Learning and analyzing spatio-temporal objects from high resolution satellite image time series

Keywords: Satellite Image Time Series, Sentinel-2, spatio-temporal objects

Context

On March 7, 2017, the European Space Agency (ESA) successfully put its latest high-resolution satellite Sentinel-2B into orbit. The two Sentinel-2 satellites are now capturing images of all emerged surfaces every 2 to 5 days at high spatial and spectral resolutions, which makes it possible to monitor the evolution of land surfaces on a global scale. Satellite image time series (SITS) extracted from Sentinel-2 constellations are useful for many applications such as land cover mapping [1], crop type mapping, soil moisture estimation, burnt area detection, and many more. Similarly, high resolution satellite image time series, such as the one provided by Pléiades, are relevant for urban area analysis, disaster risk management, and rapid mapping.

Due to their volume and complexity, the analysis of these spatio-spectro-temporal datacubes requires automatic tools. Recent advances have been marked by the use of deep learning techniques to make the most of the temporal structure of SITS. It includes temporal 1D convolutions [2], recurrent networks [3], or attention-based architectures [4]. There were also several attempts to jointly exploit the spatial and temporal dimensions of SITS data by the means of deep learning [5, 6, 7]. Although these approaches have proven their efficiency, they suffer from two main issues: (i) they require a vast amount of high-quality labelled data, and (ii) they ignore previous trends in remote sensing, especially contributions from the object-based image analysis (OBIA) [8]. Ensuring a convergence between these two distinct paradigms would allow embedding more structural and semantic information in the process. Preliminary attempts to couple the two paradigms need to be pursued [9].

Objectives

Since only a few attempts exist to jointly use temporal relationships between satellite images and their intrinsic spatial structure in deep learning, the Ph.D. aims at developing novel deep learning architectures for the generation of spatio-temporal objects with no or limited supervision. It will be composed of two main objectives: (i) developing new techniques to structure raw SITS data into spatio-temporal objects, and (ii) analysing spatio-temporal objects.

First, we will consider the task of extracting objects from SITS as either a temporal sequence of 2D objects or directly 3D (*i.e.*, spatio-temporal) objects with no or a few supervision. For this task, we will propose new unsupervised deep learning strategies that take inspiration from self-supervised strategies [10] and go much further than the segmentation networks introduced recently in the literature, e.g. based on auto-encoders such as W-Net [11], or having new loss formulation [12]. We will also consider the case where a weak reference is available as we know that prior knowledge can be used to guide the extraction of objects. In our setting, this prior knowledge can come from sparse, single-date, and low-resolution reference data (*e.g.*, Corine Land Cover) or from the application of a ruleset, applied to each image of the SITS, which enriches the objects with some semantics. Compared to computer vision algorithms used for instance segmentation such as Mask-RCNN, the novelty will be to take into account both spatial and temporal structures of SITS data, to deal with the lack of quality labelled reference data at the object level, and to provide a full partition of the input data.

Second, we will develop new methods to analyze the produced spatio-temporal objects. We will still consider the deep learning framework as a methodology to perform object-based time series analysis. To do so, we will represent objects as nodes in a spatio-temporal graph, such as Graph CNNs [13] and their formulation in the spatio-temporal domain [14]. This representation will be then used for classical applications such as land cover mapping. We will also propose and evaluate similarity measures between nodes in the graph representation to cluster the data.

Candidate profile: We are looking for a candidate with strong data analysis, machine learning, and image processing skills, who is familiar with deep learning techniques. The candidate should also have excellent programming skills in at least one language (Python, C/C++, *etc.*). Knowledge of time series analysis and remote sensing techniques will be appreciated. Good communication skills (at least in English) are required.

Supervision and Collaboration: The Ph.D. candidate will be supervised by Dr. Charlotte Pelletier and Prof. Sébastien Lefèvre, who specialized in time series analysis, deep learning techniques, data science, and image processing. The PhD thesis will be conducted at Université Bretagne Sud, IRISA lab, in Vannes, France. There will also be frequent interactions with Stéphane May from the French Spatial Agency (CNES).
The candidate will also benefit from strong international collaborations, for example with the Department of Geoinformatics, Z_GIS, Paris Lodron University Salzburg, Austria.

Funding: The Ph.D. position will be fully funded by the French space agency (CNES) and the region.

Application: To apply for this position, please send a resume, a motivation letter explaining your understanding and interest in the topic, your Bachelor (Licence) and Master (or equivalent) transcripts, and name and emails of two referees to `charlotte.pelletier@univ-ubs.fr`. You can send your application until **March 16 2022**.

References

[1] Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., & Rodes, I. (2017). Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, 9(1), 95.

[2] Pelletier, C., Webb, G. I., & Petitjean, F. (2019). Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing*, 11(5), 523.

[3] Rußwurm, M., & Korner, M. (2017). Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 11-19).

[4] Rußwurm, M., & Körner, M. (2020). Self-attention for raw optical satellite time series classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, 421-435.

[5] Rußwurm, M., & Körner, M. (2018). Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS International Journal of Geo-Information*, 7(4), 129.

[6] Interdonato, R., Ienco, D., Gaetano, R., & Ose, K. (2019). DuPLO: A DUal view Point deep Learning architecture for time series classificatiOn. *ISPRS Journal of Photogrammetry and Remote Sensing*, 149, 91-104.

[7] Saint Fare Garnot, V., Landrieu, L., Giordano, S., & Chehata, N. (2020). 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* (pp. 12325-12334).

[8] Blaschke, T., Lang, S., & Hay, G. (Eds.). (2008). Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications. *Springer Science& Business Media.*

[9] Mboga, N., Georganos, S., Grippa, T., Lennert, M., Vanhuysse, S., & Wolff, E. (2019). Fully convolutional networks and geographic object-based image analysis for the classification of VHR imagery. *Remote Sensing*, 11(5), 597.

[10] Jing, L., & Tian, Y. (2020). Self-supervised visual feature learning with deep neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence.*

[11] Xia, X., & Kulis, B. (2017). W-net: A deep model for fully unsupervised image segmentation. arXiv preprint arXiv:1711.08506.

[12] Kim, B., & Ye, J. C. (2019). Mumford–Shah loss functional for image segmentation with deep learning. *IEEE Transactions on Image Processing*, 29, 1856-1866.

[13] Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., & Vandergheynst, P. (2017). Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4), 18-42.

[14] Yu, B., Yin, H., & Zhu, Z. (2017). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In P*roceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)* (pp. 3634-3640)