Combining Remote Sensing with Webdata and Machine Learning to Support Humanitarian Relief Work

Jens Kersten, Malin Kopitzsch, Jan Bongard, and Friederike Klan
German Aerospace Center (DLR), Institute of Data Science, Citizen Science, Jena, Germany (jens.kersten@dlr.de)

Gathering, analyzing and disseminating up-to-date information related to incidents and disasters is key to disaster management and relief. Satellite imagery, geo-information, and in-situ data are the mainly used information sources to support decision making. However, limitations in data timeliness as well as in spatial and temporal resolution lead to systematic information gaps in current well-established satellite-based workflows. Citizen observations spread through social media channels, like Twitter, as well as freely available webdata, like WikiData or the GDELT database, are promising complementary sources of relevant information that might be utilized to fill these information gaps and to support in-situ data acquisition. Practical examples for this are impact assessments based on social media eyewitness reports, and the utilization of this information for the early tasking of satellite or drone-based image acquisitions.

The great potential, for instance of social media data analysis in crisis response, was investigated and demonstrated in various related research works. However, the barriers of utilizing webdata and appropriate information extraction methods for decision support in real-world scenarios are still high, for instance due to information overload, varying surrounding conditions, or issues related to limited field work infrastructures, trustworthiness, and legal aspects.

Within the current DLR research project "Data4Human", demand driven data services for humanitarian aid are developed. Among others, one project goal is to investigate the practical benefit of augmenting existing workflows of the involved partners (German Red Cross, World Food Programme, and Humanitarian Open Street Map) with social media (Twitter) and real-time global event database (GDELT) data. In this contribution, the general concepts, ideas and corresponding methods for webdata analysis are presented. State-of-the-art deep learning models are utilized to filter, classify and cluster the data to automatically identify potentially crisis-related data, to assess impacts, and to summarize and characterize the course of events, respectively. We present first practical findings and analysis results for the 2019 cyclones Idai and Kenneth.