MINERALNET: A Computer Vision Approach to Mineral Species Identification

Gawan Fiore, Jorge Cueto, Guillermo Gomez

gfiore@stanford.edu, jcueto@stanford.edu, ggomez2@stanford.edu

CS 230: Deep Learning, Stanford University

Abstract

Mineralogy and geology have not historically leveraged the technological advancements of machine learning. While scientific tests for mineral identification exist, they are costlier and more difficult to access in the field compared to simple visual identification, which is almost always done manually by field researchers and exploration geologists. Similarly, mining companies spend significant resources separating debris from the specific minerals they are mining. Automated mineral species identification has the potential to become an invaluable tool for geologists, mining companies, and even the average rockhound. In this MineralNet study, we tested a variety of computer vision approaches to identify 53 mineral species from around the world.

Dataset & Features

Mindat.org dataset:

- 53 mineral species
- 106,000 total images
- RGB images in various resolutions and dimensions
- Preprocessed to dimensions of 360x360
- Labelled with species and locality

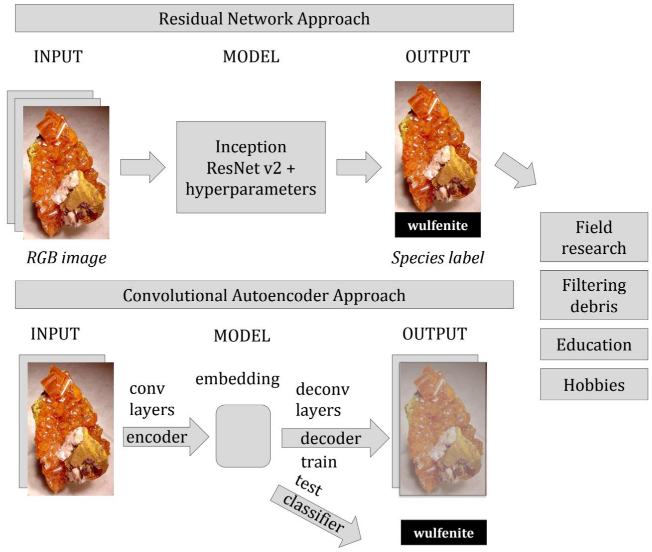
Sample dioptase image

Data augmentation:

- Flip images horizontally
- Perform random crops off the training images
- Randomly scale the size of the training images
- Randomly adjust the brightness of the training image pixels

Random crop of dioptase image

Model & Results



Model	Orig. Images Per Species	Number of Species	Random Crop	Random Scale	Accuracy (%)
Inception ResNet v2	2,000	53	Yes	Yes	34.5
Inception ResNet v2	2,000	53	No	Yes	32.1
Inception ResNet v2	2,000	53	Yes	No	35.6
Inception ResNet v2	2,000	53	No	No	30.2
Convolutional Autoencoder	2,000	53	Yes	Yes	35.4
Trained Layman Baseline	2,000	53	No	No	27.1
Expert Baseline	2,000	53	No	No	66.0

Conclusions

- Identifying mineral species is a difficult problem, both for people and computers, and is an underexplored area of the computer vision landscape.
- We developed a proof-of-concept and explored several different approaches to the task. It merits further study, and embedding-based models trained on larger datasets are likely the most effective approaches going forward.
- However, even our simplest model handily outperforms a moderately educated layman who was given a list of minerals, the training images, and a few hours to research them online. Still, our best model has a long way to go to beat an expert, but even the expert only correctly identifies about 2/3 of images!

Future Applications & Approaches

- Knowing a mineral's **specific locality** is important for understanding its significance and value. **FaceID-inspired approaches** using learned embeddings could lead to progress in identifying minerals with few examples, especially as a locality-species pair.
- Minerals in the field look very different than they do in a museum or gallery setting. Acquiring a **broader set of training data** that represents more of these modalities would make results more **generalizable**.
- Due to limited time, we retrained Inception ResNetv2 off its already learned weights, which gave a huge **initial boost due to low-level features** already baked into the model but limited the specificity of what we were able to detect. Training a **similar model end-to-end** would allow us to tune the receptive field and filter sizes to pick up features particular to minerals.
- Our **convolutional autoencoder** model showed promise, but since it did not start from pretrained weights our dataset is too small for the model to be generalizable. With a larger dataset and more time, we hypothesize this would be the best architecture to pursue.