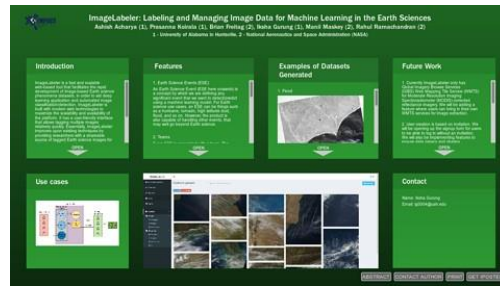


ImageLabeler: Labeling and Managing Image Data for Machine Learning in the Earth Sciences

Ashish Acharya (1), Prasanna Koirala (1), Brian Freitag (2), Iksha Gurung (1), Manil



Maskey (2), Rahul Ramachandran (2)

1 - University of Alabama in Huntsville, 2 - National Aeronautics and Space Administration (NASA)



PRESENTED AT:

INTRODUCTION

Web-based tool for the rapid development of image-based Earth science phenomena datasets

Helps build datasets for deep learning and automated image classification/detection

Built with modern web technologies to maximize the speed, scalability, and availability of the platform

Improves upon existing data management techniques by providing researchers with a shareable source of tagged Earth science images for all their machine learning needs

FEATURES

Create machine-learning-ready datasets

Team-based collaboration

Upload images or extract from the map

Create bounding boxes over images

Draw shapefiles over areas of interest

Upload and interact with GeoTIFFs

Make datasets available through APIs

Details:

Create machine-learning-ready datasets for Earth Science Events (ESE)

An Earth Science Event (ESE) is a conceptual representation of the collection of datasets (images, shapefiles, bounding boxes, etc.) which are gathered by the Subject Matter Experts (SMEs). ImageLabeler provides many features allowing users to extract and store the datasets in a machine learning ready format. Eg: images and shapefiles of hurricanes, floods, etc.

Team-based collaboration to create datasets

The ImageLabeler also allows for SMEs to bring together a team for a specific ESE and work on creating the datasets collaboratively.

Upload images or extract from the map

The ImageLabeler also allows for users to upload their own images or datasets and work on them. However, you can also create new datasets from the map view. This feature also allows users to interact with a list of satellite map layers and extract specific events to extend the ESE dataset.

Create bounding boxes over images

Users are also able to draw bounding boxes or polygons over images in the ImageLabeler. These shapes can then be downloaded to train machine learning models specially designed for object detection.

Draw shapefiles over areas of interest

Users can also draw bounding boxes or polygons in the map view. This helps retain geo-information for the extracted dataset. This dataset can then be downloaded and used for the aforementioned applications and the results can be visualized on the

map.

Upload and interact with GeoTIFFs

In some cases, the resolution of the satellite layers might not be enough for users to label and create datasets. In such cases, users can upload their own high-resolution datasets to interact with.

Make datasets available through APIs

All of the datasets are also available via API for easier programmatic access.

Video demonstration of the features:

[VIDEO]

<https://www.youtube.com/embed/w43GKBirFEU?rel=0&fs=1&modestbranding=1&rel=0&showinfo=0>

STATISTICS



ImageLabeler Statistics

Some highlights from ImageLabeler:



Flood



High Latitude Dust

Cloud Streets



All images above were extracted from MODIS satellite imagery and labeled using ImageLabeler.

FUTURE WORK

Image extraction from custom WMTS layers

Public Earth science datasets

Image extraction from GeoTIFFs

Public-facing APIs

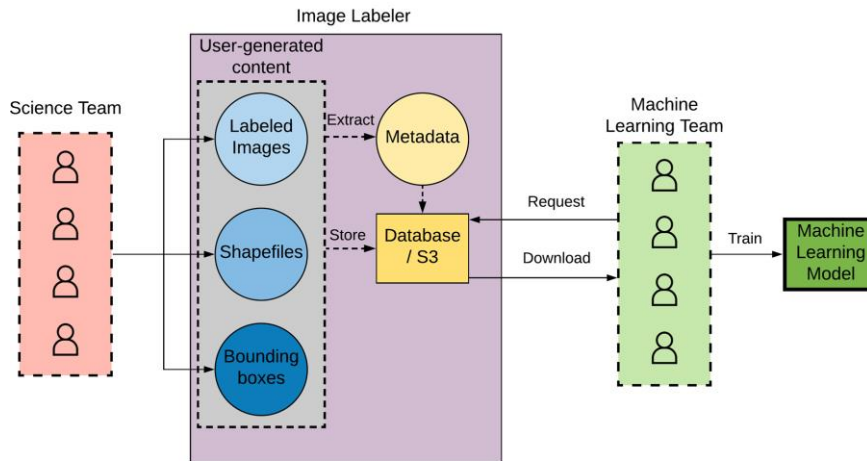
UI/UX improvements

USE CASES

Step 1: Extract images from satellite imagery or upload your own

Step 2: Label images, draw bounding boxes, create shapefiles

Step 3: Train machine learning algorithm using labeled datasets



Use case diagram

CONTACT AND ACKNOWLEDGMENT

ImageLabeler homepage

labeler.nasa-impact.net

Contact

NASA-IMPACT <teamimpact@uah.edu>

Acknowledgment

This research is supported by NASA Grant NNM11AA01A as part of the Interagency Implementation and Advanced Concepts Team (IMPACT) program

ABSTRACT

While machine learning techniques for image classification have been around for a long time, storing and managing the vast number of images required as training data is still a problem for scientists. This is especially true for the field of Earth science, where only recently have experts begun using machine learning techniques for image-based phenomena classification.

Image Labeler, a fast and scalable cloud-based tagging platform for Earth science images, seeks to improve upon existing methods of managing images and associated metadata, such as maintaining categorized folders of images on a local machine, a process that can be cumbersome and difficult to scale. The platform facilitates rapid development of image-based Earth science phenomena training datasets by allowing scientists to upload their existing imagery as well as extract new samples from open satellite imagery services made available through NASA's Global Imagery Browse Service (GIBS). Image Labeler also supports GeoTIFF data, with capabilities such as displaying GeoTIFFs on an interactive map, drawing shapefiles over them, and tagging them with additional metadata. This allows scientists to perform spatiotemporal subsetting with geographic information and develop training data more quickly.

Built using modern web technologies, Image Labeler includes additional capabilities such as team collaboration for large-scale image tagging projects. Users can download their data in a machine-learning-ready format, allowing scientists to spend time on experimentation rather than on the collection of training data. In this presentation, we demonstrate how Image Labeler seeks to become a one-stop image data management solution for machine learning applications in Earth science.