Deep learning the Euclid data

Name of the laboratory: Laboratoire d'Astrophysique de Marseille (LAM)

Thesis advisor: Nicolas Martinet
Email: nicolas.martinet@lam.fr

Co-advisor: Stéphane Arnouts

Subject's title: Deep learning the Euclid data

Subject description:

Context:

Deep learning is a powerful new tool for extracting interesting features from large datasets. With the maturity it has now acquired in observational cosmology, we propose to exploit deep learning to enhance the primary science capability of the Euclid mission. Euclid is an ESA satellite whose main objective is to understand the accelerating expansion of the Universe using the cosmic shear signal, i.e. the gravitational lensing distortion (called shear) of background galaxies due to the foreground large-scale structures of over 1.5 billion galaxies. For such an analysis, it is necessary to accurately measure the distances of the galaxies, their shapes (which carry the shear signal), and to relate them to cosmological parameters. The PhD student will develop these three research axes, in close collaboration with the new "Deep Learning team" of the Laboratoire d'Astrophysique de Marseille, and under the supervision of Nicolas Martinet (Astronome Adjoint) and Stéphane Arnouts (Directeur de Recherche CNRS).

Methodology and Work Program:

Photometric distances (photo-z) to galaxies cannot be measured from Euclid data alone as they require a combination of observations at many wavelengths. The student will first investigate the use of deep learning techniques to measure these photo-z's, focusing on the Kilo Degree Survey (membership via Nicolas Martinet) which has the most homogeneous current set of observations in the visible (u, g, r, i) and infrared (z, Y, H, J, Ks) bands over 1000 deg2. He/She will extend the deep learning photo-z's obtained by our team in the SDSS (Pasquet et al. 2019) to higher redshifts, allowing us to reveal the first cosmic web filaments at high redshift (up to z=1). He/she will then combine KiDS with Euclid's exquisite near-infrared observations to improve the redshift calibration needed for Euclid's cosmic shear analysis.

The student will then measure the shape of the galaxies and in particular the shear signal directly from the Euclid images. The deep learning approach will require the construction of a challenging set of image simulations to perform the training. By combining the individual shape measurements with the detected filaments mentioned above, he/she will characterize the possible alignments of galaxies with the surrounding cosmic web filaments, as suggested by hydrodynamic simulations but not yet confirmed by observation beyond the local Universe. These intrinsic galaxy alignments are a major source of systematic bias for cosmic shear analyses, which are still poorly constrained observationally. Finally, deep learning will allow us to measure the statistical shear signal for small patches of the sky, elegantly solving the problem of blending in shear measurements (Euclid collaboration: Martinet et al. 2019).

The last part of this thesis will focus on measuring cosmological parameters directly from galaxy distortion maps. In Martinet et al. 2021a, we showed that using a particular convolution of these maps increases the constraining power of Euclid on the dark energy equation of state (which governs the nature of the accelerating expansion) by a factor of 3. Relying on deep learning should further improve this gain as the neural network is trained to optimize a series of convolutions of the distortion maps. More precisely, the information probed by these methods is the non-Gaussian part of the matter distribution that appears at small cosmological scales due to the non-linear growth of the large-scale structures. This non-Gaussian information, currently neglected by traditional cosmic shear estimators, should further enhance the success of the mission.

Research environment:

The three parts of the thesis can be treated independently and constitute three major axes of Euclid science. The student will benefit greatly from the LAM environment with a large team of experts in galaxy evolution and cosmology and many members of the Euclid Consortium. In addition to developing skills in deep learning approaches, the selected candidate will obtain high-level training in cosmic shear, which will allow him/her to play a leading role in the coordination of Euclid's primary science. This thesis is timely as the Euclid satellite will be launched in the first year of the thesis (early 2023), allowing the PhD student to work on the first Euclid data until the first data release (DR1) expected in mid-2025. The tools developed in this thesis, together with the 2500 deg2 of DR1, will revolutionize both our understanding of galaxy populations within cosmic filaments, and the extraction of non-Gaussian cosmic shear information for cosmology.

Bibliography:

Pasquet, Bertin, Treyer, Arnouts, Fouchez 2019, A&A, 621, A26
Euclid collaboration: Martinet, Schrabback, Hoekstra, et al. 2019, A&A, 627, 59
Martinet, Harnois-Déraps, Jullo, Schneider 2021a, A&A, 646, 62