Title: Deep Learning to Find Clusters
Prof. James G. Bartlett
APC – Université Paris Diderot

The goal of this thesis is to apply deep learning to the detection of galaxy clusters in astronomical surveys. Clusters are the most massive objects in the universe, hosting hundreds of galaxies and a gaseous intra-cluster medium (ICM) at temperatures of tens of millions of Kelvin. They are detected as galaxy over-densities in optical/near-infrared (NIR) imaging, through the X-ray emission of their ICM or via the Sunyaev-Zeldovich (SZ) effect, a distortion of the cosmic microwave background spectrum generated by photon-electron scattering in the ICM. Clusters serve as powerful probes of dark energy, dark matter, neutrino physics and gravity, and as valuable laboratories for understanding galaxy formation. As a consequence, cornerstone astronomical surveys, like the future *Euclid* and WFIRST missions, and the Large Synoptic Survey Telescope (LSST), produce cluster catalogs as one of their primary science products.

The objectives of the thesis will be to develop deep neural networks to:

1. Improve the reliability of the *Planck* mission galaxy cluster catalog based on the nine-band millimeter-wave measurements of the SZ effect.
2. Improve the reliability of the *Planck* mission galaxy cluster catalog by combining *Planck* data with NIR galaxy catalogs from the WISE mission.
3. Detect clusters in simulated optical/NIR galaxy catalogs representative of the future *Euclid* and WFIRST missions, and LSST.

As part of the third objective, we will develop a practical pipeline for use in the analysis of the real data from the *Euclid* and LSST surveys, planned to begin operation just after the completion of the thesis.

Standard cluster detection methods are built on simplified models of clusters and typically rely on linear filtering and thresholding (i.e., signal-to-noise cuts). Galaxy cluster detection is an interesting application of deep learning because it presents an opportunity to apply potentially more efficient non-linear combinations of complex cluster-like features that a deep network would learn when trained on known clusters.

In this light, our project is application of deep learning to a one-class classification problem: We seek to extract a catalog of a specific object type from a large dataset containing many “not my object” configurations. We train our networks by providing examples of the object of interest taken from smaller datasets with more complete information that enables accurate identification.

The APC laboratory is heavily invested in LSST and the *Planck* and *Euclid* missions, with Prof. Bartlett co-leading the Galaxy Cluster Science Working Group for the Euclid Consortium. The student will work in a dynamic environment with colleagues at APC and including a group of laboratories in the Paris area developing a cluster cosmology pipeline for the *Euclid* mission.