Editorial

Foreword to the Special Issue on Machine Learning for Geospatial Data Analysis

Jan Dirk Wegner [1],* [ID], Ribana Roscher [2] [ID], Michele Volpi [3] [ID] and Fabio Veronesi [4] [ID]

[1] Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry, ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland
[2] Remote Sensing Group, IGG, University of Bonn, Nussallee 15, 53115 Bonn, Germany; ribana.roscher@uni-bonn.de
[3] Swiss Data Science Center, ETH Zürich, Universitätstrasse 25, 8006 Zürich, Switzerland; michele.volpi@sdsc.ethz.ch
[4] Crop and Environmental Sciences Department, Harper Adams University, Edgmond, Newport TF10 8NB, UK; fveronesi@harper-adams.ac.uk
* Correspondence: jan.wegner@geod.baug.ethz.ch

Abstract: Advances in machine learning research are pushing the limits of geographical information sciences (GIScience) by offering accurate procedures to analyze small-to-big GeoData. This Special Issue groups together six original contributions in the field of GeoData-driven GIScience that focus mainly on three different areas: extraction of semantic information from satellite imagery, image recommendation, and map generalization. Different technical approaches are chosen for each sub-topic, from deep learning to latent topic models.

Keywords: geospatial machine learning; big data; classification; remote sensing; GIS; GIScience

---

Machine learning is revolutionizing digital and data-intensive disciplines by offering tools to analyze and extract valuable information from very large quantities of unstructured data. Geographical Information Sciences (GIScience) in general are no less affected by such change, and are among the disciplines that could benefit the most from tailored machine learning solutions. Modern GIScience research is characterized by very large and unstructured sources of geolocated data, from which it is often required to extract high level information in the form of spatial semantics, spatial object relationships, trajectories, or more generally, numeric tags associated to objects embedded in geographical coordinates.

Modern machine and deep learning together with the rapid development of hardware and open source software libraries allows us to approach geospatial applications that were beyond reach a few years ago. Methods must scale to massive amounts of data, for both training and prediction steps. In parallel, the increasing availability of cloud computing services and affordable graphics processing units (GPUs) eases accessibility to huge computing power. In light of these trends, the geospatial community is moving quickly towards data-driven (deep) machine and learning tools to solve challenging open research questions.

This Special Issue assembles six novel contributions in different areas of GeoData-driven machine learning. Topics span different disciplines of GIScience: generation of street address from satellite imagery [1], land-cover classification of polarimetric Synthetic Aperture Radar (PolSAR) images [2], extraction of buildings from maps to perform generalization [3], land-cover classification from satellite image time series [4], automatic selection of buildings based on cartographic constraints [5] and satellite image retrieval and recommendation [6].

In addition to tackling different applications, the papers employ a variety of machine learning tools to accomplish their goals. Deep learning is used in two works [1,4]: in the former, a convolutional neural

network extracts roads from satellite images as an initial step to generate street addresses, while in the latter, a sequential convolutional recurrent neural network provides robust end-to-end land-cover and land-use mapping from satellite image time series. Decision trees in the form of single decision trees [3] or random forests [2] are used for building detection in the first and land-cover classification in the second contribution. Lee, et al. [3] also test other standard machine learning approaches such as support vector machines, *k*-nearest-neighbor and naïve Bayes classification to explore which methodologies are most appropriate for map generalization purposes. Map generalization is also at the core of [5], who approach building selection with genetic algorithms that are constrained by cartographic and contextual knowledge. Finally, Zhang, et al. [6] use a latent topic model including space and time to retrieve and recommend remote sensing imagery.

From these contributions it is apparent that machine learning offers solutions that are highly competitive when compared to humans. Although not yet rivalling the human performance in terms of accuracy on several applications, such methods allow us to process and interpret very large quantities of data, which would be impossible to treat manually. For instance, the address generation methods in Demir, et al. [1] could be scaled at a global level or the method presented in Russwurm and Körner [4] could be extended to very long time series and large geographical areas with minimal user intervention. Modern applications of machine learning should always allow for an open door for flexible inclusion of more data or deployment to large scale datasets.

We currently observe that machine learning tools are quickly becoming standard for analyzing geospatial data. Widespread use of machine learning across a large variety of disciplines fosters collaborative research efforts. Even colleagues new to the field can quickly learn and apply machine learning due to many well-designed free tutorials and access to open source software libraries on the web. In order to further boost research in machine learning applied to GIScience, there is a need to share data and source code publicly, to design and maintain convincing benchmarking activities, and to validate methods quantitatively on open datasets of realistic (large) size.

Author Contributions: All authors contributed equally to the writing of the editorial.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Demir, I.; Hughes, F.; Raj, A.; Dhruv, K.; Muddala, S.M.; Garg, S.; Doo, B.; Raskar, R. Generative street addresses from satellite imagery. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 84. [CrossRef]
2. Hänsch, R.; Hellwich, O. Classification of PolSAR images by stacked random forests. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 74. [CrossRef]
3. Lee, J.; Jang, H.; Yang, J.; Yu, K. Machine learning classification of buildings for map generalization. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 309. [CrossRef]
4. Russwurm, M.; Körner, M. Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 129. [CrossRef]
5. Wang, L.; Guo, Q.; Lui, Y.; Sun, Y.; Wei, Z. Contextual building selection based on a genetic algorithm in map generalization. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 271. [CrossRef]
6. Zhang, X.; Chen, D.; Liu, J. A space-time periodic task model for recommendation of remote sensing images. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 40. [CrossRef]