PhD Program in
Science & Technologies for Electronics & Telecommunication Engineering

YEAR 2023 - 2024

Curriculum in

Computer Vision, Automatic Recognition and Learning

Visione Computazionale, Riconoscimento e Apprendimento Automatico

ATTENTION

The PhD application also implies submitting a research proposal under one or more themes chosen among those below indicated. To write a proper research proposal, please follow the instructions indicated in the following file:
https://pavisdata.iit.it/data/phd/ResearchProjectTemplate.pdf

Research themes

Theme A
3D scene understanding with geometrical and deep learning reasoning
Tutor: Alessio Del Bue, Pietro Morerio

Classical multi-view geometry problems make use of geometric reasoning to infer the scene's 3D structure and its dynamic. These approaches often neglect the semantic composition of the scene that instead provides important cues about objects motion and their current spatial configuration. Instead, such context and semantic information can be given by current deep learning architectures but very few works attempted to merge geometrical reasoning with such semantic information. This research theme will have the aim to bridge this gap and to provide solutions that can be applied on intelligent systems for robotics and autonomous driving. This research will study different methods and tools involving scene graphs, object detection with either multi-view images or 3D data, deep learning methods for scene representation and large-scale 3D reconstruction in dynamic environments. The research activities will build over and expand the current expertise on these problems developed under the Spatial AI topics at PAVIS research line.

Theme B
Multi-modal learning
Tutor: Alessio Del Bue, Pietro Morerio, Vittorio Murino

This topic is related to the research and implementation of algorithms that leverage multimodal data, namely data coming from different sensors, for general purposes such as classification, recognition and, more in general, scene understanding. In particular, we primarily aim at exploiting optical (RGB) sensors, range or depth sensors, thermal sensors (thermal and near-infrared cameras), acoustic sensors to detect persons and objects, tracking and classifying objects, events and behaviors in general. The main focus will be on learning and developing new deep learning methods, as particularly suitable to merge heterogeneous information coming from different sensor modalities.

Theme C

Self-Supervised and Unsupervised Deep Learning

Tutor: Alessio Del Bue

One of the key factors behind the recent popularity of deep learning algorithms is the possibility of leveraging a large corpus of labelled data. Despite gathering massive amounts of data is nowadays not problematic, data annotation is surely a major bottleneck. In fact, not only is it time- consuming and economically expensive, but it is also prone to errors since requiring human intervention. This research topic focuses on how to relax the level of supervision required to develop (deep) machine learning algorithms. The ultimate goal is to devise new computational techniques which exploit and discover geometrical inner properties of the data (e.g., self-supervised learning, clustering), while also considering the transfer of knowledge from existing labelled datasets in order to recognize categories and classes for which a little or even no labelled data is provided (few-shot or zero-shot learning).

Theme D

Generative models for human and scene generation

Tutor: Alessio Del Bue, Pietro Morerio

Learning generative models that can explain complex data distributions is a long-standing and challenging problem in machine learning research. These problems are particularly evident in the field of computer vision, where image and 3D data generation is complex due to the high dimensionality of data. Generative models for images and 3D data are not only important for unsupervised feature learning but also enable a wide range of applications such as image editing and content manipulation. With recent advances in Generative Adversarial Networks, Transformers and Diffusion Models, the generation of realistic images and 3D structures has reached new outstanding results, but yet limitations are present. Our research aims at improving and better understanding such models, as well as exploring better alternative methods for image and 3D data generation.

Theme E

Novel Graph Operators for learning on large-scale and temporal data

Tutor: Alessio Del Bue, Stefano Fiorini

The majority of the real-world phenomena are ruled by complex internal relationships that are better represented by a graph. For instance, electronic commerce relies on graphs to model the interaction between users and products to make accurate recommendations, and drug discovery

uses interaction graphs to model molecule bioactivity.
In the Deep Learning field, Graph Convolutional Networks (GCNs) are the state-of-the-art tools used to solve tasks on data that exhibit a graph-like structure. They are specifically designed to capture and exploit the graph's topology that models the interaction between different entities in a complex system via a graph-based convolution operator. In spectral-based GCNs, a category of GCNS, one of the crucial components is constructing operators (satisfying certain mathematical properties) for convolutions on graphs. We are interested in investigating the design of novel operators to address several limitations observed in the state of the art, including handling multigraphs, temporal graphs, and hypergraphs. Moreover, we are interested in exploring the application of GCNs to large-scale graphs

Theme F
Domain adaptation and Generalization
Tutor: Vittorio Murino, Pietro Morerio

This research theme focuses on the development of machine learning models for computer vision that can be deployed into the wild. More specifically, one drawback of modern learning systems is that they strongly rely on the characteristics of the data they are trained with. This results in models that poorly generalize to context unexplored during training (for example, consider a home robot that is deployed in a new house). To overcome this liability, two main strategies are domain adaptation and domain generalization. In the former case, we can leverage non-annotated samples from a desired scenario during training, and design models that better adapt to that domain. In the latter, the goal is generalizing to domains that are utterly unseen during training. The design of new training procedures to solve these tasks and the identification of novel application settings represent the main directions of this research.