

Invitation à la soutenance de thèse

**SEMANTIC SEGMENTATION OF AIRBORNE HYPERSPECTRAL IMAGES (0.4 – 2.5 μm) FOR
MAPPING IMPREMEABLE SURFACES IN LARGE URBAN AREAS**

**SEGMENTATION SÉMANTIQUE D'IMAGES HYPERSPECTRALES AÉROPORTÉES (0,4 – 2,5 μm)
POUR LA CARTOGRAPHIE DE SURFACES IMPRÉMÉABLES DANS DE GRANDES ZONES
URBAINES**

Romain Thoreau

Le 20 novembre 2023 à 14h
Salle des thèses de l'ISAE
10 Av. Édouard Belin, 31400 Toulouse

Devant le jury composé de :

Mme Céline HUDELOT	CentraleSupélec	Rapporteure
M. Karteek ALAHARI	INRIA	Examinateur
Mme Yuliya TARABALKA	INRIA / LuxCarta	Examinatrice
M. Patrick GALLINARI	Sorbonne Université / Criteo	Examinateur
M. Laurent RISSER	CNRS - IMT	Co-directeur de thèse
M. Xavier BRIOTTET	ONERA	Directeur de thèse
Mme Véronique ACHARD	ONERA	Membre invité
Mme Béatrice BERTHELOT	Magellium	Membre invité

Résumé

In Earth observation, one of the main challenges is to map soil artificialization, which has been identified as a major public policy issue in the 2018 Biodiversity plan of the French Ministry of Ecological Transition. The sealing of natural surfaces by artificial impermeable surfaces has indeed major consequences on watershed hydrology, in particular on floods and droughts, on urban heat island effects and on soil carbon sequestration. If automatic mapping techniques have allowed to produce large scale land cover maps at the object level (e.g. forests, built-up areas, vineyards...) from satellite images, current methods are struggling to produce land cover maps of artificial impermeable surfaces (e.g. concrete, asphalt, paving stones...) at high spatial resolution over large urban areas. Airborne hyperspectral imaging has great potential for discriminating the land cover, thanks to its high spatial resolution and high spectral resolution over a wide spectral range. State-of-the-art mapping techniques optimize semantic segmentation models (i.e. models that segment the image into zones to which classes are assigned) from a subset of labeled pixels, called the training data set. Because labeling pixels require expensive field campaigns and photo-interpretation, the training data set is usually very small with respect to the diversity of materials in metropolitan areas and to the large spectral intra-class variability, that can be divided into three categories: *physics* variability induced by different illumination

conditions, *intrinsic* variability induced by slight variations in the material chemical composition and *semantic* variability resulting from the fact that different materials are actually gathered within the same land cover classes. In this context, the three main contributions of this thesis consist of leveraging the information of the millions of unlabeled pixels in the hyperspectral image in order to learn discriminating data representations that are robust to intra-class variability (and incidentally to inter-class similarities). Our first contribution is to study the potential of Active Learning (AL) algorithms to select a few additional pixels to annotate in order to improve the quality of the training data set. Our numerical experiments show that different AL strategies are complementary and that we can effectively integrate a priori semantic information through the class hierarchy of impermeable surfaces, resulting in significant improvement of the accuracy of segmentation models for an additional few hundred labeled pixels. Nevertheless, the benefits of AL are limited by the capacity of segmentation models (used by AL methods themselves) to learn discriminating representations. Thus, our second contribution is to introduce a generative model that integrates a priori knowledge derived from physical laws in order to learn, from labeled and unlabeled data, spectral representations that are robust to *physics* intra-class variations. Our experiments show that our hybrid model has better extrapolation capabilities compared to conventional machine learning models for materials with illumination conditions that were not represented in the training data set. Finally, our third contribution is to investigate the potential of unsupervised and self-supervised methods for learning spectral representations that are robust to *intrinsic* and *semantic* variations. To this end, we have built and released a very large hyperspectral database particularly suited to the evaluation of semi-supervised and self-supervised techniques, on which we are establishing two baselines for spectral representation learning.

En observation de la Terre, l'un des principaux défis est de cartographier l'artificialisation des sols, qui a été identifiée comme un enjeu majeur de politique publique dans le plan Biodiversité 2018 du ministère de la Transition Écologique. L'imperméabilisation des surfaces naturelles par des surfaces artificielles imperméables a en effet des conséquences majeures sur l'hydrologie des bassins versants, en particulier sur les inondations et les sécheresses, sur les effets d'îlots de chaleur urbains et sur la séquestration du carbone dans les sols. Si les techniques de cartographie automatique ont permis de produire des cartes d'usage des sols à grande échelle (comprenant des classes telles que des forêts, zones bâties, vignobles...) à partir d'images satellite, les méthodes actuelles peinent à produire des cartes des surfaces imperméabilisées (telles que du béton ou de l'asphalte) à haute résolution spatiale à l'échelle d'une métropole. L'imagerie hyperspectrale aéroportée présente un grand potentiel pour la discrimination de l'occupation des sols, grâce à sa haute résolution spatiale et à sa haute résolution spectrale sur un grand domaine spectral. Les techniques de cartographie de l'état de l'art optimisent les modèles de segmentation sémantique (c'est-à-dire des modèles qui segmentent l'image en zones auxquelles des classes sont attribuées) à partir d'un sous-ensemble de pixels annotés, appelé la base d'apprentissage. L'annotation des pixels nécessitant des campagnes terrain coûteuses, la base d'apprentissage est généralement très petite au regard de la diversité des matériaux dans les milieux urbains et de la grande variabilité spectrale intra-classe, qui peut être divisée en trois catégories : la variabilité *physique* induite par différentes conditions d'éclairage, la variabilité *intrinsèque* induite par de légères variations dans la composition chimique des matériaux et la variabilité *sémantique* résultant du fait que différents matériaux sont en fait rassemblés au sein des mêmes classes. Dans ce contexte, les trois principales contributions de cette thèse consistent à exploiter l'information des millions de pixels non annotés dans l'image hyperspectrale afin d'apprendre des représentations de données discriminantes et robustes aux variabilités intra-classes (et incidemment aux similitudes inter-classes). Notre première contribution est d'étudier le potentiel des algorithmes d'Active Learning (AL) à sélectionner quelques pixels supplémentaires à annoter afin d'améliorer la qualité de la base d'apprentissage. Nos expériences numériques montrent que différentes stratégies d'AL sont

complémentaires et que l'information sémantique a priori dérivée de la hiérarchie des classes perméables et imperméables peut efficacement être intégrée dans le processus d'AL, résultant dans des gains importants de la précision des modèles de segmentation pour quelques centaines de pixels annotés supplémentaires. Néanmoins, le potentiel des techniques d'AL est limité par la capacité des modèles de segmentation (utilisés par les méthodes d'AL elles-mêmes) à apprendre des représentations discriminantes. Ainsi, notre deuxième contribution est d'intégrer des connaissances a priori dérivées de lois physiques dans un modèle d'apprentissage automatique afin d'apprendre, à partir de données annotées et non annotées, des représentations spectrales qui sont robustes aux variations intra-classes *physiques*. Nos expériences montrent que notre modèle hybride a de meilleures capacités d'extrapolation que les modèles conventionnels d'apprentissage automatique pour les matériaux dont les conditions d'éclairage ne sont pas représentées dans la base d'apprentissage. Enfin, notre troisième contribution est d'étudier le potentiel des méthodes non supervisées et auto-supervisées pour l'apprentissage de représentations spectrales robustes aux variations *intrinsèques* et *sémantiques*. À cette fin, nous avons construit et publié une très grande base de données hyperspectrales particulièrement adaptée à l'évaluation des méthodes semi-supervisées et auto-supervisées, sur laquelle nous établissons deux méthodes de référence pour l'apprentissage de représentation spectrale.

Mots clés

hyperspectral imaging, semantic segmentation, machine learning, impermeable surfaces