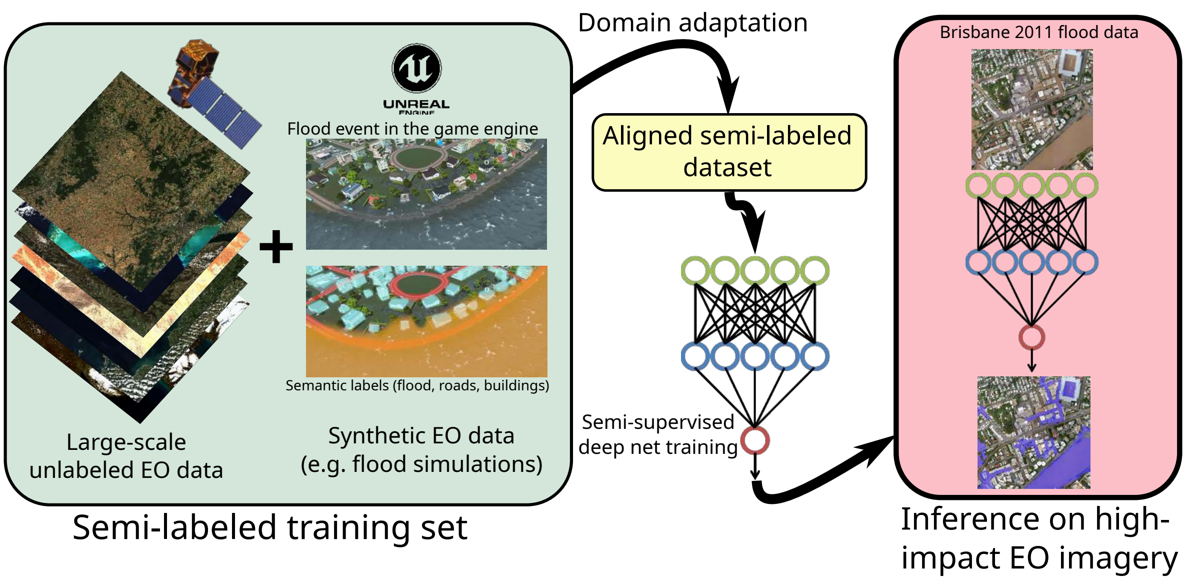
PhD position in semi-supervised learning/domain adaptation for aerial imagery

Nicolas Audebert: nicolas.audebert@cnam.fr

Application deadline: September 30, 2022

Location	Cnam, Paris, France
Salary	≈1800€ net/mois
Contract	Fixed-term (3 years)
Remote work	Partial
Starting date	January 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

This doctoral position aims to investigate semi-supervised training of deep models with a focus on two frequent use-cases in Earth Observation. More precisely, we assume that the dataset can be split in two subsets:

- a large amount of unlabeled aerial or satellite observations,
- a small number of labeled images ("gold data").

These labeled images come from either real acquisitions annotated by experts or from synthetic observations obtained by simulation. In both cases, the problem is the misalignment between the distribution of the labeled data and the target geographical area [3]. For example, how can the model map accurately flooded towns in India if the only available labels come from urban reas in Europe? A possibility is to leverage simulated images. Yet, this introduces a significant perceptual bias since simulations do not accurately model specificities of real acquisitions (atmospheric perturbations, sensor noise).

This PhD will investigate two research topics in deep learning, with a focus on mapping applications, *i.e.* semantic segmentation of remote sensing images for land cover or land use mapping. Other applications such as height estimation or change detection might also be considered depending on opportunities.

Semi-supervised learning for mapping

To accelerate the mapping capacities of statistical models over the whole Earth, it is necessary to leverage unlabeled data and therefore moving away from the supervised paradigm. Labeling remote sensing images require expert knowledge that makes it elusive that large scale datasets similar to Cityscapes or COCO can ever be available. Advances in deep learning have revealed that self-supervised pretraining is extremely effective for the fine-tuning of specialized models on small amount of data. For example, pretext tasks (jigsaws [4], contrastive learning on different augmentations of the same image [1]) allow deep models to learn useful representations of images for downstream tasks, without any human label. Designing semi or self-supervised algorithms suited to Earth Observation data would greatly facilitate transfer learning on smaller target datasets, therefore significantly improving the way we train deep networks for mapping. The plethora of small models specialized on constrained geogrpahical areas could be replaced by a large generic model that can be fine-tuned quicker and cheaper than training from scratch.

Domain adaptation for Earth Observation

Simulation of aerial and satellite images is an exciting prospect for geoscientists and geographs, especially for natural disaster management: floods, volanic events, earthquakes, storms...Indeed, these events are rare and observations are scarce, even without labels. Creating synthetic datasets is now common practice in computer vision and autonomous driving. Early works using GTA V [6] have identified that the main obstacle to using simulations to train Earth Observation models is that synthetic data does not exhibit some of the usual properties of remote sensing sensors (atmospheric perturbations, sensoir noise, illumination realism...). Consequently, there is a gap between the synthetic and real distributions. Reducing this gap is the goal of domain adaptation. We aim to study and design domain adaptation algorithms tailored to improving the realism of synthetic Earth Observation images, so that simulations can then be used to train deep models that will generalize on actual acquisitions. More specifically, methods based on generative models (*normalizing flows* [2], diffusion models [5]) constitute a promising avenue since they have been successfully applied for realism enhancement in image generation.

3 Candidate profile

The ideal applicant has the following qualifications:

- holds a **master degree or an engineering degree in computer science**, with a specialization in machine learning or computer vision,
- has some **experience with deep learning** (project, courses, publication...),
- is familiar with the **Python** programming language and at least one **deep learning framework** (PyTorch, TensorFlow, JAX, ...),
- oral and written communication skills. French is not required but can help with the everyday life of the PhD candidate.

Although not required, an experience or a will to learn about remote sensing, interpretation of aerial and satellite images and geographical applications (climate change, disaster management) is a plus. Since motivation is the main factor for a successful thesis, all applications, even those that do not entirely fit the described profile, will be considered.

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: PhD thesis in France are done under a fixed-term work contract of 3 years (36 months, "CDD"). It is a full-time position of 35 hours/week. The expected salary is about ≈1800€ net/month. The starting date is expected at the earliest on the November 1st 2023 and at the latest on January 1sc 2023, depending on the applicant.

Supervision: the PhD will be supervised by Nicolas Audebert (Cnam/Cédric) and performed under the direction of Pr. Nicolas Thome (Sorbonne Université/ISIR).

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...)

References

[1] Ting Chen et al. "A Simple Framework for Contrastive Learning of Visual Representations". June 30, 2020. arXiv: 2002.05709 [cs, stat]. URL: http://arxiv.org/abs/2002.05709.

[2] Ivan Kobyzev, Simon J.D. Prince, and Marcus A. Brubaker. "Normalizing Flows: An Introduction and Review of Current Methods". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.11 (Nov. 2021), pp. 3964–3979.

[3] Nicholus Mboga et al. "Domain Adaptation for Semantic Segmentation of Historical Panchromatic Orthomosaics in Central Africa". In: *ISPRS International Journal of Geo-Information* 10.8 (8 Aug. 2021), p. 523. URL: https://www.mdpi.com/2220-9964/10/8/523.

[4] Mehdi Noroozi and Paolo Favaro. "Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles". In: *Computer Vision – ECCV 2016*. Lecture Notes in Computer Science. 2016, pp. 69–84.

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

[5] Qinsheng Zhang and Yongxin Chen. "Diffusion Normalizing Flow". In: *Advances in Neural Information Processing Systems*. 2021. URL: http://arxiv.org/abs/2110.07579.

[6] Zhengxia Zou et al. "Do Game Data Generalize Well for Remote Sensing Image Segmentation?" In: *Remote Sensing* 12.2 (2 Jan. 2020), p. 275. URL: https://www.mdpi.com/2072-4292/12/2/275.