Leveraging Overhead Imagery Capabilities in the Nonprofit Sector through Analytics-as-a-Service and Machine Learning

by Pia Ulrich

The availability of commercial satellite imagery to the Non-Governmental (NGO) Nonproliferation world undoubtedly presents an opportunity to advance their work through satellite imagery analysis. At the same time, it becomes increasingly clear that the images by themselves will only do so much. While the availability and accessibility of satellite imagery support analysts' work, the key to and value of satellite imagery's importance is its proper analysis. Geospatial experts emphasize that what the users want from satellite imagery is synthesis — the pieces of the puzzle put in context. However, without the skills necessary to analyze and interpret the data, users remain; "data rich, information poor."

Using overhead imagery effectively requires not only specialized supporting software but also geospatial expertise. From choosing between the best types and resolutions of imagery data for the intended illustration, to the need to properly process the raw data, it is essential that users have the expertise to analyze the imagery and interpret its data. Imagery analysis and interpretation have a great margin for error, faulty analysis and interpretation. Conclusions based on flawed analysis or interpretation can easily tarnish the credibility of those relying on it.

The challenge for even skilled imagery analysts is the complex nature of working with satellite data and the lack of standardization of data. However, these challenges pale in comparison to the information overload resulting from the huge volumes of data collected by satellites. The American commercial vendor of space imagery and geospatial content, DigitalGlobe, alone collects about 73 terabytes of satellite images every day — on top of a 16-year running archive.[1] The amount of time necessary to process imagery, analyze, and interpret it — often by repeating the same processing steps for the same data over and over again — results in inefficiencies and poses limitations that can negate the unique utility of satellite data.

Fortunately, analytic methods to assess large amounts of satellite data have made significant advances in recent years, and with the help of algorithms and machine learning, analysis can be produced with greater efficiency and at reduced cost. A game-changer in maximizing the value of satellite data for end users, however, is the availability of analysis-as-a-service (sometimes also called analytics-as-a-service), or A3S. Companies providing data analytics have created a new market by providing improved access to satellite data, often in the form of real-time insights. By transforming satellite data into (accessible) information, users with less expertise or infrastructure can effectively use satellite imagery.

One form of these A3S services is Analysis Ready Data (ARD), which is defined by The Committee on Earth Observation Satellites (CEOS) as: "satellite data that have been processed to a minimum set of

[1] Steven Melendez, *The Eye in the Sky Gets a Brain that Knows What It's Seeing*, January 16, 2017, Fast Company, https://www.fastcompany.com/3066622/the-eye-in-the-sky-gets-a-brain-that-knows-what-its-seeing.

requirements and organized into a form that allows immediate analysis without additional user effort and interoperability with other datasets both through time and space."[2]

Other forms of services exist as well, with companies offering full satellite data acquisition services that include an assessment of the satellite image data needs specific to the customer's application. Offers include matches in the spatial, textural, and spectral resolutions, customized data collections, and can merge all of it with archived data. Advanced image processing techniques allow A3S providers to extract valuable information and to integrate geospatial information acquired with different spatial and spectral resolutions. The result is fused data that greatly exceeds the value of data from each of the individual sources. Other services go beyond simple technical processing of satellite images and provide customers full models and analysis of spatial data, embedded in Geographic Information Systems (GIS), or fused with information from other sources, such as news and social media.

Currently, demand for A3S is increasing across all sectors. However, its biggest limitation is the lack of an established market and the small number of specialist organizations currently developing the market. Solution development is often supplier-driven as opposed to market-led. Nonetheless, these services paired with the availability of growing amounts of satellite imagery change the dynamics of the field significantly. With that, capabilities available to the commercial sector — including the nonproliferation NGO world — include new levels of persistence that were never before available, direct tasking, advanced analytics, auto-tracking, alerts, change detection, and activity-based intelligence.

Even the U.S. government is still trying to better understand the capabilities that are developing so rapidly. Efforts aimed at understanding the value of, and how to leverage and apply, next-generation commercial imagery to government needs are underway. However, the dangers and flaws of A3S still need to be explored. A primary concern remains the lack of quality control and the incorporation of inherent and intentional bias, or incoherence, in the services provided.

Artificial Intelligence and Machine Learning: Information, Not Data

The ability to offer data analysis-as-a-service to the customer is, to a large part, a consequence of the application of artificial intelligence (AI) to the technology and underlying data science. The combination of computer vision and machine learning tools allows automation of the process of analyzing satellite image data to extract meaningful information from the data. A growing cast of Silicon Valley companies, among them Facebook, is working on perfecting one particular form of AI: machine learning. By training computers to learn a level of subject matter expertise that is comparable to that of a human, machines are learning to detect even contextual clues. Experts agree that the need for machine learning has increased and will play an increasingly important role in extracting information from satellite data. In a Belfer Center study, Greg Allen and Taniel Chan describe the potential of existing machine learning

---

[2] CEOS Plenary 2016, presentation p. 5, https://media.sa.catapult.org.uk/wp-content/uploads/2017/03/11101352/An-update-on-CEOS-and-CARD4L-NCEO.pdf.

technology that "could enable high degrees of automation in labor-intensive activities such as satellite imagery […]."[3]

Some techniques are modeled after the learning that the popular search engine Google displays in learning about its users and their preferences, and adapting search results in accordance with those. As a result, imagery analysis products with machine learning will watch their users, make suggestions, optimize, and ultimately learn. DigitalGlobe's Space Net Initiative, on the other hand, learns by being fed pre-labeled overhead imagery models; then, DigitalGlobe's GBDX platform trains computers to automatically detect relevant objects in satellite images. Large sets of data are used to develop algorithms that classify images, audio, or text, resulting in a system that can, without human intervention, accurately determine and analyze points of interest in the photos. Crowdsourcing networks help with the training and testing of the machine-learning algorithms.

Spun out of Los Alamos National Lab (LANL), startup Descartes Lab provides AI-driven analysis and recently released a public demo of its GeoVisual Search. This search engines allows its user to choose an object anywhere on the globe, as seen from space, and matches it with objects similar in appearance anywhere on the planet. The search engine combines satellite images with machine learning at an unprecedented scale. The company is working on automated counting of objects, with the idea of turning satellite imagery into a searchable database with real-time information. Ultimately, Descartes Lab aims at uncovering patterns that tie together all activities in a given pictures that might be missed by the human eye. Currently, the algorithms are being tested on videos, resulting in object detection within video content faster than real time.

Observera's Electro-Optical Change Detection (EOCD) software is the first fully automated processing capability to work with panchromatic imagery, producing reliable detections, highlighting changes, and identifying second and third order indicators, thus saving analysts time and catching those changes that the analysts might not even have noticed.

However, even industry experts recognize that, so far, it has been challenging to use the vast amount of location data available to answer questions. The large potential for business opportunities drives innovation and fuels machine learning. After all, the application of AI in data analysis applications allows for automatic analysis, predicative analysis, trends identification, and anomaly detection.

The development of computer vision and machine learning tools that are required to automate the process are just beginning. According to the co-founder and CEO of Spaceknow, a Silicon Valley data analytics firm that works with commercial satellite data, the convergence of computing power, machine learning, and satellite imagery makes for a perfect storm that is just now beginning to peak.[4] At the same time, these developments can and will be used to generate forgeries and discern the truth, creating an even greater need for independent, unbiased monitoring and verification.

---

[3] Greg Allen, Taniel Chan, *Artificial Intelligence and National Security*, Belfer Center Study, July 2017, p.1.

[4] Clay Dillow, *What Happens When You Combine Artificial Intelligence and Satellite Imagery*, Fortune, 30 March 2016; http://fortune.com/2016/03/30/facebook-ai-satellite-imagery/.