Offre de stage – Master 2 / École d'ingénieur

Intelligence artificielle et annotation sémantique d'images satellitaires : développement d'un système RAG pour l'exploration des données spatiales

Structure d'accueil : UMR Espace-Dev
Environnement scientifique : GDR PIC
Durée : 5 à 6 mois
Niveau : Master 2 ou dernière année d'école d'ingénieur
Localisation : Maison de la télédétection
Date de début : Février ou Mars

Contexte

Le développement de l'observation de la Terre génère des volumes croissants d'images satellitaires à très haute résolution spatiale, notamment issues des satellites SPOT 6/7, Pléiades et Pléiades Neo.

Ces images constituent une source importante d'informations pour l'étude des territoires, des paysages et des dynamiques environnementales. Leur exploitation nécessite cependant de développer des méthodes permettant d'identifier, de qualifier et de structurer automatiquement les informations qu'elles contiennent.

Les approches récentes d'intelligence artificielle, notamment les modèles de vision, les grands modèles de langage (LLM) et les systèmes RAG (*Retrieval-Augmented Generation*), offrent de nouvelles perspectives pour automatiser l'annotation des images et faciliter leur exploration par des requêtes en langage naturel.

Objectifs du stage

L'objectif du stage est de développer une approche d'annotation sémantique et descriptive d'images satellitaires, combinant vision par ordinateur et intelligence artificielle générative, afin de construire une base de connaissances géospatiales interrogeable par un système RAG.

Les informations étudiées porteront notamment sur :

- **Occupation du sol :** espaces urbains, agricoles, forestiers, surfaces en eau et sols nus.
- **Structures paysagères :** organisation spatiale, fragmentation, continuités écologiques et interfaces entre milieux.
- **Objets géographiques :** bâtiments, routes, infrastructures et parcelles.

- **Caractéristiques environnementales :** végétation, artificialisation et évolution des surfaces.
- **Dynamiques territoriales :** changements observables entre plusieurs acquisitions satellitaires.

Missions principales

Le ou la stagiaire sera amené(e) à :

1. Constituer et préparer un corpus d'images satellitaires SPOT 6/7, Pléiades et Pléiades Neo et leurs métadonnées.
2. Étudier et mettre en œuvre des méthodes d'annotation automatique ou semi-automatique : classification, segmentation, détection d'objets et description textuelle des images.
3. Développer une chaîne d'enrichissement sémantique associant informations visuelles, géographiques et descriptives.
4. Concevoir un prototype de système RAG permettant d'interroger les images et leurs annotations en langage naturel.
5. Évaluer la qualité des annotations et la pertinence des résultats produits.

Profil recherché

Étudiant(e) en Master 2 ou dernière année d'école d'ingénieur en informatique, intelligence artificielle, science des données, géomatique ou télédétection.

Compétences souhaitées :

- Programmation Python.
- Notions de traitement d'images et de machine learning.
- Intérêt pour les LLM, les modèles multimodaux et les architectures RAG.
- Connaissances en télédétection et données géospatiales appréciées.
- Autonomie, curiosité scientifique et goût pour l'expérimentation.

Résultats attendus

- Une méthodologie d'annotation sémantique d'images satellitaires.
- Un corpus d'images enrichies par des annotations structurées et des descriptions textuelles.
- Un prototype de système RAG appliqué aux données d'imagerie spatiale.
- Une démonstration permettant de rechercher et d'interroger les images à partir de critères géographiques, thématiques et sémantiques.

Encadrement scientifique

Le stage se déroulera au sein de l'UMR Espace-Dev, dans un environnement scientifique pluridisciplinaire, en lien avec les réflexions et activités du GDR PIC autour de l'information géographique, de la structuration des connaissances et des nouvelles approches d'intelligence artificielle appliquées aux données spatiales.

Equipe d’encadrement Stéphane Debard et Nicolas Barthes.

Contacts : stephane.debard@ird.fr, nicolas.barthes@ird.fr

Envoyez CV et lettre de motivations

Gratification : conformément à la réglementation en vigueur relative à l'accueil des stagiaires dans la fonction publique, une gratification sera versée lorsque la durée du stage y ouvre droit, selon les montants et conditions légales applicables.

Mots-clés : Télédétection, SPOT 6/7, Pléiades, Pléiades Neo, annotation sémantique, vision par ordinateur, LLM, RAG, intelligence artificielle, données géospatiales.