

---

# DÉTECTION DE CHANGEMENTS POUR L'OBSERVATION DE LA TERRE

---

**Laboratoire:** LASTIG lab. (Univ Gustave Eiffel, IGN, ENSG)

**Localisation:** LASTIG, IGN-ENSG, Saint-Mandé, France

**Encadrement:** Nicolas Gonthier, PhD & Clément Mallet, PhD HDR;

**Démarrage:** Septembre 2023

**Mot-Clés:** Apprentissage profond, apprentissage auto-supervisé, imagerie géospatiale, détection de changements, bases de données topographiques

**Environnement de développement:** Linux, Python, PyTorch

## 1 Contexte

Face à l'érosion de la biodiversité et au réchauffement climatique, des mesures s'imposent pour préserver l'état de la planète. Ces mesures de préservation nécessitent de connaître l'état de la planète mais aussi d'être capable d'en suivre l'évolution avec la meilleure actualité possible.

Les données d'observation de la Terre (OT) sont désormais abondantes grâce à des efforts internationaux tels que le déploiement de la constellation européenne Sentinel-2, qui acquiert des images en couleur de toutes les zones du globe tous les cinq jours avec une résolution de 10 mètres. Des images de plus haute résolution peuvent être obtenues en France grâce à des programmes nationaux tels que SPOT-6/7 (Airbus/CNES, 1,5m/pixel), Pléiades-Pléiades Néo (Airbus, 0,3 m/px) ou la BDORTHO (IGN, 0,2 m/px). Cependant ces images aériennes et satellites ne sont acquis qu'avec une fréquence allant d'un à trois ans.

Ainsi, l'observation de la Terre et le traitement des images sont une des sources d'information pour la mise en place de nombreuses politiques de protection des espaces naturels terrestres et marins. Cependant, l'annotation des images satellites ou aériennes est une tâche longue, coûteuse et fastidieuse qui doit être continuellement renouvelée. Maintenant, qu'on était constitué de larges bases d'informations décrivant le territoire de différentes manières (tel que la BD Topo ou la BD Forêt de l'IGN), se pose la question de leur mise à jour. Ainsi la tâche principale est d'être capable de détecter une différence entre une information présente dans une base de connaissances cartographique d'une région donnée et une nouvelle acquisition. D'une part, les changements sont des phénomènes visuellement très variés mais aussi très rares. Par exemple, au sein de l'IGN, il a été montré que sur le département du Gers, seulement 5% du territoire avait été notablement modifié en 3 ans. Cependant, le parcours exhaustif de la totalité des acquisition est nécessaire pour de ne pas manquer un changement important. L'investigation complète d'un département requièrent environ 300 heures de travail. C'est pourquoi il est crucial de fournir de nouvelles méthodes pour assister et accélérer la mise à jour de la donnée géographique.

Le domaine de l'analyse d'image a connu cette dernière décennie une progression fulgurante, grâce au renouveau d'outils relativement anciens de l'intelligence artificielle, notamment les réseaux de neurones [Goodfellow et al., 2016]. Les réseaux convolutionnels ont permis d'obtenir des performances inédites en terme de

reconnaissance d'objets, au prix d'apprentissages supervisés massifs [Russakovsky et al., 2015]. Cependant, l'apprentissage automatique pour l'observation de la Terre est presque toujours confronté à un manque d'ensembles de données annotées. Une grande partie des données disponibles est non annotée, c'est à dire sans information sémantique associée. Cette immense ensemble de données pourrait être utilisé comme une source d'information complémentaire pour l'apprentissage de modèles pour des applications telles que la surveillance de l'artificialisation des sols, de la déforestation ou du suivi des cultures.

Un point positif est que l'OT n'est pas le seul domaine où les annotations sont rares. L'annotation d'événements événements rares est un problème majeur dans la plupart, sinon la totalité, des applications d'apprentissage automatique. Une combinaison de deux solutions est adoptée pour pallier ce problème : l'augmentation de données (pour générer de nouvelles données annotées) et le pré-entraînement non supervisé ou auto-supervisé (pour exploiter des données non étiquetées).

**Le but de cette thèse est de s'appuyer sur les progrès récents dans le domaine l'apprentissage automatique pour développer des outils de détection de changement destinés à l'observation de la Terre en vue d'avoir un suivi des territoires plus fréquent.**

## 2 Sujet

Le travail de thèse consistera à développer des outils de détection de changements pour l'observation de la Terre reposant sur les progrès récents de l'apprentissage automatique. Contrairement à la plupart des travaux de détection de changements, le sujet:

- reposera sur des problématiques réelles opérationnelles issues de la surveillance du territoire, en particulier en lien avec la mise à jour des bases de données géographiques de l'IGN,
- identifiera les défis méthodologiques à relever pour bénéficier des progrès récents dans le domaine de l'apprentissage automatique, en particulier en ce qui concerne des problématiques d'augmentation de données et de pré-apprentissage mais dans le cadre de la télédétection.

Notons que la détection de changements regroupent plusieurs cas d'application. Trois cas d'usage différents seront considérés :

- La comparaison de deux images d'une même zone obtenus par le même capteur;
- La comparaison de deux images d'une même zone obtenus par deux capteurs différents (aérienne et satellite par exemple);
- La comparaison entre une image et l'état de la connaissance passée de la zone (par exemple représentation sémantique de la zone).

Méthodologiquement, ce dernier cas est le plus complexe et le plus intéressant car il permet de mettre à jour la base de connaissances que l'on a du territoire et qui a pu déjà être mise à jour partiellement par d'autres voies. Cette base de connaissances peut aussi contenir des erreurs de saisies passées qui seront alors corrigées.

### 2.1 Problématiques applicatives

#### 2.1.1 Jeu de données synthétique pour la détection de changements en télédétection

Cette première direction consiste à produire une méthodologie pour créer des jeux de données synthétiques mais réalistes pour la détection de changements en télédétection. En effet, *il n'existe actuellement aucun jeu de données massif accessible au public pour le problème de détection des changements dans les images d'OT*. Par exemple, le récent jeu de données DynamicEarth [Toker et al., 2022] ne contient que 1800 paires d'images annotées, le dataset OSCD [Daudt et al., 2018b] n'en compte qu'une vingtaine mais de haute résolution et le dataset HRSCD [Daudt et al., 2019] en contient 500 paires d'images mais accompagnées d'annotations peu exploitables. Pour

éviter d'avoir à trier et étiqueter manuellement les changements observés entre des milliers de paires d'images, il serait intéressant de proposer une procédure pour exploiter les ensembles de données d'images à grande échelle existants déjà annotés telles que la BD Topo de l'IGN ou bien des bases de détections d'objets (bâti, réseau, etc). L'avantage de considérer des bases de haute précision comme la BD Topo IGN est que celle-ci contient l'information de segmentation individuelle de certains bâtiments contrairement à des bases d'occupation du sol qui va regrouper tout le bâti dans un seul et même ensemble [Zheng et al., 2021].

Nous proposons donc de créer des jeux de données synthétiques en supprimant ou rajoutant des objets issus des annotations sémantiques d'objets d'intérêt. Il s'agit en quelque sorte de faire des copier-coller ponctuels d'éléments d'intérêt (un bâtiment, un arbre, une route, une rivière etc) d'une image vers une autre. Pour se faire, on se reposera sur des méthodes d'*inpainting* à base d'apprentissage [Liu et al., 2018] ou non, ainsi que d'édition d'images [Pérez et al., 2003] pour simuler des "changements" visuellement réalistes. Cela a déjà été étudié pour la mise en correspondance de zones d'images [Shen et al., 2022], la découverte d'objets [Arandjelović and Zisserman, 2019] ou bien pour la détection de changements dans le cadre des images naturelles [Sachdeva and Zisserman, 2023]. Il s'agit de modifier les images directement tandis que [Ji et al., 2019] propose de créer un jeu de données synthétiques de cartes de segmentation de bâti pour ensuite entraîner un modèle de détection de changement pour ce thème particulier.

Une des particularités de cette piste de recherche est qu'il sera nécessaire de définir des transformations possibles et réalisables entre les paires d'images synthétiques. Ces transformations géométriques et colorimétriques devront correspondre aux invariants que l'on souhaiterait que le modèle possède (par exemple selon le type de capteur). Ces transformations pourront être issues de modèles physiques ou bien de méthodes de transfert de style [Zhu et al., 2017]. Il sera peut être aussi nécessaire d'étudier la manière de faire ces copier-coller de manière plus efficace de l'aléatoire par exemple en prédisant où il serait pertinent de créer un nouvel élément (par exemple rajouter une maison à côté d'une route et non pas au milieu d'une montagne).

Une telle approche permettrait de *créer des jeux de données synthétiques contenant des centaines de milliers de paires d'images alignés et de la carte de changements correspondante*. Les modèles développés pourront se fonder sur les architectures déjà développées pour la détection de changements entre deux images à base de réseaux de neurones convolutionnels [Daudt et al., 2018b, Daudt et al., 2018a, Gao et al., 2019, Zheng et al., 2021] ou bien de Transformers [Bandara and Patel, 2022, Yan et al., 2022].

Cette première piste de recherche correspond à un projet pour la première année de thèse et permettra de prendre en main le sujet.

### 2.1.2 Développement de modèle de mise à niveau de base de données

La plupart des méthodes de détection de changements s'intéressent aux changements entre deux images de télédétection. Or, il existe de plus en plus de base de connaissances des régions et la tâche importante est donc maintenant de mettre à niveau celles-ci et non pas uniquement de comparer deux images à des dates différentes. Il s'agirait donc d'être capable de détecter des changements entre une base de connaissances du territoire à la date T-1 et un ensemble d'images (issues de sources diverses et variés à la date T). Pour se faire il serait certainement nécessaire de développer un nouveau type d'architecture d'apprentissage profond capable de prendre comme entrée une image du territoire à la date T et l'état de la base de connaissance à la date T-1.

Pour cela, on pourra s'inspirer de ce qui a été fait par [Audebert et al., 2017] qui propose de prendre comme information d'entrée une image satellite et une version grossière d'une carte de la zone en vue de générer une segmentation sémantique plus détaillée de la zone en question. Dans notre cas, l'application serait différente car il s'agirait de produire une carte de détection de changements entre la base de connaissance et une prise de vue plus récente de la région.

Étant donné le caractère parcimonieux du signal de supervision (la carte de changements), il sera peut être nécessaire de mettre en place une supervision annexe du modèle (de segmentation sémantique par exemple) pour obtenir de bonnes performances de détection. Il se posera aussi le problème de la représentativité de la donnée issue des bases de connaissances. En effet, des choix ou des simplifications / généralisation sont souvent faite du passage de l'information sous forme de données cartographiques (vectorisée ou non).

Le type de modèle développé devra être agnostique au type d'images en entrée (satellite ou aérienne) mais aussi à l'annotation considérée (infrastructures, type de végétation etc).

### 2.1.3 Apprentissage sur des sources de données différentes et non annotées

Cette autre direction de recherche se concentre sur le transfert des modèles pré-entraînés sur des données non annotées, issues de la même modalité ou bien d'une autre (aérienne ou spatiale). L'état de l'art des méthodes de pré-entraînement par auto-apprentissage se concentre actuellement sur la classification ou la segmentation sémantique [Wang et al., 2022], il s'agirait de trouver des tâches prétextes qui soient efficaces pour la détection de changements dans le cadre de la télédétection. En effet, certains travaux ont commencé à étudier la mise en place de tâche prétextes qui soient spécifiques au domaine de la télédétection tel que la temporalité [Mañas et al., 2021] ou bien le géo-référencement [Ayush et al., 2022]. Mais dans le cadre de l'auto-apprentissage pour la détection de changements, les travaux de recherches se sont uniquement intéressés à des tâches prétextes *génériques* issues de la vision par ordinateur pour les images naturelles [Leenstra et al., 2021, Chen and Bruzzone, 2022, Dong et al., 2020]. Il s'agirait donc de mieux prendre en compte les spécificités des images de télédétection d'un point de vue géo-spatiale et physique.

Une fois, le modèle pré-entraîné à l'aide d'auto-apprentissage, celui-ci pourra être "fine-tuné" (re-paramétré) sur la tâche de détection de changements ou bien utilisé dans un cadre semi-supervisé. Cette seconde piste consiste en des techniques de propagation d'annotations et utilisent ensuite des prédictions cohérentes comme pseudo-labels [Mei et al., 2020] pour enrichir automatiquement des paires d'images alignés mais non annotées en vue d'entraîner des modèles profonds.

En outre, cette direction de recherche pourrait rentrer en synergie avec l'autre sujet de thèse proposé par l'IGN et portant sur l'apprentissage réciproque croisé en Observation de la Terre.

### 2.1.4 Détection de changements complet

Actuellement, les méthodes de détection de changements se focalisent sur un type de changement particulier, par exemple le changement en milieu urbain [Daudt et al., 2018b], le changement d'occupation du sol [Toker et al., 2022], l'étude de mer de glaces [Gao et al., 2019], etc. Il serait intéressant d'être capable d'effectuer une détection de changements *complet* pour chaque pixel d'une paire de deux images optiques à la manière d'une segmentation panoptique. Il s'agirait pour le modèle de prédire l'origine du changement observé entre chacun des pixels. Ce changement peut avoir plusieurs raisons :

- Radiométrique (différence entre les capteurs);
- Saisonnière (changement de la végétation selon la saison, orientation du soleil);
- Anthropique ponctuelle (déplacement de véhicules ou bien de ressources);
- Anthropique pérenne (déboisement, artificialisation des sols etc.).

Pour être capable d'apprendre ce type d'informations, il est primordial de travailler sur plusieurs modalités d'images en entrée de notre modèle mais tout particulièrement des séries temporelles d'images pour capturer les effets de saisonnalité par exemple. Sur ce dernier axe de recherche, l'usage de méthodes bayésiennes pourrait permettre de mieux gérer cet ensemble de causes cumulatives de changements. Le paradigme bayésien a été expérimenté pour la tâche de détection de changements par [Gharbi et al., 2021] de manière à mieux prendre en compte l'incertitude sur les données. Il serait intéressant de considérer les modèles à l'intersection entre les réseaux de neurones profonds et les approches bayésiennes qui émergent pour la vision par ordinateur [Jospin et al., 2022].

## 3 Conclusion

La détection de changements pour l'OT est un sujet important car il permet d'avoir une meilleure connaissance de l'état de la planète, une connaissance la plus fraîche possible. Il y a cependant encore un certains nombres de verrous à faire sauter avant d'obtenir des algorithmes robustes et efficaces. Les méthodes et algorithmes développés dans le cadre de cette thèse, sur la thématique de la détection de changements pourrait aussi être appliqués à d'autres cas d'applications tel que les images de Street View [Lei et al., 2021] mais aussi pour l'imagerie médicale et le suivi des patients dans le temps ou bien l'analyse d'oeuvres d'arts pour estimer les différences entre différents tirages d'une même oeuvre.

## References

- [Arandjelović and Zisserman, 2019] Arandjelović, R. and Zisserman, A. (2019). Object Discovery with a Copy-Pasting GAN. *arXiv:1905.11369*.
- [Audebert et al., 2017] Audebert, N., Le Saux, B., and Lefevre, S. (2017). Joint Learning from Earth Observation and OpenStreetMap Data to Get Faster Better Semantic Maps. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1552–1560, Honolulu, HI. IEEE.
- [Ayush et al., 2022] Ayush, K., Uzkent, B., Meng, C., Tanmay, K., Burke, M., Lobell, D., and Ermon, S. (2022). Geography-Aware Self-Supervised Learning. In *IEEE/CVF International Conference on Computer Vision*.
- [Bandara and Patel, 2022] Bandara, W. G. C. and Patel, V. M. (2022). A Transformer-Based Siamese Network for Change Detection.
- [Chen and Bruzzone, 2022] Chen, Y. and Bruzzone, L. (2022). Self-supervised Change Detection in Multi-view Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–12.
- [Daudt et al., 2018a] Daudt, R., Le Saux, B., and Boulch, A. (2018a). Fully Convolutional Siamese Networks for Change Detection. In *25th IEEE International Conference on Image Processing (ICIP)*, pages 4063–4067.
- [Daudt et al., 2019] Daudt, R., Le Saux, B., Boulch, A., and Gousseau, Y. (2019). Multitask learning for large-scale semantic change detection. *Computer Vision and Image Understanding*, 187:102783.
- [Daudt et al., 2018b] Daudt, R. C., Le Saux, B., Boulch, A., and Gousseau, Y. (2018b). Urban Change Detection for Multispectral Earth Observation Using Convolutional Neural Networks. In *2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 2115–2118.
- [Dong et al., 2020] Dong, H., Ma, W., Wu, Y., Zhang, J., and Jiao, L. (2020). Self-Supervised Representation Learning for Remote Sensing Image Change Detection Based on Temporal Prediction. *Remote Sensing*, 12(11):1868.
- [Gao et al., 2019] Gao, F., Wang, X., Gao, Y., Dong, J., and Wang, S. (2019). Sea Ice Change Detection in SAR Images Based on Convolutional-Wavelet Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, 16(8):1240–1244.
- [Gharbi et al., 2021] Gharbi, W., Chaari, L., and Benazza-Benyahia, A. (2021). Unsupervised Bayesian change detection for remotely sensed images. *Signal, Image and Video Processing*, 15(1):205–213.
- [Goodfellow et al., 2016] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
- [Ji et al., 2019] Ji, S., Shen, Y., Lu, M., and Zhang, Y. (2019). Building Instance Change Detection from Large-Scale Aerial Images using Convolutional Neural Networks and Simulated Samples. *Remote Sensing*, 11(11):1343.
- [Jospin et al., 2022] Jospin, L. V., Laga, H., Boussaid, F., Buntine, W., and Bennamoun, M. (2022). Hands-On Bayesian Neural Networks—A Tutorial for Deep Learning Users. *IEEE Computational Intelligence Magazine*, 17(2):29–48.
- [Leenstra et al., 2021] Leenstra, M., Marcos, D., Bovolo, F., and Tuia, D. (2021). Self-supervised pre-training enhances change detection in Sentinel-2 imagery. In *Pattern Recognition and Remote Sensing (PRRS) Workshop*. arXiv.
- [Lei et al., 2021] Lei, Y., Peng, D., Zhang, P., Ke, Q., and Li, H. (2021). Hierarchical Paired Channel Fusion Network for Street Scene Change Detection. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 30:55–67.

- [Liu et al., 2018] Liu, G., Reda, F. A., Shih, K. J., Wang, T.-C., Tao, A., and Catanzaro, B. (2018). Image Inpainting for Irregular Holes Using Partial Convolutions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 85–100.
- [Mañas et al., 2021] Mañas, O., Lacoste, A., Giro-i-Nieto, X., Vazquez, D., and Rodriguez, P. (2021). Seasonal Contrast: Unsupervised Pre-Training from Uncurated Remote Sensing Data. In *IEEE/CVF International Conference on Computer Vision*.
- [Mei et al., 2020] Mei, K., Zhu, C., Zou, J., and Zhang, S. (2020). Instance Adaptive Self-training for Unsupervised Domain Adaptation. In Vedaldi, A., Bischof, H., Brox, T., and Frahm, J.-M., editors, *Computer Vision – ECCV 2020*, volume 12371, pages 415–430, Cham. Springer International Publishing.
- [Pérez et al., 2003] Pérez, P., Gangnet, M., and Blake, A. (2003). Poisson image editing. *ACM Transactions on graphics (TOG)*, 22(3):313–318.
- [Russakovsky et al., 2015] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3):211–252.
- [Sachdeva and Zisserman, 2023] Sachdeva, R. and Zisserman, A. (2023). The Change You Want to See. In *WACV*.
- [Shen et al., 2022] Shen, X., Efros, A. A., Joulin, A., and Aubry, M. (2022). Learning Co-segmentation by Segment Swapping for Retrieval and Discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 5082–5092.
- [Toker et al., 2022] Toker, A., Kondmann, L., Weber, M., Eisenberger, M., Camero, A., Hu, J., Hoderlein, A. P., Şenaras, Ç., Davis, T., Cremers, D., Marchisio, G., Zhu, X. X., and Leal-Taixé, L. (2022). DynamicEarthNet: Daily Multi-Spectral Satellite Dataset for Semantic Change Segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21158–21167.
- [Wang et al., 2022] Wang, Y., Albrecht, C. M., Braham, N. A. A., Mou, L., and Zhu, X. X. (2022). Self-supervised Learning in Remote Sensing: A Review. *IEEE Geoscience And Remote Sensing Magazine*.
- [Yan et al., 2022] Yan, T., Wan, Z., and Zhang, P. (2022). Fully Transformer Network for Change Detection of Remote Sensing Images.
- [Zheng et al., 2021] Zheng, Z., Ma, A., Zhang, L., and Zhong, Y. (2021). Change is Everywhere: Single-Temporal Supervised Object Change Detection in Remote Sensing Imagery. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 15173–15182, Montreal, QC, Canada. IEEE.
- [Zhu et al., 2017] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks Web Page. In *Computer Vision (ICCV), 2017 IEEE International Conference On*.