

Sujet : Recherche efficace et sans apprentissage de motifs spatio-temporels dans des grands cubes de données satellites

Discipline : Informatique, Traitement du Signal et des Images

Mots-clefs : Traitement d'image, télédétection, séries temporelles, hiérarchies morphologiques, histogramme de forme (Pattern Spectra), algorithmes efficaces

Contact :

IRISA – site de l'UBS - Equipe OBELIX

François Merciol

francois.merciol@irisa.fr

IRISA, Bâtiment ENSIBS, Campus de Tohannic, 56 000 Vannes

Environnement du stage

L'IRISA (Institut de Recherche en Informatique et Systèmes Aléatoires, www.irisa.fr), est une unité mixte de recherche (UMR 6074) en informatique, en traitement du signal et des images et en robotique.

Sur la conjonction de ces thématiques, l'IRISA se positionne comme le grand laboratoire de recherche en Bretagne avec une présence affirmée sur les campus de Rennes (35), Vannes (56), et Lannion (22).

Elle s'organise autour de 7 départements, dont celui « Signal, Image, Langage » dans lequel s'inscrivent les travaux de l'équipe OBELIX.

L'équipe OBELIX (Observation de l'environnement par imagerie complexe) mène ses recherches dans le domaine de l'intelligence artificielle (apprentissage automatique et vision par ordinateur) appliquée à l'observation de la Terre. Elle conçoit ainsi des solutions informatiques originales, efficaces et robustes pour faire face à la complexité inhérente aux données environnementales. En particulier, l'équipe a développé ces dernières années une expertise méthodologique dans les domaines de l'apprentissage profond, du transport optimal, et des hiérarchies morphologiques.

Cette équipe est principalement localisée à Vannes (56), et dispose de moyens de calcul (clusters CPU et GPU) qui pourront être mobilisés pendant le stage.

L'encadrement sera assuré par François Merciol (IRISA – UBS).

Contexte applicatif

Depuis 2011, le Copernicus Land Monitoring Service (CLMS land.copernicus.eu) fournit des produits pour la surveillance de l'état, des changements et des caractéristiques de la couverture/utilisation des terres végétalisées, non végétalisées, des variables biophysiques, des conditions de l'eau et de la cryosphère. Cette cartographie à large échelle utilisent les données d'observation à haute résolution de la terre (10 m de résolution spatiale, une acquisition toutes les semaines) gratuites et libres d'accès telles que Sentinel (1 et 2), ainsi que des données commerciales à très haute résolution spatiale (1 m de résolution spatiale, tous les ans).

Dans ce contexte l'UMR LETG (Unité mixte de recherche : Littoral – Environnement – Télédétection – Géomatique) explore les interactions nature/société au travers de la géographie. Nous souhaitons renforcer la compréhension et l'automatisation des liens entre outils d'analyse

numérique de série temporelle d'images satellite représente et la description ontologique de scène naturelle. Un des objectifs de ce stage est poursuivre une collaboration entre les deux laboratoires au travers d'une thèse CIFRE.

Contexte scientifique

Dans le cadre d'une demande de l'Agence Européenne de l'Environnement, OBELIX ont conçu et déployé une chaîne de production originale de la cartographie des trames vertes à l'échelle continentale pour le compte du programme Copernicus. Pour faire face au volume de données à traiter (38 000 images, soit 120 To), et à la diversité des scènes étudiées, la solution développée s'est appuyée sur des algorithmes efficaces de caractérisation multi-échelle des pixels (profils d'attributs) à l'aide des hiérarchies morphologiques, et de classification semi-supervisée par une approche ensembliste de forêts aléatoires [1]. Elle a été implantée à l'aide de composants logiciels C++ diffusés sous licence libre: TRISKELE [2] et Broceliande [3]. Une attention particulière est portée à l'optimisation systématique de toutes les étapes du processus, y compris l'extraction des descripteurs [4]. Cette étape, centrale dans le processus de cartographie automatique, est souvent mise en œuvre à l'aide des profils d'attributs calculés efficacement à l'aide des hiérarchies morphologiques [5].

Dans un autre contexte, l'équipe OBELIX a collaboré avec le CNES dans le cadre d'une étude R&T pour développer une solution efficace de recherche automatique par l'exemple dans des bases d'images satellites. Pour cela, elle a exploité les hiérarchies morphologiques pour calculer des histogrammes de formes (ou Pattern Spectra) qui permettent de mettre en œuvre des algorithmes efficaces de recherche par l'exemple. La solution ainsi développée rend possible la recherche de motifs spatiaux de taille variable dans une base de très grandes images [6]. Contrairement aux approches populaires en vision par ordinateur, fondées sur l'apprentissage automatique ou profond, elle ne s'appuie pas sur un entraînement préalable d'un modèle prédictif, et fonctionne sans recourir à des données annotées. Ce travail a abouti au démonstrateur Korrigan [7].

Au vu de la pertinence des hiérarchies morphologiques et des outils qui en découlent (profils d'attributs, histogrammes de formes) pour élaborer des solutions efficaces d'analyse semi-automatique d'images satellites, leur extension aux séries temporelles d'images satellites a également été étudiée, au travers d'une thèse de doctorat conduite conjointement par l'équipe OBELIX, le CNES, et CLS [8].

Enfin nous venons de présenter une première évaluation de ce travail en novembre 2025 dans le cadre de DGMM à Groningen en comparant ce résultat avec un celui d'un auto-encodeur [9].

Sujet

En s'appuyant sur les développements précédents de l'équipe (en particulier les histogramme de forme) appliquées à l'analyse d'images et de séries temporelles d'images satellites, l'objectif du stage sera de définir une solution originale et efficace pour la recherche de motifs spatio-temporelles. Nous souhaitons explorer une démarche fondée sur des algorithmes efficaces afin de mesurer inexplicabilité des résultats (fonctionnement en boîte blanche) et leur sobriété (faible empreinte carbone du fait de l'optimisation des ressources en calcul et mémoire).

Plus précisément, alors que nos travaux initiaux conduits avec le CNES ont donné lieu à un démonstrateur où la recherche portait sur des motifs spatiaux, il s'agira ici d'ajouter une contrainte temporelle, ce qui permettra la recherche d'objets et/ou d'évènements représentés par des motifs

spatio-temporels. Bien que des travaux préliminaires aient été menés pour fournir une définition de l'histogramme de forme pour des données spatio-temporelles.

Au cours du stage, le travail consistera à :

- Etude bibliographique des Pattern Spectra et de leurs applications
- Etude bibliographique de la recherche de motifs dans les séries temporelles d'images satellites
- Prise en main de l'environnement logiciel
- Proposition d'une méthodologie
- Prototypage de la solution
- Validation sur des jeux de données

Profile recherché

Nous recherchons des candidats, en voie d'achèvement de leur cursus de Master ou école d'ingénieurs, qui affichent d'excellents résultats académiques, et les compétences suivantes :

- Algorithmique avancée et programmation (Python et/ou C++)
- Traitement d'image
- Notions ou intérêt pour la télédétection
- Bonnes compétences verbales et écrites en anglais scientifique
- Curiosité, rigueur et esprit d'équipe

Débouchés

En fonction du niveau et de la motivation du ou de la stagiaire, une poursuite en thèse CIFRE avec la société Alkante est envisagée. La thèse élargira les approches méthodologiques envisagées à l'apprentissage profond, dans le but de comparer les avantages et inconvénients des deux paradigmes, et d'élaborer une stratégie hybride le cas échéant.

Gratification

Le stage est rémunéré conformément à la législation.

Référence

[1]

Geobia at the terapixel scale: Toward efficient mapping of small woody features from heterogeneous vhr scenes, François Merciol, Loïc Fauqueur, Bharath Bhushan Damodaran, et al, ISPRS International Journal of Geo-Information, 2019

[2]

Tree Representations of Images for Scalable Knowledge Extraction and Learning for Earth observation, <https://gitlab.inria.fr/obelix/triskele/>

[3]

Broceliande is a tools for classification base on TRISKELE and Random Forest, <https://gitlab.inria.fr/obelix/broceliande/>

BROCELIANDE: a comparative study of attribute profiles and featureprofiles from different attributes, François Merciol, Minh-Tan Pham, Deise Santana Maia et al, ISPRS, 2020

[5]

Classification of remote sensing data with morphological attributes profiles: a decade of advances, Deise Santana Maia, Minh-Tan Pham, Erchan Aptoula et al, IEEE geoscience and remote sensing magazine, 2021

[6]

Fast Pattern Spectra using Tree Representation of the Image for Patch Retrieval, Behzad Mirmahboub, Jérôme Moré, David Youssefi et al, Discrete Geometry and Mathematical Morphology 2021

[7]

Korrigan is a tools base on TRISKELE for seeking patch in big remote sensing images database, <https://gitlab.inria.fr/obelix/broceliande/>

[8]

Thèse : Morphological Hierarchies for Satellite Image Time Series, Caglayan Tuna, Université Bretagne Sud, 2020

[9]

Efficient Content-Based Time Series Retrieval using Pattern Spectra, François Merciol, Tom Avellaneda, Abdelbadie Belmouhcine, Sébastien Lefèvre, DGMM, Groningen, 2025