

## Contents

<b>1</b>	<b>Raman Spectroscopy for Mineral Detection and Sorting</b>	<b>1</b>
1.1	Summary	1
1.1.1	Milestone 1: Investigating Discrepancies	1
1.1.2	Milestone 2: Developing Transferable Machine Learning Models	1
1.1.3	Milestone 3: Building a Design Space for Single Species	1
1.1.4	Milestone 4: Creating Realistic Design Spaces	2
1.1.5	Milestone 5: Enabling Deployment	2
1.1.6	Supporting Information	2

## 1 Raman Spectroscopy for Mineral Detection and Sorting

details

### 1.1 Summary

A critical aspect of mineral processing is selecting the appropriate technique to extract minerals from mined ore, which heavily relies on accurate species identification. This task can be challenging even for experienced professionals—minerals like chalcopyrite, pyrite, and gold can be easily mistaken for each other due to their similar visual appearances (e.g., yellowish color with metallic luster). Such misidentifications can lead to incorrect processing choices, such as using flotation to extract non-existent species, resulting in substantial financial, material, and time losses. This issue is exacerbated with low-grade ore samples, which are increasingly common due to continuous mineral exploration and the declining grades and availability of mineral resources. The aim is to develop digital twins—high-fidelity computational models that replicate physical materials and processes. These models can be deployed at mining sites for the smart identification of mineral species and the selection of appropriate extraction processes.

Electromagnetic spectra and Raman spectra (fingerprints) will be used to identify species in mined ore samples. Currently, two comprehensive sets of spectral fingerprints exist:

- Elemental spectra: Hosted at NIST’s Standard Reference Database – Atomic Spectra Database (ASD); these spectra are unique for all known elements.
- Mineral spectra: Available from the U.S. Geological Survey Data Series and NASA’s Jet Propulsion Laboratory, although discrepancies exist depending on factors such as mine location and purity/contamination.

#### 1.1.1 Milestone 1: Investigating Discrepancies

We will examine the fundamental building blocks of minerals—atoms—and their unique spectral fingerprints. By correlating the elemental spectra from ASD with mineral spectra, we aim to understand and account for discrepancies.

Using an Extended Iterative Optimization Technique (EIOT), we will extract intrinsic (constant) and non-intrinsic (variable) features from the spectra. Principal Component Analysis (PCA) will identify the most relevant components while maintaining a high data correlation ( $> 99\%$ ).

#### 1.1.2 Milestone 2: Developing Transferable Machine Learning Models

Current mineral identification methods using hyperspectral sensors rely on supervised machine learning, which requires labeled data. This is limited by:

- The need to retrain models for each mine location,
- The inability to handle previously unseen mineral species.

We propose developing **unsupervised, physically informed machine learning models** that adapt to new data on-the-fly.

#### 1.1.3 Milestone 3: Building a Design Space for Single Species

We will compile data on minerals of concern at Canadian mine sites, including their phase structures, transitions, and electromagnetic spectra. When data is lacking, we will use Density Functional Theory (DFT) and Molecular Dynamics (MD) to compute spectra.

This dataset will train a Convolutional Neural Network (CNN), which will be validated using a separate test dataset.

#### 1.1.4 Milestone 4: Creating Realistic Design Spaces

While a CNN trained on pure species data can identify individual minerals, it may struggle with mixtures. To address this:

- Mixture models of the collected species will be developed,
- Random configurations will be generated and their spectra computed using Machine Learning Interatomic Potentials (MLIP) in MD,
- A CNN trained on these mixture datasets will be tested with experimentally collected hyperspectral images.

#### 1.1.5 Milestone 5: Enabling Deployment

We will explore the development of a low-cost, fast, cloud-based self-learning model for mineral detection called **otfMLesf** (on-the-fly Machine Learning models for electromagnetic spectrum fingerprints).

- The model will use the CNN from Milestone 3 to detect single species in real-time,
- For mixtures involving detected species, it will recall the CNN from Milestone 4 to quantify composition,
- For unknown species, it will build random mixture models, compute spectra as in Milestone 4, retrain the CNN, and proceed to quantification.

This work aims to revolutionize mineral identification at mine operations through the development of robust digital twins and advanced machine learning models, ensuring efficient and accurate resource extraction with minimal waste.

#### 1.1.6 Supporting Information

- [Automatic materials characterization from infrared spectra using convolutional neural networks](#)
- [Leveraging infrared spectroscopy for automated structure elucidation](#)
- [On the interplay of the potential energy and dipole moment surfaces in controlling the infrared activity of liquid water](#)  
the infrared (IR) spectrum of a generic molecular system can be obtained, within the electric dipole approximation and linear response theory, from the Fourier transform of the quantum time autocorrelation function of the system's dipole moment >
- [Raman spectroscopy of minerals and mineral pigments in archaeometry](#)
- [Stand-off Raman spectroscopic detection of minerals on planetary surfaces](#)
- [RRUFF - Comprehensive Database of Mineral Data \(old website\)](#)
- [Mindat.org is the world's largest open database of minerals, rocks, meteorites and the localities they come from](#)
- [Raman spectroscopy for mineral identification and quantification for in situ planetary surface analysis: A point count](#)