

# Mineralogy meets zero-shot computer vision

August 25 2023, by Artur Kadurin



Examples of mineral photos from different datasets. Credit: Artur Kadurin

Identifying minerals is a complex and time-consuming problem for geologists, often taking anywhere from 30 minutes to several days per sample. Further complicating the situation is the fact that a sufficient portion of minerals remain inadequately researched, leaving us with just a few hundred comprehensively characterized out of the 6,000 currently

identified minerals.

Visual diagnostics of minerals and rocks is a widespread practice in geology, because it is much cheaper and faster than other methods, such as spectroscopy and chemical analysis. However, it is time-consuming and less accurate compared to more expensive methods. Even experienced mineralogists can make mistakes when working with a rare material or low-quality sample. Incorporating machine intelligence into this process can help with error identification and reduce the time spent on routine tasks by experts.

Despite ongoing research in this area, there is a lack of clear benchmarking for mineral image analysis in the scientific literature. To address this gap, the Artificial Intelligence Research Institute, in collaboration with Sber AI and Lomonosov Moscow State University, has created a benchmark dataset for computer vision models focused on [mineral recognition](#).

We called the dataset MineralImage5k. It is based on the Fersman mineralogical museum's collection and contains 44 thousand samples. While smaller than the Mindat dataset, MineralImage5k offers greater homogeneity of photo conditions and consists of unprocessed samples that closely resemble natural minerals.

The MineralImage5k dataset is divided into three subsets of varying complexity, challenging researchers in mineral classification, segmentation, and size estimation. The simplest classification task presented in the benchmark contains ten mineral species with at least 462 examples per specie. The most tricky problem is to classify minerals to 5K classes with only one image per class available.

One problem that AI may face when working with photos of a mineral is which part of the presented rock is an actual mineral of interest. To

address this problem, we share a separate set of about 100 images with additional labels and the segmentation task in addition to the classification. Integrating the segmentation task into the classification pipeline may provide additional insights in cases when the model makes mistakes and reduce the number of such situations.

Beyond the classification and segmentation, we study zero-shot mineral size estimation. Automatic specimen size estimation could be very useful for museum specimen storage procedures. Having these data for all samples, we can plan the optimal storage system and purchase or manufacture boxes of the right size in the correct quantity. Therefore, we provide more than 18K labeled samples for the regression task in our benchmark.

To demonstrate the effectiveness of the benchmark, we evaluated a vision-language model pre-trained on general domain data. We found that fine-tuning the model on the domain-specific dataset such as MineralImage5k may significantly improve its accuracy. We also highlight the promising potential of cross-dataset evaluation for assessing mineral recognition models.

Our research is published in the journal *Computers & Geosciences*. We are happy to help with the usage of the dataset and benchmark, and we invite all interested researchers to share their ideas on making it more useful for the community.

*This story is part of [Science X Dialog](#), where researchers can report findings from their published research articles. [Visit this page](#) for information about ScienceX Dialog and how to participate.*

**More information:** Sergey Nesteruk et al, MineralImage5k: A benchmark for zero-shot raw mineral visual recognition and description, *Computers & Geosciences* (2023). [DOI: 10.1016/j.cageo.2023.105414](https://doi.org/10.1016/j.cageo.2023.105414)

Artur Kadurin is the former Chief AI Officer at Insilico Medicine, a company utilizing Deep Learning techniques for drug discovery and aging research. He is now leading the "DL in Life Sciences" research group at Artificial Intelligence Research Institute, AIRI. He and his colleague Denis Dimitrov can be contacted via email (kadurin@airi.net, dimitrov@airi.net) if you need any help running your experiments on their data.

Citation: Mineralogy meets zero-shot computer vision (2023, August 25) retrieved 22 February 2024 from <https://techxplore.com/news/2023-08-mineralogy-zero-shot-vision.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.