

- Baseline tests of a few Convolutional Neural Networks (CNNs) extracted from a kaggle competition on plankton images (working implementations available).
- Marginal improvements of the most promising CNN architecture.
- Guidance in the implementation of the best working solution in a web application for the classification of plankton images: <http://ecotaxa.obs-vlfr.fr> . This application is currently in production and used by about a dozen operators daily, from various teams around the world. The implementation will be subcontracted and the intern will just need to provide settings based on the tests performed above.

These steps should be relatively quick because most of the work is already done or prepared. They are essential, however, to pave the way for new developments. The potential avenues envisioned for this internship are:

- Exploration of mixed models including a CNN for feature extraction and classic machine learning for classification, hence allowing the inclusion of ancillary data (such as location, date, depth, etc.).
- Test of the generality of CNN-based models by training models across several instruments. Exploration of input images modification that can help generalise the problem across various sensors with different resolutions, dynamic ranges, etc.
- Exploration of the representation of images along the layers of the neural network, to determine which features are relevant for classification, with the aim of pre-processing the images and simplifying the network for its future implementation on embedded platforms
- Exploration of hierarchical classification (rather than classifying all classes simultaneously) to improve classification reliability and accuracy.

All these research questions would be of interest for the task at hand. It is unlikely that all can be tackled within the internship but the most interesting/promising will be identified through a discussion between the partners and the intern and then pursued.

The intern will be supervised by Marc Picheral (Ingénieur de Recherche) and Jean-Olivier Irisson (Maître de Conférence) at LOV and will collaborate with researchers at I3S. He/she will work in the framework of an existing convention between the two labs and a common project funded by CNRS (DL-PIC: Deep Learning for Plankton Images Classification). The DL-PIC project also involves international partners: ecologists at the Hatfield Marine Science Center of Oregon State University (OSU) and a deep learning expert at Facebook Artificial Intelligence Research lab (FAIR).

Practically, housing facilities may be available in the lab directly (depending on dates). The student will need to travel between Villefranche and Sophia-Antipolis to meet with all team members and a car would be useful (but not required). The student will be compensated for such travels in addition to the usual internship stipend. The student will be provided with a Linux computer but would be able to use his/her own. All actual work will be done on a remote Linux server equipped with a Tesla GPU entirely dedicated to the project.