

## **Chaire Data Science and Artificial Intelligence for Digitalized Industry and Services**

### **Offre de stage**

#### **Sujet**

**Foundational models for adaptive semantic segmentation of satellite imagery.**

#### **Possibilité de poursuivre sur une thèse**

OUI

#### **A propos de la chaire**

La Chaire Artificial Intelligence for Industry (AI4I), portée par Stéphane Cléménçon, enseignant-chercheur dans le département Image, Données, Signal de Télécom Paris, réunit cinq partenaires industriels : Airbus Defence & Space, Engie, Idemia, Safran et Valeo. Son objectif général est de développer, en liaison étroite avec les partenaires, une formation et une recherche de niveau international.

Ses quatre principaux axes de recherche sont :

1. Analyse et prévision de séries temporelles (Predictive Analytics on Time Series) ;
2. Exploitation de données hétérogènes, massives et partiellement étiquetées (Exploiting Large Scale and Heterogeneous, Partially Labelled Data) ;
3. Apprentissage pour une prise de décision robuste et fiable (Learning for Trusted and Robust Decision) ;
4. Apprentissage dans un environnement dynamique (Learning through Interactions with a Changing Environment).

#### **Description du stage**

##### **Encadrement**

Stéphane Lathuillière, Florence Tupin, Denis Maraud, Benjamin Cepas

##### **Lieu et dates du stage**

Adresse : Télécom Paris, 19 Place Marguerite Perey, 91120 Palaiseau

Date de début du stage : mars - avril 2024

##### **Équipe(s) d'accueil de la thèse**

Département IDS

##### **Mots clés**

Semantic segmentation, foundation models, earth observation, remote sensing, optical imaging

## Sujet détaillé

Semantic segmentation is the process of assigning a class to each pixel in an image. In the context of Earth Observation, the objective can be to extract the accurate footprint of buildings, roads, low or high vegetation, water areas, but also objects like cars, trucks, ships, planes, ... The objective of the internship is to study the operational interest of large foundation models to support the adaptive semantic segmentation of high resolution satellite images (30 to 50cm) . The adaptiveness of the semantic segmentation will be studied for two settings:

- The interactive settings for which an image analyst is in the loop and may provide interactive inputs to orient the segmentation (visual prompting)
- The automatic setting for which we want to fine tune / adapt an existing semantic segmentation on a given domain to a new domain.

Both cases can be seen as domain adaptation problems: in the first case, the target domain is a single image that the image analyst intends to segment with as few inputs as possible. In the second case, the target domain is supposed to be represented by the distribution spanned by the available target images (the objective here being to request a as small as possible proportion of annotated target images among all available target images).

Foundational models have shown impressive generalization capabilities for a large number of tasks. One typical example of such models is Meta's Segment Anything model (SAM) (1). These models get their generalization capability in the vast amounts of data used to train them. Even if this data is usually taken from the internet (pictures with their captions for multimodal models), they still achieve decent performances on completely new domains such as Earth Observation (2), (3), (8). This performance is however not sufficient for operational usage. Besides, it is not fully understood how the specialization of these generic models is best performed for a given new task (here interactive semantic segmentation or new domain semantic segmentation of Earth Observation images): either through fine tuning (and then, which layers?), through visual prompting (like in SAM (1) or SLiME (4)), test time training, knowledge distillation in a new lighter model, ... In parallel, foundational models dedicated to Earth Observation imagery begin to appear (e.g. NASA/IBM geospatial model (5), RemoteCLIP (6) or RingMo (7)) and we may wonder whether it is really beneficial to start from these models rather than the previous ones.

The objectives of this internship are thus to:

- establish a state of the art of foundational models assessed in the literature on information extraction in satellite images (object detection, semantic segmentation)
- select the most promising approach to be used and evaluated for the internship
- identify publicly available datasets on which the adaptive semantic segmentation will be tested
- define the semantic segmentation pipelines for the two settings: interactive semantic segmentation and automatic domain adaptation.
- implement the approaches and run the evaluations
- propose potential evolutions to address the identified limitations.

## Profil du candidat

Etudiant titulaire d'un master 2 recherche ou école d'ingénieurs en mathématiques appliquées ou informatique avec spécialisation en science des données et traitement d'images.

## Candidatures

A envoyer à [florence.tupin@telecom-paris.fr](mailto:florence.tupin@telecom-paris.fr), [stephane.lathuilier@telecom-paris.fr](mailto:stephane.lathuilier@telecom-paris.fr);

- Curriculum Vitae
- Lettre de motivation personnalisée expliquant l'intérêt du candidat sur le sujet (directement dans le corps du mail)
- Relevés de notes des années précédentes
- Contact d'une personne de référence

Les candidatures incomplètes ne seront pas examinées.

## Références

- (1): SAM: Segment Anything - <https://segment-anything.com/>
- (2) The Segment Anything Model (SAM) for Remote Sensing Applications: From Zero to One Shot: <https://arxiv.org/abs/2306.16623>
- (3) Have Foundational Models seen Satellite Images? <https://hal.science/hal-04112634/document>
- (4): SLIME: Segment Like Me - <https://arxiv.org/abs/2309.03179>
- (5): NASA/IBM Geospatial model - <https://huggingface.co/ibm-nasa-geospatial>
- (6) RemoteCLIP: A Vision Language Foundation Model for Remote Sensing - <https://arxiv.org/abs/2306.11029>
- (7) RingMo: A Remote Sensing Foundation Model with Masked Image Modeling - <https://ieeexplore.ieee.org/abstract/document/9844015>
- (8) SAMRS: Scaling-up Remote Sensing Segmentation Dataset with Segment Anything Model: <https://arxiv.org/abs/2305.02034>