
DEEP LEARNING FOR INVERTEBRATES DETECTION AND CLASSIFICATION

INTERSHIP OFFER



Supervision: [Hadrien Hendrikx](#), Inria (Thoth team), hadrien.hendrikx@inria.fr

In collaboration with François Postic and Mickael Hedde (INRAE, Eco&Sols team).

Context. More than 60% of European soils are unhealthy, with degradation caused by unsustainable land management, soil sealing, contamination, overexploitation, and climate change. Degraded soils reduce ecosystem services such as food production, carbon sequestration, and water regulation. Soil rehabilitation is essential to preserve ecosystem health.

In other ecosystem components, passive sensor approaches combined with computer vision and machine learning techniques have revolutionized our ability to monitor and analyze biodiversity and its role in ecosystem functioning rapidly and on a large scale. Digital recording of observational data increases the reproducibility of studies and allows for the exploitation of larger datasets while maintaining an objective view of the biological events being studied. These monitoring methods are implemented primarily in aquatic and terrestrial environments, but very rarely for the monitoring of soil organisms.

The Eco&Sols lab from INRAE Montpellier has developed a new methodology based on scanners to obtain a large corpus of high-quality soil images (4 images a day for each scanner). Counting and classifying invertebrates is a necessary step in better understanding soil biodiversity. Yet, manual processing is highly time-consuming and prevents large-scale studies over multiple sites using various scanners. The goal of this internship is to participate in the automation of an image processing pipeline for soil biodiversity monitoring.

Methods. A preliminary image processing pipeline has already been developed using manually annotated data. This internship will improve it by focusing on two main tasks:

- Using **unsupervised pretraining** methods to leverage the increasing amount of images that could not be annotated manually. This will allow to produce meaningful representations for image patches and improve detection and classification performances.
- Using state-of-the-art **object detection** models to directly detect and classify invertebrates on new images.

Skills developed by the candidate. By the end of the internship, the candidate will have developed a strong experience in building reliable pipelines for (self-supervised) training of state-of-the-art deep learning models. The intern will also develop a strong research experience, and the results of the internship might lead to a scientific publication. Given the increasingly important role of machine learning in biodiversity monitoring and beyond, this internship will give the candidate a relevant background for continuing with a PhD or becoming ML engineer.

Environment. The internship will take place at Inria Grenoble, in the [Thoth team](#) for at most 6 months (roughly March-August 2025, but flexible). This is a large team focused on machine learning, and in particular computer vision. Particular topics of interest include visual comprehension, hyperspectral imaging, numerical and parallel optimization, and unsupervised learning. All Thoth researchers have strong track records in top-tier ML conferences. Supervision will be ensured by a pluri-disciplinary team composed of experienced researchers in ecology and machine learning, and the internship will include visits to the Eco&Sols team, at INRAE Montpellier. The intern will also be in contact with ML researchers from the LECA lab.

This project is part of the CASCA project, funded by the ADEME (French agency for the ecological transition). This internship provides the unique opportunity to discuss with scientists from other fields and to improve their workflows

through AI research.

Requirements. We seek candidates strongly motivated by challenging research topics in machine learning for science. Applicants should have a strong coding and mathematical background with knowledge of machine learning. Proficiency in Python is expected and preliminary experience in Pytorch is a plus. The candidate should also be willing to engage in pluri-disciplinary collaborations and interested in contributing to answering research questions in ecology.