Large-scale Mapping of Arctic Coastal Infrastructure using Copernicus Sentinel Data and Machine Learning and Deep Learning Methods

Pointner, G.; Bartsch, A.; Ingeman-Nielsen, T.

Link to article, DOI:
10.5281/zenodo.3826154

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
The climate change induced increased warming of the Arctic is leading to an accelerated thawing of permafrost, which can cause ground subsidence. In consequence, buildings and other infrastructure of local settlements are endangered from destabilization and collapsing in many Arctic regions. The increase of the exploitation of Arctic natural resources has led to the establishment of large industrial infrastructures that are at risk likewise. Most of the human activity in the Arctic is located near permafrost coasts. The thawing of coastal permafrost additionally leads to coastal erosion, which makes Arctic coastal settlements even more vulnerable.

The European Union (EU) Horizon2020 project "Nunataryuk" aims to assess the impacts of thawing land, coast and subsea permafrost on the climate and on local communities in the Arctic. One task of the project is to determine the impacts of permafrost thaw on coastal Arctic infrastructures and to provide appropriate adaptation and mitigation strategies. For that purpose, a circumpolar account of infrastructure is needed.

We present an automated workflow for downloading, processing and classifying Sentinel-2 (optical) and Sentinel-1 (Synthetic Aperture Radar) data in order to map coastal infrastructure with circumpolar extent, developed on a highly performant virtual machine (VM) provided by the Copernicus Research and User Support (RUS). We further assess the first classification results mapped with two different methods, one being a pixel-based classification using a Gradient Boosting Machine and the other being a windowed semantic segmentation approach using the deep-learning framework Keras.

**Data & Methods**

- Sentinel-2 data is a common choice for land cover mapping. Most land cover products only include one class for built-up areas, however. The fusion of optical and Synthetic Aperture Radar (SAR) data for land cover mapping has gained more and more attention over the last years. By combining Sentinel-2 and Sentinel-1 SAR data, the classification of multiple types of infrastructure can be anticipated. Another emerging trend in the application machine learning and deep learning methods for land cover mapping.
- Preliminary steps:
  - Sentinel-1 IW GRO: border-noise removal, slice-assembly, radiometric calibration, thermal noise removal, terrain correction, incidence angle normalization
  - Sentinel-2 L1C: atmospheric correction, super-resolution, cloud masking, merging of individual scenes, calculation of spectral indices
- Resulting images were stacked and two different approaches are tested for classification with varying polarization combinations:
  - Machine Learning: pixel-based classification using a Gradient Boosting Machine (GBM) from XGBoost (Chen and Guestrin 2016)
  - Deep learning (DL): a windowed semantic segmentation approach using the deep-learning framework Keras (Chollet et al., 2015).
- Calibration data: OpenStreetmap for Deep Learning, Manually collected samples selected with respect to Google Earth imagery for the XGBoost classifier
- Validation data:
  - Prudhoe Bay Cumulative Impact Map, Alaskan North Slope (Raynolds et al. 2014)
  - Dedicated observations within NUNATARYUK for several settlements on Greenland (Ingeman-Nielsen et al. 2018) and for Longyearbyen (Lu et al. 2018) - Vector geometries thoroughly annotated with information on infrastructure types and materials

**Preliminary Results**

- Prudhoe Bay background mostly tundra, some river beds only Sentinel-1 VV/VH available
- Ilulissat background mostly bedrock, only patches of vegetation only Sentinel-1 HH/HV available

**Summary and future work**

- Both classification approaches provide useful information but differ strongly in error of omission and error of commission. GBM is considered slightly better in terms of error of omission, DL better in terms of errors of commission (Bartsch et al., in preparation)
- Both classification approaches require manual post processing in order to provide a useful data product
- A method that incorporates a combination of both approaches may be able to make use of the respective advantages of the presented methods
- Detailed validation datasets will enable assessment of classification results with respect to type and material of infrastructure objects

---

**References:**


Tianhong Chen and Colén Grönqvist (2017). XGBoost in Proceedings of the 32nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

François Chollet et al. (2015). Keras http://keras.io

Tianhong Chen and Colén Grönqvist (2017). XGBoost in Proceedings of the 32nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Tianhong Chen and Colén Grönqvist (2017). XGBoost in Proceedings of the 32nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.