1D Profiles vs. Spectral Images: A Comparative Study of Machine Learning Models for Mineral and Rock Classification

Sai Puneeth Reddy Gottam[1]
sai.gottam@unileoben.ac.at

Martin Johannes Findl[2]
martin.findl@unileoben.ac.at

Robert Galler[3]
robert.galler@unileoben.ac.at

Klaus Philipp Sedlazeck[2]
philipp.sedlazeck@unileoben.ac.at

Elmar Rueckert[1]
elmar.rueckert@unileoben.ac.at

Abstract

The rapid identification of minerals is critical for real-time geological analysis. This study investigates the efficacy of machine learning models in classifying mineral and rock samples using high-speed Raman sensors. We evaluate three distinct data representation strategies: **(1) 1D spectral profiles**, **(2) 2D Raman spectral images**, and **(3) a fused multi-modal approach** combining both spatial and spectral features. Using a diverse dataset of geological samples, we benchmark several model architectures to determine the trade-offs between computational efficiency and classification accuracy. Our results demonstrate how spatial context from imaging can enhance identification compared to traditional 1D methods, while also identifying the scenarios where signal-only processing remains optimal. This work provides a framework for selecting the most effective data representation for high-speed, automated mineralogical mapping.

1 INTRODUCTION

Rapid and accurate identification of mineral phases and rock compositions is a cornerstone of modern geological exploration, planetary science, and industrial mining. Traditional mineralogical analysis often relies on manual petrography or X-ray diffraction (XRD) [1], which, while accurate, are time-consuming and difficult to implement in high-throughput or autonomous environments. **Raman spectroscopy** has emerged as a powerful alternative due to its non-destructive nature and its ability to provide a unique "fingerprint" of a material's molecular structure [1, 2].

Recent advancements in sensor technology have shifted the paradigm from single-point Raman probes to **high-speed spectral imaging**. These modern sensors allow the collection of large amounts of data in short timeframes, capturing both the chemical signature (1D spectral profiles) and the spatial distribution (2D spectral images) of mineral grains. However, the influx of high-cadence data presents a significant computational challenge[5]: how to effectively process and classify these signals in real-time.

[1]Chair of Cyber Physical Systems, Montanuniversität Leoben, Austria
[2]Chair of Waste Processing Technology and Waste Management, Montanuniversität Leoben, Austria
[3]Chair of Subsurface Engineering, Montanuniversität Leoben, Austria

The Third Austrian Symposium on AI and Vision (AIROV26).

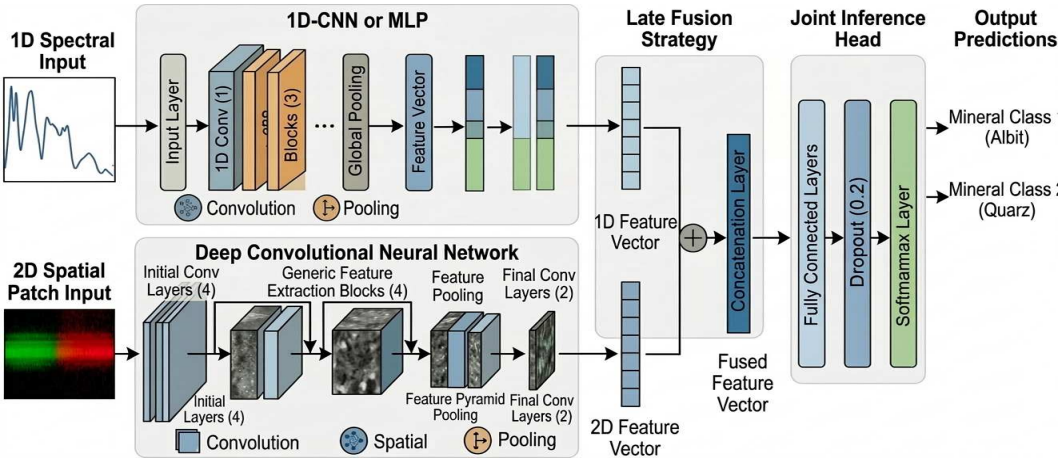

Figure 1: Proposed multi-modal late-fusion architecture. The model processes 1D spectral inputs and 2D spatial patches through independent feature extraction branches. These high-dimensional vectors are concatenated and processed by a joint inference head with dropout regularization to provide final mineral classifications.

Machine learning (ML) and deep learning (DL) have shown immense promise in automating spectral analysis [1, 4]. Although 1D Convolutional Neural Networks (1D-CNNs) are efficient in processing individual spectra, they often ignore the textural and morphological context provided by the surrounding mineral matrix. Conversely, 2D imaging models capture spatial relationships but may introduce unnecessary computational overhead if the spectral signal alone is sufficiently discriminative [6].

This paper investigates the trade-offs between different data representations for mineral classification. We conduct a comprehensive experimental study using high-speed Raman data to evaluate three distinct approaches:

- **1D Profile Analysis:** Classifying minerals based purely on individual spectral signatures.
- **Image-Based Classification:** Utilizing 2D spatial-spectral maps to capture mineral texture and grain boundaries.
- **Fused Multi-Modal Learning:** Investigating whether the integration of both 1D and 2D features yields superior robustness.

By benchmarking these methods, we aim to identify the optimal balance between classification accuracy and processing speed, providing a roadmap for the deployment of autonomous mineralogical mapping systems.

2 Method

2.1 Dataset and Experimental setup

The dataset consists of Raman measurements collected under controlled laboratory conditions. It includes multiple mineral classes as well as heterogeneous rock samples composed of mixed mineral phases.The mineral dataset comprises 12 classes with 3000 images and profiles of each mineral class. The rock dataset has 12 classes with 3000 images and profiles of each class. Some of these rock classes are the variations of different version of same rock type such as varities of Granite and Sandstone.

The dataset was split into training, validation, and test sets using a standard 70/15/15 ratio to ensure unbiased evaluation. All models were evaluated on a held-out test set.

2.2 Data Preparation and Normalization

All Raman spectra were normalized using min-max scaling to ensure consistency across samples and mitigate variations due to sensor conditions such as laser power and acquisition distance.

- **1D Profiles:** Represented as intensity vectors across the wavenumber domain
- **2D Spectral Images:** Constructed from spatial Raman scans to capture texture and morphology

2.3 Model Architectures

We evaluated five architectures:

- **Random Forest (RFC):** Baseline model using 100 trees for 1D classification
- **MLP:** Fully connected network with ReLU activations and dropout
- **1D-CNN:** Convolutional model capturing local spectral features
- **ResNet-18 (2D) [3]:** Image-based classifier leveraging spatial context
- **Multi-Modal Fusion:** Late fusion architecture combining 1D and 2D features as shown in 1. The fusion model concatenates latent representations from both branches before final classification.

3 Preliminary Results and Discussion

The benchmarking phase evaluated five distinct architectures across two primary geological scales: pure mineral phases and complex rock aggregates. The results reveal a significant disparity in how different models handle the high-cadence 30fps data stream. All models are trained on mineral dataset but only 2 models are trained on Rock dataset.

Table 1: Classification Performance across Model Architectures and Mineral Dataset

Model Architecture	Data Input	Mineral Acc (%)	Rock Acc (%)	Complexity
Random Forest (Baseline)	1D Profile	**99.96%**	-	41,450 nodes
MLP	1D Profile	99.10%	-	Medium
1D-CNN	1D Profile	80.92%	-	Low
ResNet-18	2D Image	>99.9%	99.77%	11.2M Params
Multi-Modal Fusion	Hybrid (1D+2D)	99.85%	**99.98%**	High

3.1 Performance on Pure Mineral Dataset

For the identification of individual mineral phases, the high-resolution spectral fingerprints proved to be highly discriminative.

- **Classical Baseline Excellence:** Surprisingly, the **Random Forest Classifier (RFC)** emerged as the top-performing model for 1D mineral profiles, achieving an accuracy of **99.96%** as shown in table 1. This suggests that for normalized laboratory data, the ensemble of decision trees effectively captures the essential vibrational peaks without the computational overhead of deep learning.
- **Deep Learning Performance:** The **MLP** and **ResNet (2D Image)** models also performed exceptionally well, both exceeding **99.9%** accuracy. This confirms that the spectral "images" extracted from the sensor provide sufficient morphological detail to distinguish pure minerals with near-perfect reliability.
- **The 1D-CNN Outlier:** Interestingly, the **1D-CNN** was the only architecture that struggled in this phase, achieving only **80.92%** accuracy. This indicates that a simple convolutional approach may be overly sensitive to spectral shifts or lacks the global feature integration required for this specific dataset. Also there are very few parameters in this model which fail to capture the characteristics needed for classifying accurately.

The high accuracy values are attributed to the controlled laboratory environment and high-quality spectral signals.

3.2 Performance on Rock Dataset

When transitioning to rock classification—where minerals are found in complex, heterogeneous aggregates—the performance characteristics shifted. The models were able to distinguish between same rock type and effectively classifying the different variations of same rock.

- **Robustness of Spatial Data:** The **ResNet-based Image Classifier** proved highly robust. By freezing the early layers of the ResNet-18 backbone and fine-tuning the final layers, the model successfully identified rock types based on texture and mineral associations, even when individual spectral pixels were noisy.
- **Multi-Modal:** The **Multi-Modal Fusion** model (combining 1D Profiles and 2D Images) achieved a high accuracy of **99.98%**. It proved to be the most comprehensive solution for rocks, as it might cross-reference chemical signatures with spatial morphology.

4 Conclusion and Future work

This study compared three data representation strategies for the automated classification of minerals and rocks using high-speed Raman sensors. Our results indicate that, while 1D spectral profiles provide a computationally lightweight solution for rapid identification, they are susceptible to noise in chemically complex samples where overlapping peaks occur. In contrast, 2D spectral imaging architectures significantly improved classification accuracy by capturing better context of the spectral shift.

The fused multi-modal approach demonstrated the highest robustness, successfully identifying mineral phases in heterogeneous rock samples. However, for industrial applications, the choice of model must balance this accuracy against the latency requirements of high-cadence data streams. The results were very accurate, possibly due to the data collected under ideal conditions.

Future work will shift towards the challenges of dynamic, moving environments, where minerals must be classified in real-time as they pass under high-speed sensors in a continuous flow. We aim to optimize these 1D-2D fused architectures to maintain high precision despite the motion-induced artifacts typical of industrial transport systems. This will involve implementing edge-computing solutions to minimize the processing delay between data acquisition and automated sorting decisions.

5 Acknowledgment

The presented research was funded by the Austrian Research Promotion Agency (FFG) under Project NNATT.

References

[1] Chris Carey et al. Machine learning methods for classifying mineral species from Raman spectra. *Journal of Raman Spectroscopy*, 46(10):894–903, 2015.

[2] Larry A Haskin, Alian Wang, Kaylynn M Rockow, Bradley L Jolliff, Randy L Korotev, and Karen M Viskupic. Raman spectroscopy as a tool for in situ lunar mineralogical analysis. *Journal of Geophysical Research: Planets*, 102(E8):19293–19306, 1997.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[4] Jun Liu/m et al. Deep learning for rapid mineral identification using Raman spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 252:119511, 2021.

[5] J. Sun et al. High-speed Raman imaging for real-time monitoring of mineralogical processes. *Sensors and Actuators*, 380:133314, 2023.

[6] Xiao Zhang et al. A comparison of 1d and 2d deep learning models for Raman spectral-spatial classification. *Remote Sensing*, 11(12):1473, 2019.