PhD Thesis Proposal
Recruitment Campaign 2022

Title: Computational Visual Perception and Learning Framework for Semantic Scene Understanding

Brief description of the topic

The analysis and interpretation of dynamic scenes is one of the most active and challenging research topics in computer vision, robotics and artificial intelligence. Various theories and concepts have been introduced since the early development of the computer vision field. Recently, with the renewed interest in artificial intelligence and the development of intelligent sensors technologies as well as high performance computing resources, we are witnessing a proliferation of deep learning-based solutions. Most of the current developed algorithms for scene analysis and interpretation are based on deep neural networks (DNNs).

Indeed, even if we are far from being able to model and justify in a convincing and explainable mathematical framework the use of DNN-based approaches, it remains that this trend is really supported by the efficient solutions developed so far for the resolution of various problems and in particular in the field of computer vision and robotics, e.g., face recognition, pedestrian detection and pose estimation. The main goal of this research project is to contribute to the reformulation of new approaches, based on convolutional neural network architectures, for the processing and analysis of visual information. More specifically, it concerns scene analysis for solving some real-world problems of great interest. The main idea is to combine the perceptual approach, based on the modelling of retino-cortical mechanisms, with deep learning techniques. The originalities and contributions made in this work will be highlighted through several concrete actions on both the theoretical and application levels. One of the objectives of this project is to redefine the loss functions and the different connectionist architectures by drawing on sufficiently established knowledge on the architecture and functioning of the visual cortex. Theoretical outputs are, on the one hand, a contribution to the development of explainable and justified deep learning models; and, on the other hand, the advancement of research in the field of scene analysis and understanding, through a more elaborate modelling of the processing and analysis mechanisms of visual signals. In terms of applications, two major fields will be considered: public health and security. The methods developed will be more particularly evaluated and validated on concrete cases such as diagnosis in the context of medical imaging, computer-guided surgery and security based on video surveillance. In summary, in this thesis, we will introduce a new framework by building a bridge between computational modeling of perceptual vision and machine learning paradigm for solving real-world computer vision problems.

MAIN SUPERVISOR (DIRECTEUR DE THESE)

Prof. Azeddine Beghdadi

L2TI, Institut Galilée, Université Sorbonne Paris Nord

e-mail : beghdadi@sorbonne-paris-nord.fr

Personal Webpage: https://www-l2ti.univ-paris13.fr/~beghdadi/

Hosting institution

The recruited candidate will be hosted at the “Laboratoire de Traitement et Transport de l’Information” (L2TI) at Université Sorbonne Paris Nord (USPN), Villetaneuse, France.