

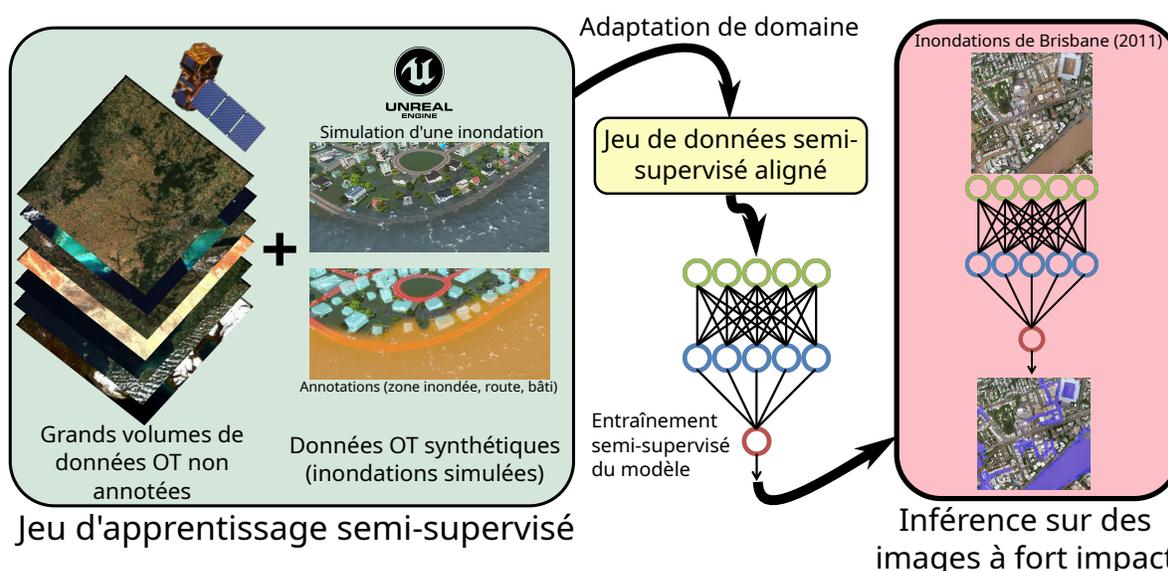
Offre de post-doctorat en apprentissage semi-supervisé/adaptation de domaine pour l'imagerie aérienne

Nicolas Audebert : nicolas.audebert@cnam.fr

Date limite de candidature : 30 septembre 2022

Localisation	Cnam, Paris 3 ^e (75)
Rémunération	≈2100€ à 2500€ net/mois
Contrat	CDD (1 an, renouvelable)
Télétravail	Partiel
Prise de fonction	Au plus tard 1 ^{er} février 2023
Date limite de candidature :	30 septembre 2022

Mots-clés: modèles génératifs, adaptation de domaine, apprentissage semi-supervisé, apprentissage non-supervisé.



1 Le projet de recherche

Le projet MAGE (*Mapping Aerial images using Game Engine data*) porte sur l'utilisation de données synthétiques pour l'entraînement à très grande échelle de modèles de cartographies d'images aériennes et satellitaires par apprentissage profond. D'importants volumes de données d'Observation de la Terre (OT) sont disponibles grâce aux satellites européens Sentinel-2 et aux programmes français d'imagerie SPOT et BDORTHO. Néanmoins, cette masse de données est non-étiquetée et ne contient pas d'information sémantique utile pour l'apprentissage automatique de modèles pour des applications concrètes.

La branche du programme Copernicus dédiée à la gestion des catastrophes réalise des tâches de cartographie rapide qui pourraient fortement bénéficier de temps de réponse grâce à l'IA. L'interprétation des données de télédétection après une catastrophe (inondation, séisme...) relève actuellement du travail des experts. Ces incidents étant des événements peu fréquents, les jeux de données d'apprentissage sont particulièrement rares. La communauté de la conduite autonome a pallié ce problème en utilisant les jeux vidéo pour simuler des situations difficiles à observer dans la réalité. Simuler des données d'OT est devenu possible grâce à des outils tels que CityEngine en créant des clones virtuels de villes entières.

Ce projet vise à allier ces outils et permettre à l'apprentissage profond d'exploiter les données de simulation pour la cartographie rapide post-catastrophe. Nous nous appuyons sur des moteurs de jeux pour simuler des vues aériennes de villes avant et après incident (inondation, séisme, incendie). Ces images serviront de données supervisées complémentaires à la masse de données non-étiquetée existante. Nous rendrons les images simulées plus réalistes en développant des approches d'adaptation de domaine basées sur les modèles génératifs et nous développerons

des méthodes d'apprentissage semi-supervisé sur le principe de l'auto-supervision pour la segmentation sémantique. Ceci nous permettra d'entraîner des modèles profonds capables de généraliser pour la cartographie des structures endommagées et l'identification des quartiers les plus touchés, facilitant la navigation des services d'urgence.

2 Les missions scientifiques

Le projet MAGE s'intéresse à deux verrous scientifique liés à l'apprentissage sur des données synthétiques. Le premier verrou cherche à résoudre le problème de l'adaptation de domaine entre les données synthétiques et les données réelles. Aussi réaliste le moteur de jeu soit-il, il existera toujours un écart statistique entre les images générées et les acquisitions par des capteurs réels. Afin d'éviter une dégradation des performances, une solution peut passer par l'usage de modèles génératifs pour augmenter le réalisme des images synthétiques. Un deuxième verrou est l'apprentissage à partir d'un mélange de données synthétiques étiquetées en faible nombre et d'un très grand volume de données réelles mais non annotées. Cette configuration semi-supervisée nécessite des schémas d'apprentissage adaptés afin d'exploiter les images non annotées, par exemple au travers de pré-entraînement auto-supervisé, de techniques de propagations d'annotations ou d'approches hybrides comme par exemple le *consistency learning*.

3 Le profil recherché

Le ou la candidate idéale est titulaire d'un doctorat en informatique, dans le domaine de l'apprentissage machine ou de la vision par ordinateur. Il ou elle démontre une expérience de l'apprentissage profond (projet, publication...), ainsi qu'une familiarité avec la programmation en Python et au moins une bibliothèque d'apprentissage profond (TensorFlow, PyTorch, JAX...). Sans être indispensable, une première expérience en analyse de données d'imagerie aérienne ou satellitaire est un plus.

4 L'environnement d'accueil

Le *Centre d'études et de recherche en informatique et communications* (Cédric) est le laboratoire d'informatique du Conservatoire national des arts et métiers (Cnam). Il comporte environ 90 enseignant-chercheurs et enseignantes-chercheuses permanentes, pour un effectif total d'environ 180 personnes. Ses huit équipes recouvrent la science des données, les médias interactifs, l'optimisation combinatoire, les télécommunications ou encore les objets connectés. La personne recrutée intégrera l'équipe *Données Complexes, Apprentissage et Représentations*¹. Elle effectuera ses recherches au sein du projet MAGE, porté par Dr. Nicolas Audebert.

Organisation: Le contrat de travail est un CDD de 12 mois (35h/semaine, temps plein), possiblement renouvelable une fois. La rémunération prévue est de 2100€ à 2500€ mensuels nets en fonction de l'expérience et du niveau de qualification de la personne retenue. La prise de fonction est prévue au plus tôt le 1^{er} novembre 2022 et au plus tard le 1^{er} février 2023, avec possibilités d'adaptation en fonction de la personne retenue.

Localisation: Le laboratoire se situe en plein cœur de Paris dans le 3^e arrondissement, au 2 rue Conté (métro Arts & Métiers, lignes 3 et 11, ou métro Réaumur-Sébastopol, lignes 3 et 4).

Procédure: le recrutement se fait en deux étapes : un premier entretien téléphonique d'environ une demi-heure, suivi d'un deuxième entretien technique d'environ une heure.

Candidature: envoyer un CV à nicolas.audebert@cnam.fr

Avantages:

- 44 jours de congés annuels
- cantine subventionnée
- télétravail partiel possible
- association des personnels (salle de sport, cours de musique...)

1. <https://cedric.cnam.fr/lab/equipes/vertigo/>