complex network) thanks to the recent evolution of high-density multi-electrode arrays (MEAs). Nevertheless, independently of the granularity, the relationships between the functional dependencies and the structural connectivity remains still unresolved. In this research theme the topic under investigation is therefore the integration of functional and structural connectivity information with different applications ranging from the characterization of mental diseases' biomarkers (e.g., schizophrenia, autism, multiple sclerosis, etc.) to studies for the analysis of treatments, from the investigation of cognitive functions to their correlation with behavioral patterns. The integration of functional and structural connectivity information will require the development of tools for the analysis of multi-modal data at macro-scale (e.g., functional MRI, diffusion imaging, etc.) and meso-scale (MEA electrical activity, fluorescent microscopy imaging). Machine learning and pattern recognition techniques will be the principal investigated domains, with particular attention to deep learning models, aiming at integrating/fusing these multimodal information sources.

## 10. Animal behavior analysis

Tutors: Diego Sona, Vittorio Murino

Description Behavioural neuroscience is a fundamental research field studying the biological bases of behavior providing insights into the mechanisms of the nervous systems producing anomalous behavior, and experimental subjects mostly involve animals. We are, therefore, involved in a multidisciplinary research activity, which need the development of techniques and systems for the automatic analysis of actions, postures and social behaviours of mice in home cages. In this framework, we aim at designing methods for tracking 24/7 the mice and objects from video recorded from multiple camera with different orientations. Computer vision and machine learning constitute the focus of this research, with particular interest in methodologies exploiting the spatio-temporal information. To date, we already have developed a series of algorithms based on Restricted Boltzmann Machines, but the investigation of other deep learning models is envisioned.