Knowledge Extracted from Copernicus Satellite Data

Corneliu Octavian Dumitru[1], Gottfried Schwarz[1], Torbjørn Eltoft[2], Thomas Kræmer[2], Penelope Wagner[3], Nick Hughes[3], David Arthurs[4], Andrew Fleming[5], Manolis Koubarakis[6], and Mihai Datcu[1]

[1]Remote Sensing Technology Institute, German Aerospace Center, Wessling 82234, Germany (email: corneliu.dumitru@dlr.de; gottfried.schwarz@dlr.de; mihai.datcu@dlr.de)

[2]The Arctic University of Norway, Hansine Hansens veg 18, 9019 Tromsø, Norway (email: torbjorn.eltoft@uit.no; thomas.kramer@uit.no)

[3]Norwegian Meteorological Institute, Kirkegårdsveien 60, P.O. Box 6314 Langnes, 9293 Tromsø, Norway (email: penelopew@met.no, nick.hughes@met.no)

[4]Polar View Earth Observation Ltd, Fermi Avenue Harwell Oxford, Didcot Oxfordshire OX11 0QR, UK (email: david.arthurs@polarview.org)

[5]United Kingdom Research and Innovation, British Antarctic Survey, High Cross Madingley Road, Cambridge, UK (email: ahf@bas.ac.uk)

[6]National and Kapodistrian University of Athens, Athens 157 72, Greece (email: koubarak@di.uoa.gr)

Abstract

ExtremeEarth is a European H2020 project; it aims at developing analytics techniques and technologies that combine Copernicus satellite data with information and knowledge extraction, and exploiting them on ESA's Food Security and Polar Thematic Exploitation Platforms.

In this publication, we focus on the Polar case which requires the selection of validation areas, the generation of a training dataset, the development and testing of deep learning algorithms, and the demonstration of regional results.

During the development of deep learning algorithms, a key activity is to establish a large amount of referenced Earth Observation data. They need to be sufficiently diverse to cover the major target areas of satellite images under varying imaging conditions and across all seasons. For doing this, we propose to select overlapping target areas from Synthetic Aperture Radar and multispectral images acquired with rapid succession. Such a combination approach already demonstrated its applicability for monitoring seasonal snow cover (https://ieeexplore.ieee.org/document/8518203).

By applying an already established active learning approach based on a Support Vector Machine with relevance feedback (http://wiki.services.eoportal.org/tiki-index.php?page=EOLIB+Project), we can limit ourselves to a limited number of typical satellite images to extract their information content, and to generate semantic annotations for them.

This approach is also a simple way to generate benchmarking datasets that can be used for testing and validating different algorithms, and for creating additional bigger datasets for large-scale demonstrations. The proposed methodology uses new paradigms from Recurrent Neural Networks and Generative Adversarial Networks, supported by Bayesian and Information Bottleneck concepts.

Keywords

Satellite Data, Knowledge Extraction, Classification