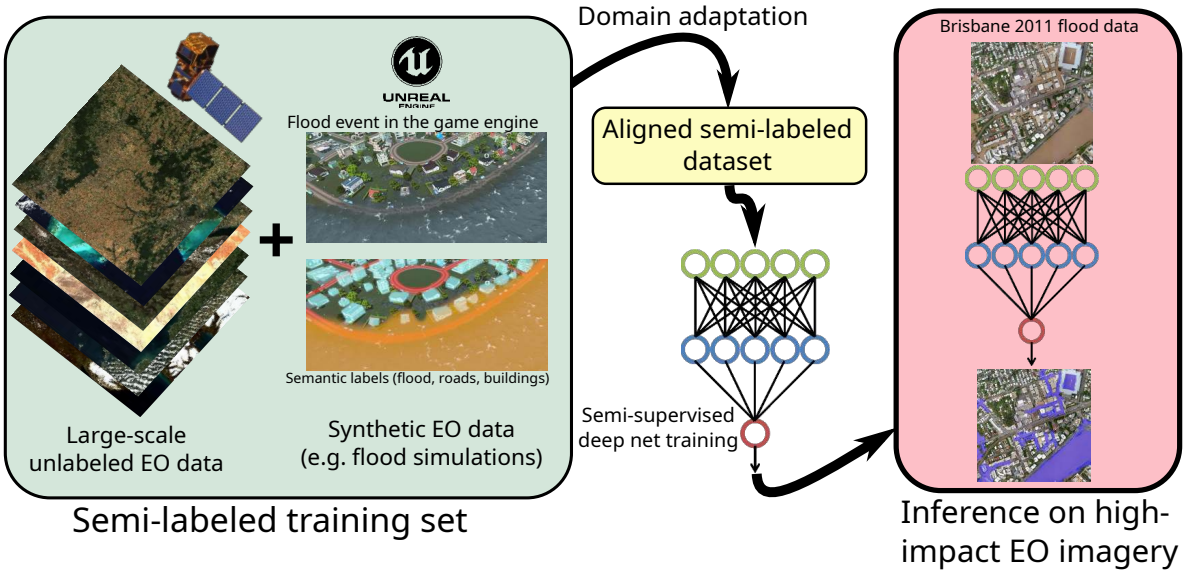
Postdoctoral position in semi-supervised learning/domain adaptation for aerial imagery

Nicolas Audebert: nicolas.audebert@cnam.fr

Application deadline: September 30, 2022

Location	Cnam (Paris, France)
Salary	≈2100€ to 2500€ net/month
Contract	Fixed-term (1 year, renewable)
Remote work	Partial
Start	February 1st 2023 at the latest

Keywords: generative models, domain adaptation, semi-supervised learning, unsupervised learning.



1 The research project

The MAGE project (*Mapping Aerial images using Game Engine data*) investigates the use of synthetic data to train very large-scale deep models for land cover and land use mapping from aerial and satellite imagery. Huge amounts of Earth Observation (EO) data are available thanks to european satellites Sentinel-2 and to French observations programs SPOT and BDORTHO. However, these big data are unlabeled and therefore contain no useful semantic information that could be used to train downstream machine learning models.

However, the Copernicus disaster management program, that delivers rapid mapping of sinistered areas, could strongly benefit from the faster response time that AI is able to bring. Interpretation of remote sensing data after a disaster (flood, earthquake...) is currently done by hand by experts. Since these events are exceptional, there are few labeled datasets that can be used to train supervised models. Meanwhile, the autonomous driving community has dealt with this problem by using video games to simulate events that are rarely observed in real life. Simulation of EO data is now possible thanks to softwares such as CityEngine that can procedurally generate entire cities.

This project aims to bring together both those tools and using simulations to train deep networks for rapid post-disaster mapping. We will leverage modern video game engines to simulate aerial views of cities before and after a disaster (flood, earthquake, fire). These images will constitute the labeled dataset that will complement the large amount of existing unlabeled data. We will improve the realism of the simulation using adaptation domain techniques based on generative models, and we will develop semi-supervised learning algorithms based on self-supervision for semantic segmentation. This will allow us to train deep models able to generalize for large-scale mapping of damaged structures, to identify the most affected areas and improve how emergency services navigate the city after a disaster.

2 Scientific objectives

The MAGE project deals with two scientific problems related to learning on synthetic data. The first objective aims to solve domain adaptation between synthetic and real datasets. As photorealistic rendering engines have become in the last years, there will still be a statistical gap between simulated images and acquisitions from actual sensors. To avoid degrading the model's performances, one solution is the use of generative models to improve the realism of synthetic images, therefore closing the domain gap. A second objective is to investigate effective training schemes from a mix of a small amount of labeled but synthetic images and a large quantity of real unlabeled data. This semi-supervised setting requires suitable learning algorithms that can leverage the unlabeled observations, for example using self-supervised pretraining schemes, label propagation algorithms ou hybrid approaches such as consistency learning.

3 Candidate profile

The ideal applicant holds a PhD in computer science, either with a specialization in machine learning ou computer vision. They demonstrate a strong experience in deep learning (project, publication, ...) and familiarity with the Python programming languages and its deep learning frameworks (PyTorch, TensorFlow, JAX...). Although not required, a first experience or a will to learn more about Earth Observation and remote sensing (aerial/satellite images) is a plus.

4 Where you will work

The *Center for research and studies in computer science and communications* (Cédric) is the computer science laboratory of the Conservatoire national des arts et métiers (Cnam), a prestigious French higher education institute. It is comprised of 80 faculty members and researchers, for a total of more than 180 people including postdoctoral fellows and PhD students. Its eight teams cover most areas in computer science, from data science to interactive media, discrete optimization, telecommunications and the Internet of Things. The new hire will join the *Complex Data, Machine Learning and Representations* team[1]. Their research will be performed inside the MAGE project under its principal investigator Dr. Nicolas Audebert.

Organization: the work contract is a fixed-term contract of 12 months ("CDD"). It is a full-time position of 35 hours/week. It can optionally be renewed once. The expected salary is 2100€ to 2500€ per month depending on the qualifications and the experience of the applicant. The starting date is expected at the earliest on the November 1st 2022 and at the latest on February 1st 2023, depending on the applicant.

Location: the laboratory is located in the heart of Paris, in the third "arrondissement", at 2 rue Conté (subway "Arts & Métiers", lines 3 and 11).

Hiring: the application process is done in two steps: first a short half-hour interview by phone or videocall, then a longer technical interview of about one hour.

To apply: send a resume to `nicolas.audebert@cnam.fr`

Benefits:

- 44 days of paid holidays
- on-site subsidized restaurant
- partial remote work is possible
- employees' association (music classes, on-site gym...)

[1] `https://cedric.cnam.fr/lab/equipes/vertigo/`