



Adaptation de modèle de fondation pour la détection d'objets sur images satellitaires

Adaptation of foundation models for object detection on remote sensing images

Soutenance de thèse de Clément BARBIER

2 décembre 2025 à 14h

Auditorium du Lumen - 8 Av. des Sciences, 91190 Gif-sur-Yvette

Devant le jury composé de :

Mme Javiera CASTILLO-NAVARRO
M. Vincent GRIPON
M. Sébastien LEFÈVRE
M. Clément MALLET
Mme Anissa MOKRAOUI

CNAM
IMT Atlantique
Université Bretagne Sud
IGN
Université Sorbonne Paris Nord

Examinatrice
Rapporteur
Examineur
Rapporteur
Examinatrice

M. Stéphane HERBIN
M. Baptiste ABELOOS

ONERA
ONERA

Directeur de thèse
Encadrant

Résumé

La télédétection connaît une révolution quantitative avec plus de 1200 satellites d'observation terrestre générant quotidiennement des pétaoctets d'images. Face à ce déluge de données, l'exploitation manuelle devient impossible et l'apprentissage profond s'impose comme solution incontournable. Cependant, le paradigme supervisé traditionnel, dépendant d'annotations massives et coûteuses, ne peut répondre aux besoins opérationnels de flexibilité et de réactivité. Cette thèse explore une voie alternative : l'adaptation de modèles vision-langage pré-entraînés sur données généralistes pour la détection d'objets en télédétection optique. Ces modèles de fondation, alignant représentations visuelles et textuelles dans un espace latent commun, permettent théoriquement une détection sans apprentissage préalable par simple requête textuelle. Le défi central réside dans le fossé sémantique et visuel séparant les images naturelles du web des vues aériennes satellitaires, caractérisées par leurs perspectives zénithales, leurs objets de petite taille et leur vocabulaire spécialisé.

Trois axes méthodologiques complémentaires sont investigués pour surmonter ce fossé sans recourir aux annotations massives. Premièrement, une méthode d'optimisation automatique des requêtes textuelles par descente de gradient est développée, apprenant des formulations sémantiquement alignées avec les spécificités visuelles de la télédétection. Deuxièmement, une analyse systématique des stratégies d'adaptation parcimonieuse identifie les composantes critiques du détecteur à ajuster pour maximiser les performances tout en minimisant le coût computationnel, particulièrement en régime de rareté des données. Troisièmement, une approche novatrice sans phase d'entraînement est proposée, exploitant des prototypes purement visuels pour conditionner le modèle et contourner les limitations du conditionnement textuel.

Ces contributions établissent les fondements méthodologiques pour le déploiement efficace des modèles de fondation en télédétection, ouvrant des perspectives pour l'analyse réactive et adaptative des images satellitaires dans des domaines critiques : surveillance environnementale, gestion des catastrophes, agriculture de précision et sécurité territoriale.

Mots clés : Détection d'objets, Apprentissage frugal, Télédétection, Intelligence Artificielle

Vous êtes invité à rejoindre la web-conférence via le lien ci-dessous :

<https://universite-paris-saclay-fr.zoom.us/j/95067077794?pwd=sc2Tla4oJ6DvD7zJbotzicabsTDh5c.1>

ID de réunion: 950 6707 7794, Code secret: 923334