



**HAL**  
open science

# Adaptation de domaine pour la cartographie de l'occupation des sols à partir de séries temporelles d'images satellitaires optiques

Emmanuel Capliez, Dino Ienco, Raffaele Gaetano, Nicolas Baghdadi, Adrien Hadj Salah

► **To cite this version:**

Emmanuel Capliez, Dino Ienco, Raffaele Gaetano, Nicolas Baghdadi, Adrien Hadj Salah. Adaptation de domaine pour la cartographie de l'occupation des sols à partir de séries temporelles d'images satellitaires optiques. ORASIS 2021, Centre National de la Recherche Scientifique [CNRS], Sep 2021, Saint Ferréol, France. hal-03339618

**HAL Id: hal-03339618**

**<https://hal.science/hal-03339618v1>**

Submitted on 9 Sep 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Adaptation de domaine pour la cartographie de l'occupation des sols à partir de séries temporelles d'images satellitaires optiques

Emmanuel Capliez<sup>1</sup> Dino Ienco<sup>1</sup> Raffaele Gaetano<sup>2</sup> Nicolas Baghdadi<sup>1</sup> Adrien Hadj Salah<sup>3</sup>

<sup>1</sup> INRAE, UMR TETIS, Univ. Montpellier, France

<sup>2</sup> CIRAD, UMR TETIS, Univ. Montpellier, France

<sup>3</sup> Airbus Defence and Space, France

emmanuel.capliez@inrae.fr

## Résumé

*Source de données à fortes valeurs intrinsèques, les séries temporelles d'images satellitaires sont devenues indispensables dans l'élaboration de produits tels que les cartographies d'occupation des sols. Leur production automatique par apprentissage machine est possible sous réserve de disposer des données de référence nécessaires. Le facteur limitant concerne bien souvent la quantité et la qualité de ces données. Étudiée depuis une dizaine d'années, l'adaptation de domaine non supervisé est un champ de recherche prometteur mais peu décliné et efficace dans le cas de cette problématique. Aussi, nous proposons de rechercher une nouvelle méthode d'adaptation de domaine adaptée aux données d'observation de la Terre et tirant profit de la dimension temps des séries temporelles d'images satellitaires.*

## Mots Clés

Adaptation de domaine, Série temporelle d'images satellitaires, Occupation des sols, Apprentissage profond, Apprentissage machine, Télédétection optique, Sentinel-2.

## Abstract

*As a source of data with high intrinsic value, satellite image time series have become essential in the development of products such as land cover mappings. Their automatic production by machine learning is possible if the necessary reference data are available. The limiting factor is often the quantity and quality of this data. Unsupervised domain adaptation, which has been studied for the last ten years, is a promising field of research, but it's not well developed and efficient in the case of this issue. Therefore, we propose to search for a new domain adaptation method adapted to Earth observation data and taking advantage of the time dimension of satellite image time series.*

## Keywords

Domain adaptation, Satellite Image Time Series, Land cover, Deep learning, Machine learning, Optical remote sensing, Sentinel-2.

## 1 Contexte

L'imagerie satellitaire est devenue une source incontournable de données à fortes valeurs intrinsèques. En effet, ses atouts sont nombreux : grands volumes de données multimodales acquis en un passage au dessus d'une zone géographique large, absence de contrainte de survol, gratuité parfois (ex : programme Sentinel de l'Union Européenne) et surtout une forte revisite (5 à 10 jours pour Sentinel-2<sup>1</sup>). Ce dernier avantage offre la possibilité d'obtenir des séries temporelles d'images satellitaires (STIS), ce qui autorise le suivi dynamique et l'évolution temporelle des observables. Les STIS sont grandement utilisées dans de nombreux secteurs d'activité importants comme l'agriculture [2], l'écologie<sup>2</sup>, la mobilité, la santé, l'évaluation des risques, l'aménagement du territoire, le suivi des forêts et de l'habitat naturel.

Cependant, à l'instar de nombreuses applications nécessitant de grands volumes de données, les utilisateurs de STIS sont confrontés à un défi important, celui de l'extraction efficace de la connaissance et l'exploitation de la dimension spatio-temporelle pour en tirer le meilleur parti. L'efficacité recherchée ici est relative à l'augmentation de l'automatisme et à la diminution des coûts de l'extraction.

Nous nous intéressons à un produit, considéré depuis longtemps comme une donnée d'entrée d'importance vitale pour de multiples activités de recherche et de management, et dérivable à partir de séries temporelles d'images satellitaires : l'occupation des sols [3].

## 2 Problématique

Une carte d'occupation des sols peut être produite automatiquement grâce à des algorithmes d'apprentissage machine supervisé (apprentissage profond et forêts d'arbres décisionnels). Les recherches récentes ont montré que l'utilisation de STIS à la place d'un image mono-date permettaient de réaliser des cartes plus précises, ainsi que de clas-

1. <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>

2. Références pour cette activité et les suivantes non citées pour raison de taille globale du document

sifier des types d'occupation qui ne peuvent l'être avec cette seule image. Par exemple, le soja et le maïs sont des cultures hivernales qui apparaissent semblables sur une seule image. L'utilisation d'une STIS met en évidence la différence de leur taux de croissance, ce qui permet de les distinguer.

Malheureusement, les très bonnes performances de ces méthodes de classification reposent sur la disponibilité de données de référence, sur la zone d'étude, en quantité et qualité suffisante. Cependant, c'est loin d'être toujours le cas ; les raisons en sont multiples : financières, disponibilité de l'expertise humaine pour la labellisation et accès aux instruments de mesure. Il y a parfois quelques données de référence, mais le plus souvent aucune donnée de vérité terrain n'est disponible.

Face à cette problématique et au besoin récurrent de constituer et mettre à jour une carte d'occupation des sols à moindre effort (temps, ressources humaine et financière), un des principes mis en oeuvre consiste en la réutilisation de connaissances acquises [7]. On parle ainsi de transfert de connaissances, de modèles.

Partant de ce principe, on se rend compte que très rares sont les cas où l'on peut réutiliser directement le modèle dans un contexte d'analyse de données de télédétection. En effet, on est vite confronté à des challenges spécifiques liés aux différences de conditions météorologiques et climatiques et d'occupation du sol lors de l'acquisition, lorsque l'on cherche à appliquer efficacement un modèle construit sur un site donné ou sur une année donnée, à une autre zone géographique ou sur le même site d'apprentissage mais à une autre date.

### 3 Méthodologie de recherche

Cette problématique générale fait partie d'un champ de recherche particulier de l'apprentissage automatique : l'adaptation de domaine [1]. Rattaché à l'apprentissage par transfert, cela consiste à transférer de la connaissance entre deux domaines : le domaine source et le domaine cible. Ce qui revient à faire apprendre un modèle à partir des données labellisées du domaine source et à l'utiliser sur des données le plus souvent non labellisées du domaine cible. L'hypothèse est qu'il existe un lien entre la distribution de données du domaine source et celle du domaine cible. Cela permet, par exemple, de palier un manque de données labellisées dans le domaine cible ou un décalage entre les deux distributions. Cette technique se décline aussi pour les algorithmes d'apprentissage profond [7]. L'étude menée se restreindra à l'adaptation de domaine non supervisée i.e. aucune donnée labellisée n'est disponible pour le domaine cible.

Bien qu'il existe une quantité conséquente de travaux menés sur l'adaptation de domaine non supervisé, peu traitent de séries temporelles, et encore moins acquises par télédétection [6]. De plus, une étude très récente montre que les méthodes de l'état de l'art d'adaptation de domaine non supervisé sont inefficaces avec les STIS [5]. Ce travail de

recherche sera un challenge particulièrement intéressant à relever.

Ainsi, tout en gardant à l'esprit la problématique opérationnelle de construction et mise à jour d'une carte d'occupation des sols, notre idée de recherche aura pour objet de proposer une nouvelle méthode de classification s'appuyant sur l'hybridation de méthodes à l'état de l'art en matière d'adaptation de domaine (priorité) et de classification de STIS.

A ce titre, notre méthodologie de recherche s'articulera en cinq phases : i) évaluation de différents classificateurs en mode supervisé permettant d'obtenir une référence basse de performance : TempCNN\*, CNN1D [4], Hob2sRNN\*..., ii) évaluation de différents compétiteurs pour l'adaptation de domaine : DANN\*, ADDA\*, GFK\*, iii) proposition d'une nouvelle méthode d'adaptation de domaine s'appuyant sur l'état de l'art et visant de meilleures performances, iv) validation de celle-ci sur un jeu de données (Balaruc les Bains) et v) expérimentation avec un autre jeu de données sur l'île de La Réunion. Cette nouvelle méthode devra exploiter au mieux la caractéristique temporelle d'une STIS. \*Référence non citée pour gain de place mais aisément trouvable sur Internet.

### Références

- [1] Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., Vaughan, J. W. (2010). A theory of learning from different domains. *Mach. Learn.*, 79(1-2), 151-175.
- [2] Bellón, B., Bégué, A., Seen, D. L., de Almeida, C., Simões, M. (2017). A remote sensing approach for regional-scale mapping of agricultural land-use systems based on NDVI time series. *Remote Sens.*, 9(6), 600.
- [3] Ienco, D., Interdonato, R., Gaetano, R., Ho Tong Minh, D. (2019). Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS J. of Photo. and Rem. Sens.*, 158, 1183-1192.
- [4] Ienco, D., Gbodjo, Y. J. E., Gaetano, R., Interdonato, R. (2020). Weakly Supervised Learning for Land Cover Mapping of Satellite Image Time Series via Attention-Based CNN. *IEEE Access*, 8, 179547-179560.
- [5] Lucas, B., Pelletier, C., Schmidt, D., Webb, G. I., Petitjean, F. (2020). Unsupervised Domain Adaptation Techniques for Classification of Satellite Image Time Series. *IGARSS*, 1074-1077.
- [6] Tardy, B., Inglada, J., Michel, J. (2017). Fusion Approaches for Land Cover Map Production Using High Resolution Image Time Series without Reference Data of the Corresponding Period. *Remote Sensing*, 9(11), 1151. doi :10.3390/rs9111151
- [7] Wang, M., Deng, W. (2018). Deep visual domain adaptation : A survey. *Neurocomputing*, 312, 135-153.