Domain invariant interpretable representation learning for satellite image time-series

A fully funded PhD position is open at the University of Strasbourg (ICube). The position will be jointly funded by the French National Centre for Space Studies (CNES) and the Chair SDIA. The candidate will join the SDC research team under the supervision of Dr Thomas Lampert, the Chair of Data Science and Artificial Intelligence, and join his international team to develop novel deep learning approaches to domain invariant representation learning for satellite image time-series (SITS).

There are two areas in which research into SITS can have impact:

1. Monitoring long-term changes – this is important for government agencies to track climate change, urbanisation (the expansion of cities), desertification, deforestation, and similar threats to the environment.
2. Monitoring short-term changes – this is crucial in monitoring and reacting to natural catastrophes such as forest fires, earthquakes, and floods. The output of this can be used to plan rescue efforts, study their spread so that proper action can be taken, and to study any lasting effects.

The goal of the project is to develop models for learning domain invariant representations using deep learning for the analysis of satellite image time-series.

It is difficult and expensive to annotate the huge amount of data generated by satellites, but this is needed for the success of deep learning algorithms. To overcome this, transfer learning and domain adaptation techniques will be developed to exploit unlabelled data. These techniques allow an algorithm's performance to be improved with minimal (or potentially no) additional annotation, lowering the cost of deployment.

Current work on domain adaptation for time-series uses either weak supervision [1] or attention-based mechanisms [2, 3] for classification, or focus on the related problem of time-series forecasting [4]. However, none of these tackle the problem of learning domain invariant representations that can be applied to multiple locations without further adaptation.

Existing collaborations

The project will rely on the longstanding collaboration with ICUBE-SERTIT. And will draw upon time-series expertise in the former ANR FOSTER and the current ANR HERELLES and ANR HIATUS projects.

Proposed Approach

Initial inspiration will be taken from existing work in the group, a domain adaptation model for remote sensing image data named Semi-Supervised Heterogeneous Image Domain Adaptation (SS-HIDA) [5]. SS-HIDA learns a domain invariant representation of images from heterogeneous data, i.e. image sensors with different characteristics. Two input branches bring the data to the same dimensionality and using a domain critic (in semi-supervised domain adaptation), or contrastive learning (in unsupervised domain adaptation), the representation is forced to be invariant to both inputs. Therefore, the labels from the source domain enable the target domain data to be classified. SS-HIDA's current architecture, however, is not suitable for time-series, which require specific architectures to exploit the particularities of the data type.

Shapelets [6, 7] are powerful time-series representations that are interpretable. The group is developing semi-supervised approaches to learn shapelets in a neural network. These can be integrated into deep learning models and give the potential for shapelet domain adaptations mechanisms and domain invariant representations (using the SS-HIDA framework). Long Short-Term Memory [8] and Transformers with Attention mechanisms [9] will also be investigated.

The resulting representation should be invariant to the domain, applicable to heterogenous/various domains, transferrable, and useable in multiple downstream tasks (clustering, multitask learning, and domain adaptation).

Candidate

The successful candidate will have (or will soon obtain) an MSc in Computer Science or related subject. Experience with deep learning is required and experience with time series and/or remote sensing is a bonus.

Location: Strasbourg is a beautiful medieval city (its historic city centre is a UNESCO World Heritage Site) on the crossroads of Europe, with Germany a tram ride away, and both Switzerland and Luxembourg short train trips away. The University of Strasbourg traces its roots back to the 16th century, has numerous Nobel laureates, and is a member of several prestigious research networks and France's *Initiative d'Excellence*.

ICube and SDC: created in 2013, ICube laboratory brings together 650 researchers in the fields of engineering and computer science. The Data Science and Knowledge research team (SDC) covers a large spectrum of research in artificial intelligence, particularly data science, machine learning, and their applications. The team has close collaborations with several hospital research departments, both in Strasbourg (IHU Strasbourg) and abroad, ICube's remote sensing platform (SERTIT), and several multinational and local companies. Through these, our research impacts medical, biological, and remote sensing research, to name but a few.

Send a letter of motivation, transcript of grades, and your CV to Thomas Lampert l1ampert@uni2stra.fr - !remove the numbers! with [CNES PhD] at the beginning of the subject.

The application deadline is 15/3/2022 and the starting date will be September 2022 (or soon after).

References

[1] G. Wilson, J. R. Doppa and D. Cook, "Multi-Source Deep Domain Adaptation with Weak Supervision for Time-Series Sensor Data," in ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020.

[2] X. Jin, Y. Park, D. C. Maddix, Y. Wang and X. Yan, "Domain Adaptation for Time Series Forecasting via Attention Sharing," in ArXiv, 2021.

[3] R. Cai, J. Chen, Z. Li, W. Chen, K. Zhang, J. Ye, Z. Li, X. Yang and Z. Zhang, “Time Series Domain Adaptation via Sparse Associative Structure Alignment,” in arXiv, 2020.

[4] X. Jin, Y. Park, D. C. Maddix, Y. Wang and X. Yan, “Attention-based Domain Adaptation for Time Series Forecasting,” in ArXiv, 2021.

[5] M. Obrenović, T. Lampert, F. Monde-Kossi, M. Ivanović and P. Gançarski, “SS-HIDA: Semi-Supervised Heterogeneous Image Domain Adaptation,” in MACLEAN: MAChine Learning for EArth ObservatioN Workshop co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), 2021.

[6] J. Grabocka, N. Schilling, M. Wistuba and L. Schmidt-Thieme, “Learning time-series shapelets,” in Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, 2014.

[7] R. Tavenard, J. Faouzi, G. Vandewiele, F. Divo, G. Androz, C. Holtz, P. Marie, R. Yurchak, M. Rußwurm, K. Kolar and E. Woods, “Tslearn, A Machine Learning Toolkit for Time Series Data,” Journal of Machine Learning Research, vol. 21, no. 118, pp. 1-6, 2020.

[8] H. Sepp and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, p. 1735–1780, 1997.

[9] A. Vaswani, “Attention Is All You Need,” in NIPS, 2017.

[10] S. Dupond, “A thorough review on the current advance of neural network structures,” Annual Reviews in Control, vol. 14, p. 200–230, 2019.

[11] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar and P.-A. Muller, “Transfer learning for time series classification,” in IEEE International Conference on Big Data, 2018.