Grid based CNNs for spatially variant image categorisation
Internship proposal
U2IS Robotics & Autonomous Systems Lab.
ENSTA ParisTech (Palaiseau – FRANCE)

The development of deep Convolutional Neuron Networks (CNN) has brought a long step forward to the problem of supervised classification, particularly in images [1] . However, these methods suffer of long training and testing times, requiring the use of GPUs and highly parallel processing. Typically, they consist of taking a large dataset of images and, assuming position invariant of the image statistics, learn filter banks that optimize the representation of the images.
On the other hand, the second order statistics of images have been shown to be position dependent [2,3], and more recently, also third and fourth order statistics have been shown to depend on the position in the visual field [4]. In the application side, grid based approaches have been shown to be highly efficient in representation and classification of the gist of images [5] .

In autonomous vehicles, the semantic classification of the environment (city, high-way, country side, etc) is crucial for the successful control of the system, furthermore, it should be in real time to contribute to a save driving mode.
To solve this problem, in this project we will study the impact of grid based approaches on CNNs, their performance and their computational cost on semantic classification of images taken from a camera embedded in a vehicle [6].

Contacts : Daniela.Pamplona@ensta-paristech.fr, Antoine.Manzanera@ensta-paristech.fr

[1] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
[2] Pamplona, D., Triesch, J. and Rothkopf, C.A., 2013. Power spectra of the natural input to the visual system. *Vision research*, *83*, pp.66-75.
[3] Pamplona, D., 2014. Ecological perspectives on local image statistics. PhD Thesis
[4] Daoud, B., 2017. Les Statistiques d'ordre supérieur des Images naturelles, Rapport de PRe, ENSTA Paristech
[5] Oliva A., Torralba A., 2001.Modeling The Shape of the Scene: a Holistic Representation of the Spatial Envelope, International Journal of Computer Vision
[6] Geiger, A., Lenz, P., Stiller C. and Urtasun R., 2013. Vision meets Robotics: The KITTI Dataset. International Journal of Robotics Research