

AI-vergreens - a new multi-level labelled multi-temporal Sentinel-2 image patch-based training dataset optimized for Northern Circumboreal forest

Wednesday 16 October 2024 14:50 (2 minutes)

Circumboreal forests are located primarily in Alaska, Canada, and Northern Eurasia representing close to 30% of all forested land areas and are strongly changing in response to climate and increasingly frequent disturbances such as fires and drought. Remote sensing applications for landcover in high latitudes are possible but remain challenging for optical satellite sensors due to long lasting snow coverage and frequent cloud coverage and wildfires. In addition, the accuracy of landcover mapping strongly depends on the quality and the correct context of the training data. Forest inventories are commonly used as training data for forest remote sensing. For the Northern boreal realm, specifically the subarctic treeline ecotone, publicly available reference data remain rare and are even not adequately labeled for machine learning and remote sensing-related applications. In our previous work SiDroForest (Siberian Drone Mapped Forest Inventory), we produced a training dataset containing forest-type labeled Sentinel-2 image patches linked to AWI expedition plots in the Eastern Siberian summergreen-evergreen and treeline transition zones (van Geffen et al., 2021, <https://doi.org/10.1594/PANGAEA.933268>, van Geffen et al. 2022). The SiDroForest forest-type labels despite linked to 54 reference plots contain as real-world reference data in some cases only few members per class and are imbalanced. As forest landcover is characterized by high heterogeneity due to topographic and disturbance constraints, measuring forest parameters requires to cover a large variation across environmental gradients.

In Enguehard et al. 2024a,b <https://doi.org/10.1594/PANGAEA.964699>, within the frame of the AI-vergreens BMWK project we could extend the Sentinel-2 training dataset for Eastern Siberia to 79 labelled patches including more expedition data and photointerpretation to expand on the rare classes. Incorporating now the reference data from new expeditions to Alaska and NW Canada since 2022, we are currently building up an extensive labeled training dataset based on multi-temporal optical Sentinel-2 satellite data for Northern Circumboreal forest remote sensing applications. We assign multi-level labels based on our field data on tree species and crown cover percentages. Additionally, we derive the tree crown cover from LiDAR 3-D point clouds (YellowScan Mapper on M300 DJI drone), and structure from motion from RGB (DJI Phantom 4).

In addition, based on the extensive UAV data collection from expeditions in Siberia, Alaska and Canada, Kruse et al. (in prep.) are preparing a forest structure training data set derived from LiDAR 3-D point clouds from AWI expeditions, optimized for machine learning to detect individual trees and species in northern boreal forests and the tundra transition (BorFIT, AWI DataHub Information Infrastructure funds). We will be able to extend our image patch collection using BorFIT structural forest information extracting reference data from larger forest areas in addition to the AWI expedition plots.

We anticipate our dataset to be a starting point for a significantly more extensive one with the addition of SAR sensor image patches (Sentinel-1, TanDEM-X). Based on field data and UAV (multispectral, hyperspectral and LiDAR) we can provide the contextual understanding, detail and specificity for a vast region where publicly available reference data remain scarce. We will make the consistently and uniformly labelled patch-based training dataset optimized for machine-learning publicly available as FAIR data publication including a WebGIS visualization of the location of the training data in the Circum-Boreal as part of the AI-vergreens BMWK project.

Primary author: HEIM, Birgit (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme)

Co-authors: Prof. DEMIR, Begüm (Technische Universität Berlin, Remote Sensing Image Analysis Group, Faculty of Electrical Engineering and Computer Science); Ms VAN GEFFEN, Femke (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme); Mr SCHLADEBACH, Jacob (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Alfred-Wegener-Institut Helmholtz-Zentrum für Po-

lar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme); PAPATHANASSIOU, Konstantinos (Deutsches Zentrum für Luft- und Raumfahrt, Institut für Hochfrequenztechnik und Radarsysteme); Mrs ENGUEHARD, Lea (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme); Dr WIECZOREK, Mareike (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme); Dr HÄNSCH, Ronny (Deutsches Zentrum für Luft- und Raumfahrt, Institut für Hochfrequenztechnik und Radarsysteme); Dr KRUSE, Stefan (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme); Prof. HERZSCHUH, Ulrike (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme)

Presenter: HEIM, Birgit (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Sektion Polare Terrestrische Umweltsysteme)

Session Classification: Posters