

EGU22-4300

EGU General Assembly 2022

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



Synergetic use of Sentinel-1 and Sentinel-2 data for large-scale Land Use/Land Cover Mapping

Melanie Brandmeier^{1,2}, Maximilian Hell¹, Eya Cherif³, and Andreas Nüchter⁴¹University of Applied Sciences Würzburg, Germany (melanie.brandmeier@fhws.de)²Esri Deutschland GmbH³University of Leipzig⁴Julius-Maximilian-University Würzburg

One of the largest threats to the vast ecosystem of the Brazilian Amazon Forest is deforestation and forest degradation caused by human activity. The possibility to continuously monitor these degradation events has recently become more feasible through the use of freely available satellite remote sensing data and machine learning algorithms suited for big datasets.

A fundamental challenge of such large-scale monitoring tasks is the automatic generation of reliable and correct land use and land cover (LULC) maps. This is achieved by the development of robust deep learning models that generalize well on new data. However, these approaches require large amounts of labeled training data. We use the latest results of the MapBiomass project as the 'ground-truth' for developing new algorithms. In this project, Souza et al. [1] used yearly composites of USGS Landsat imagery to classify the LULC for the whole of Brazil. The latest iteration of their work became available for the years 1985–2020 as Collection 6 (<https://mapbiomas.org>). However, this reference data cannot be considered real ground truth, as it is itself generated from machine learning models and therefore requires novel approaches suited to overcome such problems of weakly supervised learning.

As tropical regions are often covered by clouds, radar data is better suited for continuous mapping than optical imagery, due to its cloud-penetrating capabilities. In a preliminary study, we combined data from ESA's Sentinel-1 (radar) and Sentinel-2 (multispectral) missions for developing algorithms suited to act on multi-modal and -temporal data to obtain accurate LULC maps. The best performing proposed deep learning network, DeepForestM2, employed a seven-month radar time series combined with a single optical scene. This model configuration reached an overall accuracy of 75.0% on independent test data. A state-of-the-art (SotA) DeepLab model, trained on the very same data, reached an overall accuracy of 69.9%.

Currently, we are further developing this approach of fusing multi-modal data with a temporal aspect to improve on LULC classification. Larger amounts of more recent data, both Sentinel-1 and Sentinel-2 from 2020 are included in training experiments. Additional deep learning networks and approaches to deal with weakly supervised [2] learning are developed and tested on the data. The need for the weakly supervised methods arises from the reference data, which is both inaccurate and inexact, i.e., has a coarser spatial resolution than the training data. We aim to improve the

classification results qualitatively, as well as quantitatively compared to SotA methods, especially with respect to generalizing well on new datasets. The resulting deep learning methods, together with the trained weights, will also be made accessible through a geoprocessing tool in Esri's ArcGIS Pro for users without coding background.

- Carlos M. Souza et al. "Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine". en. In: *Remote Sensing* 17 (Jan. 2020). Number: 17 Publisher: Multidisciplinary Digital Publishing Institute, p. 2735. DOI: 10.3390/rs12172735.
- Zhi-Hua Zhou. "A brief introduction to weakly supervised learning". In: *National Science Review* 5.1 (Jan. 2018), pp. 44–53. ISSN: 2095-5138. DOI: 10.1093/nsr/nwx106.