Microsoft

AI for Earth Grantee Profile

Digamma.ai

Classifying land cover to improve maps of landslide susceptibility, water, and carbon storage

Summary

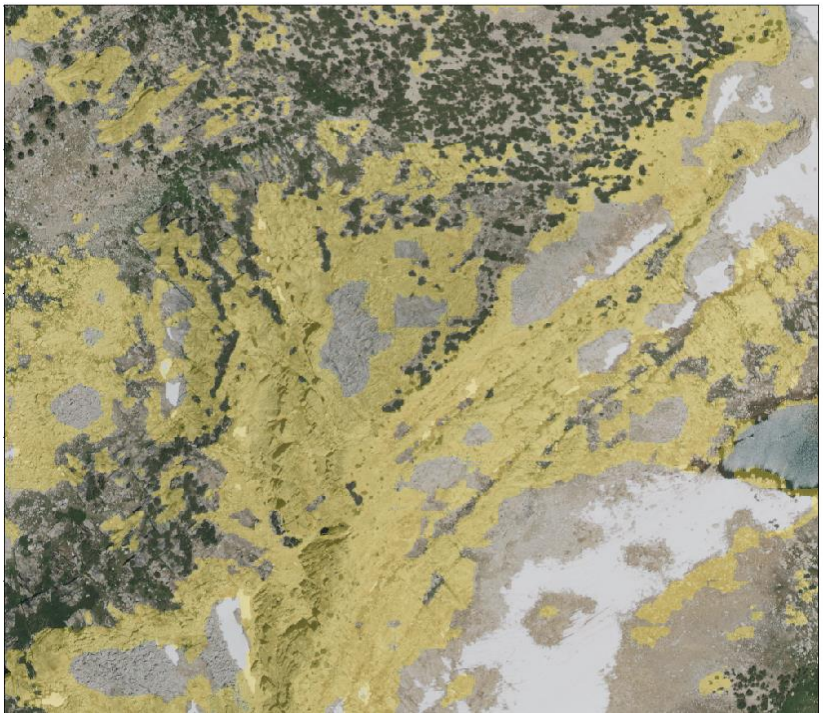
Due to technological advances, scientists can now capture data on Earth and off-world at rates that greatly exceed the ability to interpret it. These massive datasets challenge the delivery of timely maps and analysis to the nation. Artificial intelligence (AI) tools, including machine learning, allow us to rapidly interpret these data sets to solve national challenges. Digamma.ai and the U.S. Geological Survey's National Innovation Center (USGS NIC) are using machine learning to dramatically improve land cover models, with the intent of improving maps of landslide susceptibility, water, and carbon storage. In addressing this challenge, Digamma.ai used geologic mapping and USDA imagery to train machine learning models to discriminate between bare rock and exposed soil, improving land-cover maps across the Sierra Nevada in California.

Improving the Sierra Nevada land-cover map through AI

Land-cover maps are used to inform models that calculate the storage of water, carbon, and other nutrients, as well as the susceptibility of the land to mass movements like landslides. However, existing land-cover maps (such as those from USGS National Land Cover Database or Natural Resources Conservation Service) are typically based on imagery with lower resolution (10 to 100 meters) or are intended to portray low-relief agricultural areas.

> Distinguishing between rock and other barren terrain, especially in natural, non-urban environments, can be used for resource, erosion, or hazard modeling.

Global, sub-meter imagery presents an opportunity to improve the resolution and accuracy of land-cover maps. Publicly available Earth surface imagery now makes it easier to find data for almost all places in the world, often on a meter or higher resolution, but the interpretation of these images remains challenging.



Processed land cover map—yellow areas are bare rock [Image courtesy Digamma.ai]

Digamma.ai and the U.S. Geological Survey (USGS) are using machine learning approaches and resources provided by Microsoft AI for Earth to increase the quality of land-cover maps.

Digamma.ai, the artificial intelligence and machine learning branch of Codeminders, is focused specifically on improving mapping of bare rock surfaces, as this specificity is currently lacking or combined into other classes in existing surface cover datasets. Distinguishing between bedrock and other barren terrain, especially in natural, non-urban environments, can be used for resource, erosion, or hazard modeling. For example, the USGS is working to improve landslide maps statewide in California, particularly for places like the Sierra Nevada Mountains where fast-moving, damaging landslides can occur during heavy rainfall. These types of landslides only occur in soil-covered areas and masking out areas without soil improves the ability of agencies to focus resources.

Mapping bedrock with deep learning

Digamma.ai is exploring the utility of using wider-coverage, sub-meter scale aerial imagery from the USDA's 2016 NAIP collection. This four-band (red, green, blue, near-infrared), publicly available imagery covers all of California with cloud-free, 0.6-meter pixels. With this availability, this workflow could potentially be expanded statewide in the future, or at least across the entire Sierra Nevada range. Accurate classification of such heterogeneous landscapes is difficult and relies on sufficient high-quality labels in the training data. Using the KNN matting technique, Digamma.ai's algorithm significantly decreases the effort needed to perform high-quality ground-truth labeling.

Among recent land-cover classification approaches, Digamma.ai explored advances in image-based classification models and built a nine-layer convolutional neural network (CNN) to classify the likelihood of each pixel being labeled as rock by analyzing that central pixel and its vicinity. This method offers a significant improvement over what is currently available for barren or rock classification, as well as classical approaches. This initial model was tested on a relatively small patch of the Sierra Nevada, in which each image

corresponded to approximately 72 square kilometers, and it is very important to extend the application to a wider area. Approximately 2,000 images are needed to cover the whole Sierra Nevada range, requiring significant computing resources.

Microsoft Azure provides cloud computing resources to make that scalability possible. Through a grant from the AI for Earth program, Digamma.ai has access to Azure, which provides both the computational power and data platform. Another major benefit of switching to Azure is the ready availability of the NAIP dataset which Digamma.ai used for the project. Previously, each prediction took 30 minutes per image, which is time-consuming over a large area. The Azure Data Science Virtual Machine gave Digamma.ai the opportunity to optimize the image classification pipeline, which previously couldn't be done due to computational limitations. The company replaced its CNN classification model with a new segmentation model based on the U-Net architecture, which has many more parameters than the previous model. Each prediction now takes 4 minutes per image.

Going forward

USGS and Digamma.ai will release the resulting land-cover map of the Sierra Nevada range as a publicly available resource. This map could be used by the Earth science community to improve shallow landslide susceptibility models, and improve estimates of water and carbon storage. Digamma.ai also plans to finish developing an internal tool for labeling satellite or aerial imagery and release it as an open-source tool. Training data can be mapped more quickly and include some mixed pixels.

About Digamma.ai

Digamma.ai is an artificial intelligence and machine learning company from California. The group is made up of everyone from machine learning consultants, partners, and engineers to startup founders and enterprises interested in integrating AI and machine learning into their products. This proposal will expand their joint work with USGS under our existing Collaborative Research and Development Agreement (CRADA-18-6190).

About USGS National Innovation Center

The National Innovation Center (NIC) connects USGS to new technology opportunities through partnerships with industry, academia, non-governmental organizations, and other federal agencies. Its partnership building capacity enables NIC to align science capability goals across multiple USGS programs and external entities, using scarce resources to fill important national observation gaps.

Resources

Primary Contact

Helen Petliak, project principal investigator
lpetlyak@digamma.ai
(General enquiries: hello@digamma.ai)

Websites

Digamma.ai home site
USGS home site

Publications

Petliak, H.; Cerovski-Darriau, C.; Zaliva, V.; Stock, J. "Where's the Rock: Using Convolutional Neural Networks to Improve Land Cover Classification." *Remote Sens*. 2019, 11, 2211. https://www.mdpi.com/2072-4292/11/19/2211