



HAL
open science

Analyse de techniques d'apprentissage sur graphes pour la segmentation sémantique de séries temporelles d'images satellitaires

Corentin Dufourg, Charlotte Pelletier, Stéphane May, Sébastien Lefèvre

► To cite this version:

Corentin Dufourg, Charlotte Pelletier, Stéphane May, Sébastien Lefèvre. Analyse de techniques d'apprentissage sur graphes pour la segmentation sémantique de séries temporelles d'images satellitaires. ORASIS 2023, Laboratoire LIS, UMR 7020, May 2023, Carqueiranne, France. hal-04219303

HAL Id: hal-04219303

<https://hal.science/hal-04219303v1>

Submitted on 27 Sep 2023

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.

Analyse de techniques d'apprentissage sur graphes pour la segmentation sémantique de séries temporelles d'images satellitaires

C. Dufourg¹

C. Pelletier¹

S. May²

S. Lefèvre¹

¹ Université Bretagne Sud, IRISA, UMR CNRS 6074, Vannes, France

² Centre National d'Études Spatiales (CNES), Toulouse, France

corentin.dufourg@univ-ubs.fr

Résumé

Les nouvelles constellations de satellites permettent l'acquisition d'images à hautes résolutions temporelle et spatiale en tout point de la Terre. Ces données, assemblées sous forme de séries temporelles d'images satellitaires (SITS), représentent une source d'informations importante pour suivre l'évolution de la surface terrestre. L'apprentissage profond est une des solutions les plus prometteuses pour permettre l'analyse automatique des gros volumes de données acquises par les nouvelles générations de satellites. Cependant, ces techniques n'exploitent bien souvent que les structures temporelles ou spatiales. Afin de tirer bénéfice de la complémentarité temporelle et spatiale des données, nous utilisons dans ce travail une modélisation sous forme de graphes dans le cadre de l'apprentissage profond. Nous proposons ainsi une comparaison de six réseaux de neurones pour graphes issus de l'état de l'art appliqués aux SITS.

Mots Clef

Graphe spatio-temporel, réseaux de neurones pour graphes, séries temporelles d'images satellitaires, télédétection.

Abstract

The new satellite constellations allow the acquisition of high temporal and spatial resolution images at any point on the Earth. These data, assembled in the form of satellite image time series (SITS), are an important source of information for monitoring the evolution of the Earth's surface. Deep learning is one of the most promising solutions for the automatic analysis of large volumes of data acquired by new generations of satellites. However, these techniques often only exploit temporal or spatial structures. In order to take advantage of the temporal and spatial complementarity of the data, we use in this work a graph-based modeling in the context of deep learning. We thus propose a comparison of six graph neural networks applied to SITS.

Keywords

Spatio-temporal graph, graph neural network, satellite image time series, remote sensing.

Analyse par graphe pour des séries temporelles d'images satellitaires

Les progrès significatifs des techniques d'observation de la Terre au cours des dernières années permettent aux systèmes d'acquisition modernes, *e.g.* les constellations de satellites Sentinel-2 et Pléiades, de capturer à hautes résolutions spatiales et temporelles l'ensemble des surfaces émergées. Grâce à la capacité de revisite de ces satellites, les données peuvent être organisées sous forme d'un cube spatio-spectro-temporel, aussi appelé séries temporelles d'images satellitaires (SITS) et formé par une séquence d'images acquises pour la même localisation mais à des instants différents. Cette disponibilité des données ouvre ainsi la perspective d'un suivi précis des surfaces terrestres à l'échelle mondiale, comme la cartographie de l'occupation des sols [9].

En raison de leur volume et de leur complexité, les SITS ne peuvent pas être traitées manuellement à grande échelle et nécessitent donc le développement d'outils de traitement automatique. Une solution adoptée par la communauté de la télédétection est l'utilisation de techniques d'apprentissage automatique et profond. En particulier, les avancées récentes dans l'analyse des SITS ont conduit au développement de nouvelles méthodes d'apprentissage profond permettant de tirer le meilleur parti de la structure temporelle contenue dans les SITS [15] ou encore d'exploiter conjointement les dimensions spatiales et temporelles des SITS [16]. Ces approches adaptées à la classification de séries temporelles sont toutefois incapables de produire une prédiction pour chaque instant de la série, ce qui réduit leur utilité pour analyser précisément dans le temps l'évolution des paysages nécessaire, par exemple, pour le suivi de systèmes de culture multiples [6]. De plus, ces techniques nécessitent de travailler avec les images complètes à haute résolution, ce qui peut être coûteux en mémoire car les SITS peuvent être volumineuses. Pour résoudre ce problème, l'approche orientée objets est traditionnellement utilisée en télédétection [2].

L'approche orientée objets ne se soucie plus uniquement des caractéristiques des pixels indépendamment les uns des autres. La structure euclidienne régulière de l'image est abandonnée afin de prendre en compte des zones spatialement homogènes, les objets, comme unités de travail. Le traitement de ces objets

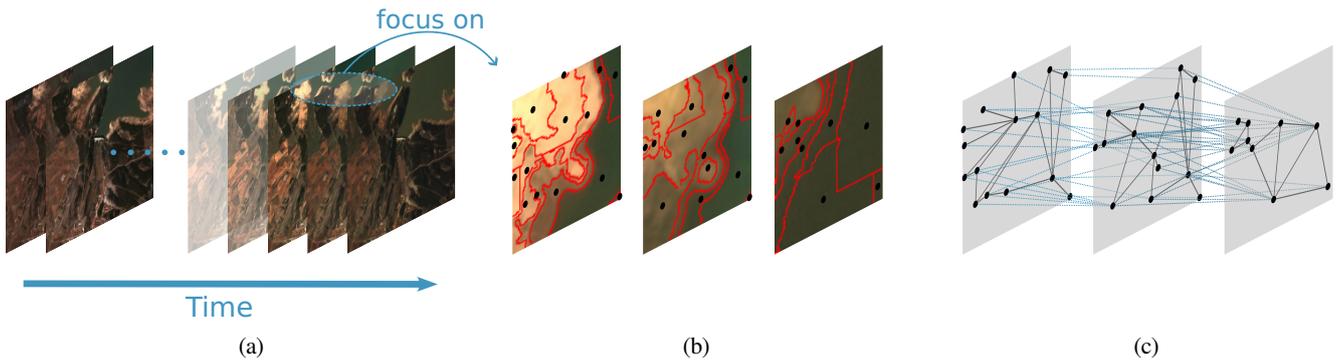


FIGURE 1 – Exemple de graphe spatio-temporel à partir d’une série temporelle d’images satellitaires extraite du jeu de données DynamicEarthNet. Dans la série temporelle d’images satellitaires (a), les objets (b) sont extraits à chaque instant (par un algorithme de superpixels par exemple). Ensuite, un graphe spatio-temporel est construit en reliant ces objets avec des arêtes spatiales (en gris) et temporelles (en bleu) (c).

peut être réalisé indépendamment les uns des autres, ou via un graphe d’adjacence de régions pour prendre en compte les relations spatiales entre eux ; les nœuds du graphe représentent les objets et une arête relie deux nœuds si leurs régions sont adjacentes. En considérant la dimension temporelle en plus de la dimension spatiale, ce paradigme peut être étendu aux SITS pour obtenir des graphes dynamiques comme cela a été fait dans certains travaux antérieurs [7].

Dans la littérature, un graphe est dit dynamique si sa structure ou ses attributs dépendent du temps [10]. Parmi les représentations sous forme de graphe existantes, le graphe spatio-temporel (STG), contenant des arêtes spatiales et temporelles comme le montre la Figure 1, est adapté à l’analyse des SITS. Ainsi, avec l’approche orientée objets, une SITS peut être représentée et utilisée de différentes manières grâce à une représentation en graphe dynamique afin de réduire la complexité mémoire tout en limitant la perte d’information.

Apprentissage profond sur graphes spatio-temporels pour la classification de nœuds

Pour apprendre à partir de cette représentation spécifique des données, une approche intéressante est l’utilisation des réseaux de neurones pour graphes (GNN). Ceux-ci reposent sur le même principe d’optimisation que les réseaux de neurones classiques appliqués aux images et aux textes mais ils prennent en compte les caractéristiques structurelles des graphes, comme la notion de voisinage des nœuds et leur invariance à la permutation. Les approches GNN les plus récentes reposent sur des convolutions spectrales et spatiales [8, 11], la récurrence [3, 14], et les mécanismes d’attention [13, 18]. En télédétection, seuls quelques travaux ont utilisé les GNN sur des STG extraits de SITS, en se limitant aux architectures basées sur les mécanismes d’attention [4, 5].

Ainsi, l’objectif de ces travaux est d’étudier la performance des GNN sur les SITS. Nous proposons donc une **étude comparative des réseaux de neurones pour graphes sur les séries temporelles d’images satellitaires** avec les deux contributions suivantes :

1. un nouveau jeu de données de graphes spatio-temporels construit à partir de SITS,
2. ainsi qu’une comparaison des performances de six modèles de GNN pour la classification de nœud (GCN [11], GraphSAGE [8], GatedGCN [3], EvolveGCN [14], GAT [18], SAN [13]) sur ce nouvel ensemble de données afin de déterminer quels modèles s’adaptent aux SITS.

Il est à noter que l’identification d’objets sur des images satellitaires n’est pas aisée et rend la construction de STG complexe. En effet, un grand nombre de techniques de segmentation d’objets existent et son choix impacte fortement la création des STG. Bien qu’une étude approfondie serait nécessaire ultérieurement, l’utilisation d’une approche standard, comme l’algorithme de superpixels SLIC [1], permet de mettre en lumière l’intérêt des STG pour l’analyse des SITS. Un tel ensemble de données adapté aux SITS fait en effet défaut à la communauté de télédétection pour évaluer les futures contributions dans ce domaine. Bien qu’il existe de nombreux jeux de données pour la classification des SITS [12, 15, 19], aucune de ces données ne peut être utilisée pour une tâche de classification de nœuds sur les STG car les annotations sont données pour la série temporelle entière. Nous proposons donc un jeu de données de graphes spatio-temporels dérivé de DynamicEarthNet [17]. L’avantage de DynamicEarthNet est de fournir des annotations pour chaque pixel et à chaque instant. Ainsi, il permet de cartographier précisément l’occupation des sols sur plusieurs dates, ce qui représente une tâche plus difficile. Les SITS sont composées d’images multispectrales PlanetScope avec des réflectances dans le visible et le proche infrarouge, et ont des résolutions spatiales (3 mètres par pixel) et temporelles (acquisition mensuelle) élevées. Ces caractéristiques permettent donc l’utilisation d’une structure de graphe spatio-temporel.

A l’issue de nos travaux, le jeu de données de graphes spatio-temporels construits à partir des SITS, le code, les modèles pré-entraînés ainsi que les résultats obtenus seront librement mis à disposition de la communauté.

Références

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11) : 2274–2282, 2012.
- [2] T. Blaschke, S. Lang, and G. Hay. *Object-based image analysis : spatial concepts for knowledge-driven remote sensing applications*. Springer Science & Business Media, 2008.
- [3] X. Bresson and T. Laurent. Residual gated graph convnets. *arXiv preprint arXiv :1711.07553*, 2017.
- [4] A. M. Censi, D. Ienco, Y. J. E. Gbodjo, R. G. Pensa, R. Interdonato, and R. Gaetano. Attentive spatial temporal graph CNN for land cover mapping from multi temporal remote sensing data. *IEEE Access*, 2021.
- [5] F. Chaabane, S. Réjichi, and F. Tupin. Self attention deep graph CNN classification of times series images for land cover monitoring. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 279–282, 2022.
- [6] I. Del’Arco Sanches, R. Q. Feitosa, P. M. Achanccaray Diaz, M. Dias Soares, A. J. Barreto Luiz, B. Schultz, and L. E. Pinheiro Maurano. Campo verde database : Seeking to improve agricultural remote sensing of tropical areas. *IEEE Geoscience and Remote Sensing Letters*, 15(3) :369–373, 2018.
- [7] F. Guttler, D. Ienco, J. Nin, M. Teisseire, and P. Poncelet. A graph-based approach to detect spatiotemporal dynamics in satellite image time series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130 :92–107, 2017.
- [8] W. Hamilton, Z. Ying, and J. Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.
- [9] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes. Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, 9(1), 2017.
- [10] S. M. Kazemi, R. Goel, K. Jain, I. Kobyzev, A. Sethi, P. Forsyth, and P. Poupart. Representation learning for dynamic graphs : A survey. *Journal of Machine Learning Research*, 21(70) :1–73, 2020.
- [11] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2017.
- [12] L. Kondmann, A. Toker, M. Rußwurm, A. Camero Unzueta, D. Peressuti, G. Milcinski, N. Longépé, P.-P. Mathieu, T. Davis, G. Marchisio, et al. DENETHOR : The DynamicEarthNET dataset for Harmonized, inter-Operable, analysis-Ready, daily crop monitoring from space. In *35th Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, pages 1–13, 2021.
- [13] D. Kreuzer, D. Beaini, W. Hamilton, V. Létourneau, and P. Tossou. Rethinking graph transformers with spectral attention. *Advances in Neural Information Processing Systems*, 34 :21618–21629, 2021.
- [14] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. Schardl, and C. Leiserson. Evolvegn : Evolving graph convolutional networks for dynamic graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 5363–5370, 2020.
- [15] M. Rußwurm, C. Pelletier, M. Zollner, S. Lefèvre, and M. Körner. Breizhcrops : A time series dataset for crop type mapping. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences ISPRS*, 2020.
- [16] V. Sainte Fare Garnot, L. Landrieu, S. Giordano, and N. Chehata. Satellite image time series classification with pixel-set encoders and temporal self-attention. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [17] A. Toker, L. Kondmann, M. Weber, M. Eisenberger, A. Camero, J. Hu, A. P. Hoderlein, C. Senaras, T. Davis, D. Cremers, G. Marchisio, X. X. Zhu, and L. Leal-Taixe. DynamicEarthNet : Daily Multi-Spectral Satellite Dataset for Semantic Change Segmentation. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21126–21135, 2022.
- [18] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio. Graph attention networks. *International Conference on Learning Representations (ICLR)*, 2018.
- [19] G. Weikmann, C. Paris, and L. Bruzzone. Timesen2crop : A million labeled samples dataset of Sentinel 2 image time series for crop-type classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14 :4699–4708, 2021.