Soil Crust Analysis With Remote Sensing Data Fusion

Cade Smith and Grayson Morgan
Department of Geography, Brigham Young University

Introduction

Biological soil crusts, sometimes called cryptobiotic soil crusts, are an important part of arid and semi-arid ecosystems throughout the world. These crusts are composed primarily of microscopic organisms that cannot be seen with the naked eye. Most crusts consist of mosses, cyanobacteria, lichens, algae, and microfungi. These microorganisms form a thin layer on exposed soil surfaces. Biocrust is an important means by which arid soils resist erosion by wind and water. Biocrust also aids in the process of nitrogen fixing, meaning they chemically convert atmospheric nitrogen into a form usable by other plants.

Biological soil crusts are quite fragile, especially during drier seasons. Small amounts of pressure will break through the crust and expose the loose sand or soil beneath to the forces of erosion. As a result, humans can have a major impact on these delicate crusts. Damaged and compacted crusts can take anywhere from a few years to several decades or longer to recover. Without the crusts stabilizing the soil, erosion can eventually kill nearby plants. New plants also have a much harder time getting established without the crusts providing increased moisture and nitrogen. As plant life becomes sparse, animal life that relies on plants will also be threatened. When soil crust is destroyed over a large area, the cascading effects ripple up through the ecosystem.

Because of the fragility and necessity of these organisms, our goal was to develop a tool to map and monitor the health of biological soil crusts over time using remotely sensed imagery. This can help identify areas where soil crust has degraded over time, and facilitate the study and restoration of these important organisms,

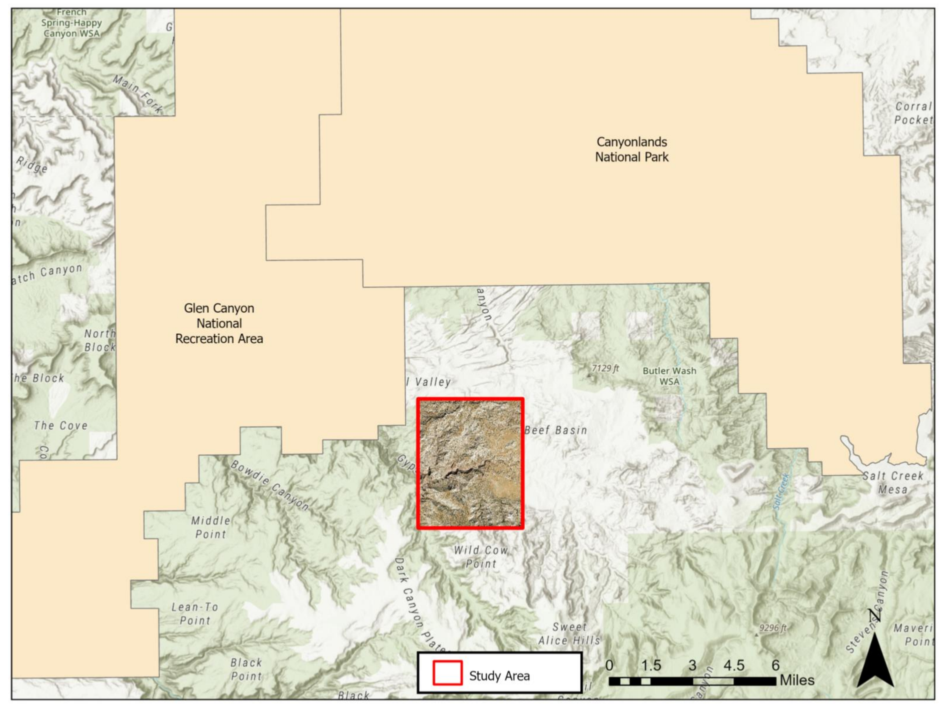

Figure 1
Location of Study Site in relation to Canyon Lands National Park and Glen Canyon National Recreation Area

Methods

We acquired data from a study done by the United States Geological Survey that contained 10 classified raster images identifying the distribution and condition of biological soil crusts using high resolution imagery from unmanned aerial systems. Classified drone images were reclassified into Light crust, dark crust, and other (including vegetation, rocks, etc). The reclassified rasters were converted to polygon data and then used to train a support vector machine and random trees classifier. Polygons larger than 1 m^2 were selected as training polygons, resulting in 1,107 training polygons across four areas. The training data were used to train multiple machine learning models. A total of 853 polygons were collected as validation polygons while undergoing the same criteria.

Two machine learning classifiers were used to compare results: random tree (forests) and support vector machine. Random trees is a machine learning method that creates conditional statements and finding constancy between the different categories. A series of decisions regarding classes and pixel values are made toward terminal nodes that result in the final characteristics of the three classes. Support Vector Machine classifier us based on the maximum margin classifier which uses a 'slab' to statistically differentiate between the different classes. Benefits of SVM include less samples required and they don't need to be normally distributed.

The trained model was applied to imagery from 2021 from the National Agriculture Imagery Program (NAIP) in order to compare the accuracy of mapping soil crusts. The study area selected was just south of Canyonlands National Park, Utah, and the training data from the USGS was acquired just south of the National Park.

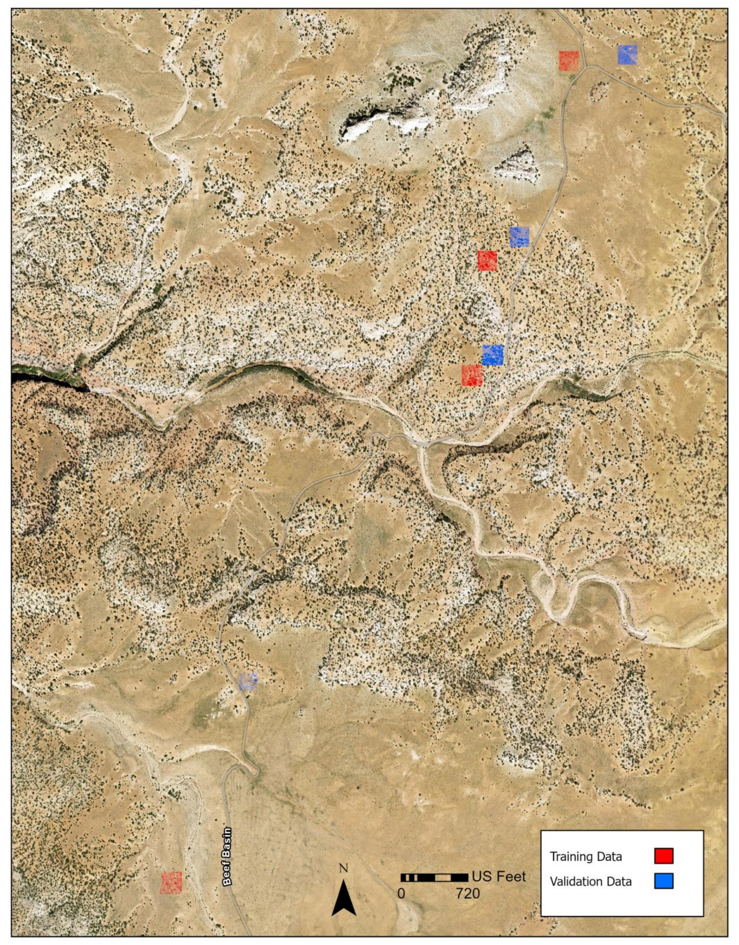

Figure 2
Location of training and validation data used for classification

Results

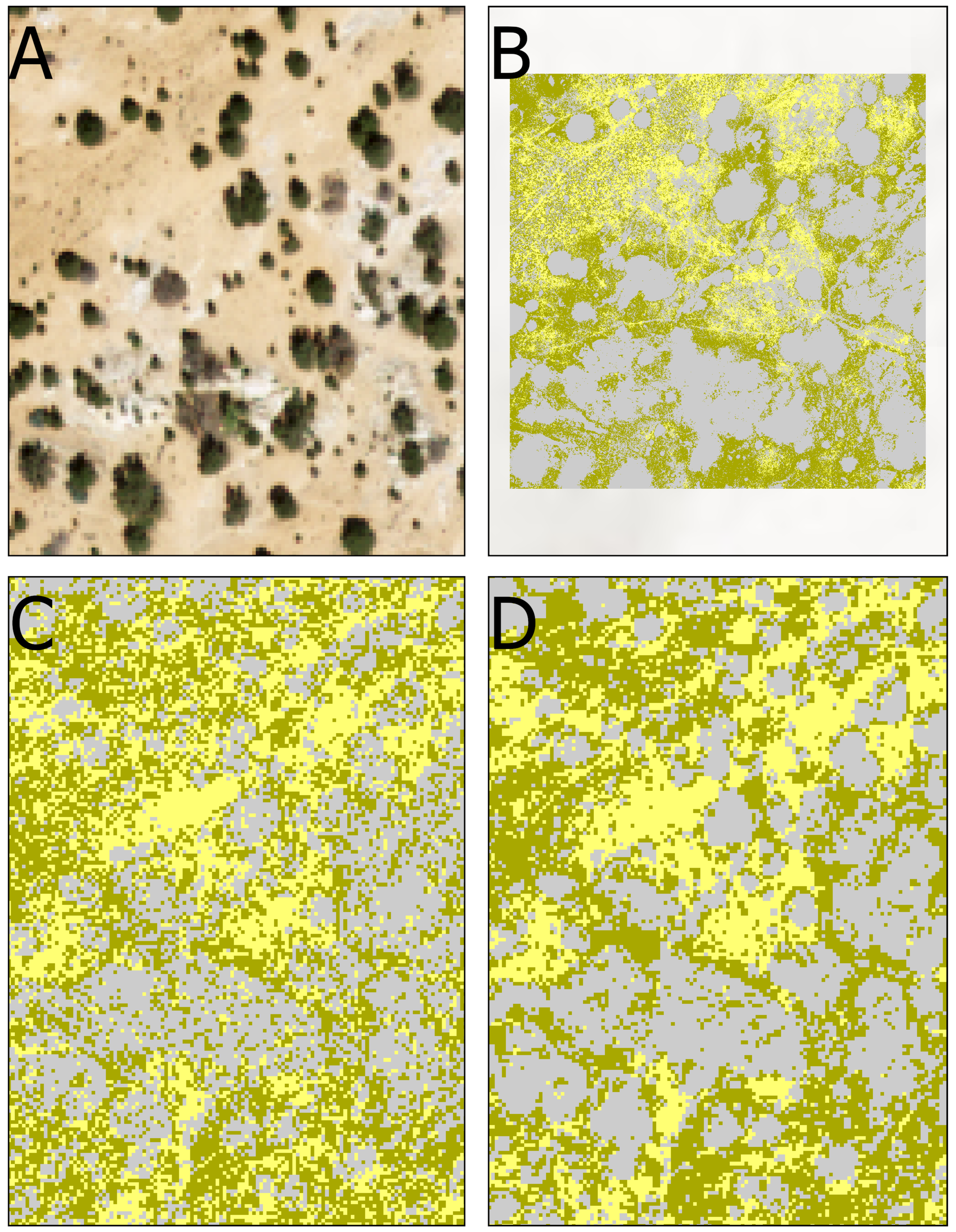

Figure 3
A) Original Image
B) Validation Dataset
C) Random Forest Classifier Results
D) Support Vector Machin Classification Results

Light Cya

Preliminary results indicate several visual similarities between the support vector machine (SVM- Figure 3D), Random Forest (RF- Figure 3C) and the validation dataset (Figure 3B). Light cyanobacteria crust is indicated in yellow, while dark/mixed biocrust is in gold. The classifications smoothed out and made drastic generalizations with larger pixel sizes, which may be part of the challenge of using drone data with larger spatial resolution aerial imagery.

Random Forest classification took less than half of the time to complete the training and classification and provided a better overall accuracy. Despite the better accuracy, the RF classified image shows more of the salt and pepper phenomenon consistent with pixel-based classifiers. For the Dark/Mixed biocrust class, both classifiers provided similar accuracy. The largest discrepancy was with the producers accuracy for the Light Cyanobacteria class (60% vs 31%).

Solid Crust (Light Cyanobacteria biocrust) Classification

Type	Time	Users accuracy	Producers accuracy
Random Forest	35:14	47.13%	60.05%
Support Vector Machine	1:15:13	47.21%	31.75%

Solid Crust (Dark/Mixed biocrust) Classification

Type	Time	Users accuracy	Producers accuracy
Random Forest	35:14	35.23%	40.57%
Support Vector Machine	1:15:13	40.31%	30.25%

Discussion and Conclusions

The random forest classifier performed the best overall. The random forest classifier is not always known for being the most complex, but in this case was sufficient to provide up to 60% classification accuracy. In other instances, SVM was more accurate for mapping land cover (Liu et al. 2017). More rigorous comparison is needed to determine the best way forward.

While progress has been made, future research is needed to continue refining the process by which we can quickly map soil crusts on large scale. The initial results are promising, but accuracies can be improved. We will seek to do this by:

- Including multispectral data to classify base off of spectral indices designed for biocrusts (Rodriguez-Caballero et al 2015)
- Test deep learning classification methods that have shown to produce accurate maps of other phenomenon (Morgan et al 2021)
- Testing the alignment of georeferenced materials
- Testing other remote sensing data sources
- Object based classification instead of pixel-based.

Once the most accurate mapping method has been validated, we seek to identify spatial and temporal patterns of soil crust change over time, especially in areas of high traffic in national parks.

Reference List available from First Author upon request. References include sources for background data on Pyramid Lake and Lake Mead.