Convolutional Neural Networks as a Tool for Raman Spectral Mineral Classification Under Low Signal, Dusty Mars Conditions

Genesis Berlanga, Quentin Williams, and Nathan Temiquel

Abstract NASA's Mars 2020 and ESA's ExoMars will collect Raman measurements in dusty field conditions obscuring underlying rocks. This presents a challenge for remote Raman measurements at distances where mechanical or ablative sample cleaning is not straightforward. Historically, probing broad lithostratigraphic suites has been thwarted by the need for pristine targets and high-quality spectra. We provide a means of identifying Raman spectra of common rock-forming silicate, carbonate, and sulfate minerals under low signal-to-noise-ratios, Mars-like conditions using a convolutional neural network (CNN). The CNN was trained on the Machine Learning Raman Open Data set data set with 500,000+ Raman spectra of hand samples/powder mixtures (5,000+ spectra/mineral class). Diversity in sample microtopography, orientation, and crystallinity simulated varying laser focuses and spectral quality, and no traditional spectral preprocessing such as cosmic ray or baseline removal was employed. The CNN identified low-intensity Raman scatterers (micas and amphiboles), mixed minerals, and distinguished between mineral endmembers with +99% success. We present among the first known implementations of “big data” machine learning using varied, high-volume Raman spectral datasets. The pattern recognition abilities of CNNs can facilitate scientist Raman spectral interpretation on Earth and autonomous rover decision-making on planets like Mars; increasing scientific yield, correcting human classification errors, reducing the need for thorough target dust removal during evaluative measurements, and streamlining the data communications pipeline—saving time and resources. This study examines an end-to-end development process for creating a deep learning algorithm sensitive to varieties of Raman spectra and provides guidelines for CNN model development at the interface of Raman spectroscopy, deep learning, and planetary science.

Plain Language Summary We collected thousands of spectra from rocks and minerals that are found on Earth and Mars and trained a computer to automatically identify them using an algorithm called a convolutional neural network or CNN. This is an effective but difficult algorithm to fine tune for Raman spectra as it requires a lot of training. A CNN works like an artificial brain with firing neurons. This brain can be uploaded onto Mars rovers such as NASA's Curiosity and Perseverance so that they can automatically identify rocks and minerals on Mars without human guidance. Doing this will accelerate scientific discovery while saving crucial mission time and energy. We essentially programmed the “brains” of a rock and mineral identification “tricorder”.

1. Introduction

Understanding the geology of a rocky planetary or lunar surface provides key insights into the terrestrial body's petrogenesis, surface evolution, and habitability. This process requires detailed understanding of morphological, physical, chemical, and mineralogical analysis of rocks and minerals within stratigraphic layers, units, outcrops, and landforms. Due to the risk involved in exploring the remote terrestrial reaches of Earth or extraterrestrial surfaces of the Moon or Mars, robotic rovers serve as valuable alternatives to the field geologist. Several relevant planetary rovers are currently exploring Mars or set to begin exploration in the coming years: NASA's Mars 2020 Perseverance rover and the Mars Science Laboratory Curiosity rover are currently acquiring thousands of images and spectra on the surfaces of Jezero and Gale Crater, respectively. (Farley et al., 2020; Mars Exploration Program and Jet Propulsion Laboratory for NASA's Science Mission Directorate, 2020a, 2020b; Maurice et al., 2012; Mustard et al., 2013) ESA's 2022 ExoMars rover is expected to begin data collection at Oxia Planum in late 2022 or early 2023 (Lopez-Reyes, Rull, et al., 2014). With the exponential increase in the acquisition of
The intended applications are three-fold: (a) to expedite and facilitate scientists’ Raman spectral interpretations, thereby increasing scientific yield; (b) to propose a method that can be used for future autonomous rover decision making; and (c) to develop a technique that can facilitate interpretation of low signal-to-noise, suboptimal spectra. In particular, we demonstrate the classifier's robustness by simulating Mars' dusty environment and testing it on gabbro and granite rocks covered in varying levels of basaltic dust. Several independent laboratory groups have conducted excellent applied machine learning and Raman spectroscopy studies using smaller spectral datasets (~1,000–5,200 total spectra) and employing network architectures fine-tuned to their respective data set sizes (Carey et al., 2015; Ishikawa & Gulick, 2013; Kwiatkowski et al., 2010; Liu et al., 2017). We report an initial application of statistically significant, high-volume, “big data,” machine learning using over 100,000 individual Raman spectra (Emmanuel & Stanier, 2016). The effects of hyperparameter tuning and neural network architectures are discussed in light of varying data set sizes and quality. We also present among the first machine learning-specific Raman datasets (Machine Learning Raman Open Data set (MLROD)) for open use to the scientific community (Berlanga, Williams, & Temiquel, 2022). Moving Raman spectral characterization into the “big data” regime enables the use of a sufficiently large teaching set for the deep learning algorithms to maximize detection and identification of weak scatterers or low-abundance constituents, and to create the prospect that mineralogic issues, such as solid solutions or multiple juxtaposed phases, can be treated as accurately as possible within potentially low signal-to-noise spectra. This study examines an end to end development process for creating an appropriate deep learning algorithm sensitive to the varieties of Raman spectra and provides general guidelines for targeted CNN model development at the interface of Raman spectroscopy, deep learning, and planetary science.

1.1. Raman Spectroscopy for Planetary Science Applications

Raman spectroscopy is an active-laser spectroscopic technique that provides bonding and symmetry information about the target material; practically, it is often deployed as a diagnostic identification technique for different minerals or mineral assemblages. A monochromatic laser light, such as a 532 nm green laser, is directed at a sample and the resulting scattered light is collected. A small fraction of the scattered light (typically of order <0.000001%) is shifted in energy from the laser frequency due to interactions between the incident electromagnetic waves and the bonds and atoms within the spectroscopic unit cell of the crystalline sample. This inelastic shift is called the Raman effect and occurs instantaneously at picosecond timescales (Berlanga et al., 2019; Blacksberg et al., 2020; Johnson et al., 1985). A Raman spectrum is then produced with peaks showing the intensity and wavelength position of the Raman scattered light. Each peak corresponds to the energy levels of different vibrational modes dictated by the analyte’s symmetry and bonding vibrations, thus enabling unique sample identification akin to a barcode or fingerprint (Lewis & Edwards, 2001). Typically, stronger and more distinctive Raman scattering occurs from minerals or materials containing molecular-type units (e.g., carbonate-, sulfate-, and silicate-groups), but low-frequency lattice-associated modes may also prove to be diagnostic (Sharma & Egan, 2019; Sobron et al., 2019).

1.1.1. Complexity of Rock and Mineral Raman Spectra

Raman spectra provide basic molecular information about the targeted sample, and additionally, various aspects of rock or mineral petrogenesis, crystalline order, textural imprints, chemical variations, possible bounds on grain size, opacity, and fluorescence centers can be ascertained from the data. While informative, these nuanced spectroscopic signatures often have the effect of complicating actual mineral or rock identification by introducing Raman peak shifts, peak broadening, or baseline enhancement. Doppler, collisional, thermal, chemical, and physical effects homogeneously or inhomogeneously broaden Raman peaks (Artur et al., 2011). These broadening effects can be generated by poor crystalline order (Tuschel, 2017), shock effects (Heymann & Cellucci, 1988), and smaller-sized grains (Foucher et al., 2013). Opaque, mafic minerals limit Raman visible laser penetration...
and reduce internal scattering, decreasing the resultant available Raman signal for peak detection. Fluorescence centers commonly contribute broad peaks and spectral baselines that sometimes saturate whole spectrum intensities and can obscure or entirely mask Raman peaks (Jones et al., 2019); a few narrow-line fluorescent centers (such as the rare earths) can generate peaks that are comparable in width to Raman peaks (Lenz et al., 2013). Sample morphology or surface scattering, preferred orientation, and impurities contribute additional skewing of characteristic spectral signatures. With these complexities in mind, accurate mineral and rock identification through Raman spectroscopy alone can potentially become a very time-consuming task, requiring an experienced spectroscopist familiar with both the technique and the materials being analyzed.

1.1.2. Challenges With Raman Spectroscopy on a Dusty Mars

The surface of Mars is covered with vast expanses of dust and sand ranging from nanophase to millimeter-sized grains. The coarser material is ~55–71% crystalline and of basaltic origin, and the finer amorphous material is ~29–45% nanophase iron oxide (Bish, Blake, Vaniman, Sarrazin et al., 2014; Bishop et al., 2002; Blake et al., 2013; Grotzinger, 2013). Much of this dust is mechanically derived from weathered Martian mafic or basaltic rocks, whose amorphous component is compositionally similar to Earth volcanic soils such as those from Mauna Kea volcano, Hawai`i (Bish, Blake, Vaniman, Chipera et al., 2014; Certini et al., 2020; Morris et al., 2001; Ojha et al., 2018). This dust circulates globally and remains suspended in the atmosphere for months after dust storms (Kahre et al., 2017). Eventually, the aeolian dust settles on rocks and boulders strewn across the surface, leaving a fine veneer that may impart a compounding optical scattering and eclipsing of Raman peaks.

NASA's Mars 2020 Perseverance rover and the Mars Science Laboratory Curiosity rover are traversing through these challenging and dusty conditions common to Mars. Both rovers carry dust removal tools that employ either pressurized gas (Farias et al., 2020) or wire-bristle brushes (Schmidt et al., 2018) to clean off fine dust veneers, but in situations where target samples are beyond the reach of the turret arm, laser ablation is required for sample cleaning. Curiosity comes equipped with the ChemCam instrument suite which includes a laser-induced breakdown spectroscopy (LIBS) system that has the ability to ablate target samples prior to actual analysis (Maurice et al., 2012). This ability has the potential to remove veneers or dust coverage obscuring the target (Lanza et al., 2015). Perseverance comes equipped with the next generation SuperCam instrument that includes a dual remote Raman and LIBS spectrometer capable of ablation and the Scanning Habitable Environments for Raman and Luminescence for Organics and Chemicals (SHERLOC) instrument, a deep ultraviolet (DUV) Raman spectrometer. Due to the varying spot diameters used for LIBS, remote Raman, and DUV Raman techniques, analyzing dusty rocks on Mars will involve repeated LIBS ablation using a 500 µm maximum diameter beam across the surface area of the larger Raman laser's 8 mm maximum diameter collimated beam (SuperCam laser at distances over 12 m). It takes approximately five LIBS shots, or the equivalent of 17% of an average 30-shot LIBS observation, before being able to adequately sample the clean target area (Bhartia et al., 2021; Wiens et al., 2021; Wiens, Maurice, and Rull Perez, 2017). The time and laser power required to clean a target before actual analysis can take place provides a major impediment to the number of possible analyses. As the progressive investment of time and power accumulates and the available consumable laser shots declines, serial laser ablation over multiple Mars years may ultimately become unsustainable. Here, we present a computational means of enhancing Raman spectral analysis and sample identification under dusty, low signal-to-noise (SNR) conditions, that saves resources by minimizing the use of extensive laser ablation during cursory exploratory spectral acquisitions.

1.1.3. Efficiency of Earth-Mars Communications

A communications lag and data bottleneck exist between Earth and Mars that hinders scientific data acquisition and human decision-making in directing rover operations. Round-trip interplanetary communications between Earth and Mars can take up to 48 min, and data rates range from 32,000 bits/s (about half as fast as a current standard home modem) to 2 million bits per second. Maximum bandwidth is not always available due to rover-orbiter-Earth lines of sight, rover power limitations, or conflicts with other planned activities (Mars Exploration Program and Jet Propulsion Laboratory for NASA's Science Mission Directorate, 2020a, 2020b). This communications time lag limits the number of commands sent to Mars and the bandwidth bottleneck slows the transmission of large data volumes. Our proposed computational enhancement of Raman spectral analysis and sample classification can partially offload decision-making from scientists on Earth to the rover, reducing the number of commands sent to Mars, and streamlining data acquisition, increasing the quality and decreasing the volume of data relayed to Earth.
1.2. Machine Learning as a Tool for Spectral Interpretation and Robotic Autonomous Decision Making

Machine learning solves problems by discovering the rules, patterns, and statistical structure underlying a large data set (Bishop, 2006; Chollet, 2018; Goodfellow et al., 2016; Witten et al., 2011). This opens the door for the creation of a new kind of toolset or assistant to the scientist—one that optimizes and expands problem solving and scientific output. We address the two-pronged case of (a) increasing volumes of spectral data arriving from remote robotic rovers on Mars and the Moon and (b) the data uplink and downlink bottleneck constrained by limited interplanetary communications bandwidth between distant planetary rovers and scientists on Earth (Figure 1). The former case tasks scientists with a repetitive and ever-growing analysis and interpretation workload, while the latter interplanetary communications bottleneck increases the number of simplistic command cycles required to acquire and produce viable scientific knowledge. Machine learning can facilitate automation of spectral sorting, analysis, and interpretation in terrestrial laboratory settings as well as enhance autonomous decision-making in remote rovers equipped with intelligent onboard systems. This has the effect of streamlining scientific data analysis and reducing the number of command cycles between human and rover. We provide a method that tackles these problems and ultimately increases the volume, speed, and quality of scientific output.

The Perseverance and Curiosity rovers are equipped with existing machine learning software (Autonomous Exploration for Gathering Increased Science) for autonomous hazard navigation and blind laser targeting (Francis et al., 2017). On Perseverance this software uses a RAD 750 processor that runs at 200 MHz and has access to 16 MB of RAM. This hardware is sufficient to run complex image recognition and segmentation of rock and boulder 2-dimensional images. Running a 1-dimensional spectral classification pre-trained CNN algorithm using such hardware is not inconceivable, and would provide a significant enhancement to an AEGIS-type software that chooses targets pre-spectral analysis, and superimpose a decision-making layer post-analysis that can reduce repetitive blind targeting.

1.2.1. Neural Networks

In this work we introduce the use of high-volume datasets and convolutional neural networks (CNNs) in the context of Raman spectroscopy. Similar to synaptic communication in biological neural pathways, a neural network receives inputs and feeds them forward to subsequent layers of neurons. This process is mathematically modeled through CNNs using a standard sequence. First, input variables are multiplied by random weights at each node in the input layer. Nodes in each following inner layer sum the weighted inputs from the previous layer and generate an output by applying an activation function. The final output layer sums the weighted inputs from the last inner layer and generates a numerical output, or prediction, via an activation function. Back-propagation recursively adjusts the weights of the network connections so as to minimize loss (Rumelhart et al., 1986). Convolutional neural networks have the capacity to capture complex data relationships and produce accurate predictions for high-volume datasets. This makes them ideal for synthesizing the multiple input variables found in spectra into a concise output.

Historically, Raman analyses pipelines have employed traditional statistics and machine learning requiring extensive spectral preprocessing such as cosmic ray removal, baseline smoothing and removal (Baek et al., 2014;
incorporate more realistic Mars dust simulants. Effects as dust on the surface of Mars. Future studies will include the effects of Mie and Rayleigh scattering and (such as Mojave Mars Simulant (MMS)-1 or MMS-2 (Peters et al., 2008)) will have similar signal:noise reduction effects as dust on the surface of Mars. Future studies will include the effects of Mie and Rayleigh scattering and incorporate more realistic Mars dust simulants.

2. Methodology

Roughly polished gabbro and granite slabs, ground and sieved basaltic dust (a heterogeneous size distribution, limited to 50–150 μm grain diameters), and 12 hand sample mineral specimens, four of which were mixed into three binary powders, were used as the test data set for our Raman spectroscopic and machine learning investigation. The irregularly shaped and distributed dust allowed for a range of different dust, sample, and beam configurations, including partial eclipsing by dust of the Raman beam, near total eclipsing, and minimal (or zero) eclipsing. Sample chemistries were confirmed using X-ray fluorescence (XRF), and sample mineralogy was confirmed through both Raman spectroscopy and powder X-ray diffraction (XRD). Over 20,000 spectra were collected at progressively shifted locations on the gabbro and granite slabs, with half of these collected under idealized laboratory conditions, and the other half under a ~50% basaltic dust coverage that mimicked dusty, in situ field acquisitions on the Martian surface. Over 80,000 mineral spectra were collected to represent major silicate, carbonate, and sulfate mineral structures found on Earth and Mars: silicates containing isolated tetrahedra (nonsilicates), single and double chain tetrahedra (inosilicates), sheets (phyllosilicates), framework silicates (tectosilicates), carbonates, and sulfates. Again, these were collected from varying locations within polycrystalline samples, and hence sampled various crystallographic orientations, and varying surface topologies. Some, but not all, of these minerals were primary mineral components in the gabbro and granite slabs. Training reference sets also included 12 pure mineral samples and four binary powder mixtures: quartz, albite, anorthite, microcline, hornblende, biotite, muscovite, forsterite, augite, enstatite, calcite, gypsum, a 1:1 quartz albite mixture, a 1:1 forsterite augite mixture, and a 1:1 forsterite albite mixture. Mineral localities and number of spectra obtained are tabulated in Table 1. This initial sample suite was chosen to encompass broad primary Raman peak coverage across the measured Raman wavenumber range as well as to include closely overlapping primary Raman peak instances in order to prevent training set biases while finetuning the CNN. Moreover, the distribution of samples incorporated the common silicate structural groups, as well as those bearing molecular-type units likely to be present on Mars. Future work will include additional Mars-relevant minerals, including intermediate plagioclase compositions, low-Ca and high-Ca pyroxenes like diopside and pigeonite, sulfates such as kieserite and anhydrite, and further exploration of Martian sedimentary minerals and phyllosilicate phases (Rogers & Hamilton, 2015).

Hawaiian basalt was finely ground and sieved to serve as a spectroscopic obscurant similar to what we expect from the fine Martian dust veneer found on the planet’s surface. Care was taken that the dust was: (a) sufficiently fine-grained that it generated no resolvable crystalline spectral features, (b) coarse enough to not contribute Mie or Rayleigh scattering effects, (c) not detectably Raman-active, and (d) exhibited low fluorescence. The basaltic dust thus functions as an optical scatterer eclipsing Raman scattering from the underlying rock. In this sense, the dust functions as a spectroscopically passive reducer of the signal: noise ratio of the Raman spectra. Hence, the precise nature or chemistry of the dust (as long as it remains spectroscopically inactive) is not critical for our results: our expectation is that similarly fine-grained basaltic dust or modestly compositionally tuned basaltic dust (such as Mojave Mars Simulant (MMS)-1 or MMS-2 (Peters et al., 2008)) will have similar signal:noise reduction effects as dust on the surface of Mars. Future studies will include the effects of Mie and Rayleigh scattering and incorporate more realistic Mars dust simulants.

Dust coverage percentages were measured out by mass and evenly spread and flattened across a 9 cm² surface whose perimeter contained all of the Raman measurements. Dust coverage percentages were computationally verified by acquiring a high definition black and white image of the 3 × 3 cm area and passing the image through
a pixel color indexing filter. The shift in the integrated pixel value histogram (proportional to the dust percentage) and the increase or decrease in the RGB value bins corresponding to gray (Bins 0–20) basaltic dust were compared between the clean and dusty gabbro and granite slabs.

3. Experimental Design

3.1. Raman Measurements

Raman measurements were performed using a Horiba LabRAM HR Evolution single stage spectrometer equipped with an air-cooled solid-state laser emitting at 532 nm/100 mW and filtered through a long range 532 nm notch filter. The spectrometer includes high throughput achromatic coupling optics, a spectrometer focal length of 800 mm, and a grating of 1,800 lines/mm with a spectral resolution of ∼1 cm$^{-1}$. An Open-Electrode charge-coupled device detector thermoelectrically cooled to −60°C collected the measurements. Raman scattering from the samples was collected using a long-distance aperture of 50× (NA = 0.5). Samples were placed on an automated Märzhäuser XY mapping stage (resolution 0.1 μm, reproducibility 1 μm) and spectra were collected using scan and point counting techniques (Haskin et al., 1997; Wang et al., 1999) at ambient temperature and pressure. A synthetic silicon standard was used for spectral wavenumber calibration. High quality raw Raman spectra were collected in conjunction with relatively short collection times of 1–120 s, and introduced into the CNN with only minimal preprocessing to standardize wavelength scales, as described below in Section 4.1. Raman measurements were collected between 100 and 1,800 cm$^{-1}$ to incorporate the fundamental range of vibrational modes for silicate, carbonate and sulfate minerals (Wang et al., 2015; Williams, 1995). Representative spectra for each training set are plotted in Figure 2. These spectroscopic measurement and laser excitation specifications fall within those of the SuperCam Raman measurements, enabling the use of the MLROD-trained CNN algorithm with SuperCam data. MLROD spectral resolutions can be downsampled within the software build to better match SuperCam or other spectrometer resolutions. The future addition of Raman training sets at a 248.6 nm laser excitation will allow the algorithm to be used on SHERLOC data. Relevant laser and spectrometer
specifications between current Mars payloads implementing Raman spectroscopy and the Horiba spectrometer used to create MLROD are tabulated in Table S3 of Supporting Information S1.

3.2. X-Ray Diffraction and X-Ray Fluorescence Measurements

To confirm the structures of our mineral standards, XRD patterns were measured between 5.0 and 65.0 2Θ in 0.02° steps with a count time of 2 s per step on an inXitu Terra diffractometer at NASA Ames Research Center. XRF measurements were collected on an inXitu Terra diffractometer to measure chemical compositions. Elemental compositions confirmed the chemistry and solid solution series position of each mineral specimen. This mineralogical information, comprising both symmetries and unit cell parameters and elemental compositions, established the phase and compositional ground truthing for the Raman analysis. XRD/XRF data and mineralogical information are available in the Supplement section.

4. Software Design

Open-sourced and freeware software and environment options were used for our investigation. Programming was completed using Python 3.8 with Jupyter Notebooks on Google Collaboratory as our development environment. Data processing and manipulation was done using the Pandas and Numpy packages (Harris et al., 2020; Reback et al., 2021). Traditional statistics and spectral signal processing were done using the SciKit Learn package (Pedregosa et al., 2011). Machine learning algorithms were developed using the Keras library on top of Tensorflow (Abadi et al., 2016).

Two datasets provided inputs to our machine learning classifiers: (a) the experimentally acquired raw Raman spectra of rocks and pure minerals and (b) a subset of excellent to poor quality publicly available spectra of pure minerals from the RRUFF Project Raman database (Lafuente et al., 2015) for independent CNN testing. No synthetic data augmentation was necessary due to our acquired data volumes being sufficiently extensive that they satisfied the training set size and variability needs for a CNN. The experimentally acquired pure and mixed mineral Raman spectra were used for CNN training, while the gabbro and granite spectra and the independent RRUFF Project Raman database spectra were used for testing. Two versions of the training and tests sets were inputed: a raw version and a minimally processed continuous wavelet transformed version.
4.1. Raw Raman Spectra Preprocessing

Prior to running the Raman spectra through the CNN, we standardized the wavelength range and interpolation intervals. This was the only pre-processing conducted on the spectra, with only truncations of data and sampling intervals being standardized. No cosmic ray removal or baseline corrections were applied. All spectra were trimmed to a wavelength range of 150 to 1,100 cm\(^{-1}\) in order to reduce file sizes and accelerate automatic processing and statistical analysis using standard commercial laptop processing speeds of 2.8 GHz and an Intel Core i7, while still including relevant primary Raman peaks for silicate or functional group identification.

Continuous wavelet transform-based (CWT) pattern matching was applied to the spectra due to the approach's sensitivity to peak shape and amplitude. Most peak detection algorithms identify peaks based solely on amplitude, sacrificing low-amplitude peak detection, and increasing the rate of false positives in low signal to noise situations (Du et al., 2006; Yang et al., 2009). By transforming the spectral data into wavelet space, we were able to simplify peak matching, increase the apparent SNR, and bypass baseline removal and peak smoothing steps. Raman shift resolution and ranges were standardized by interpolating to discrete steps of 1 cm\(^{-1}\) in order to account for variations between data set wavelength ranges and spectrometer resolution. For each 1 cm\(^{-1}\) bin, the minimum Raman intensity was chosen as the new intensity value for the interpolated pair. This step standardizes all input spectra and also has the added natural consequence of removing single pixel cosmic rays from the spectra. Care was taken to preserve adequate spectral peak information for identification. Figure 3 illustrates the Ricker continuous wavelet transform (CWT) implementation.

Two data sets are outputted from the Ricker continuous wavelet transform:

1. A 10-feature matrix with the highest peak intensities in CWT space in descending order from highest to lowest for each spectrum:

   | PeakPosition1 | PeakPosition2 | PeakPosition3 | PeakPosition4 | PeakPosition5 |
   |--------------|--------------|--------------|--------------|--------------|
   | Intensity1   | Intensity2   | Intensity3   | Intensity4   | Intensity5   |

   This set was used to test traditional statistics and machine learning classifiers.

2. The Ricker CWT of each spectrum with 190 features (one for every 1 cm\(^{-1}\) between 150 and 1,100 cm\(^{-1}\). These data are only used in the CNN. Traditional models like support vector classifiers or logistic regression cannot input this type of data due to the large number of features and noise. The general format is, where \(x_j\) and \(y_{j,k}\) denote wavenumber and intensity, respectively:

   \[
   \begin{align*}
   150 & 155 & x_j & \ldots & 1100 \\
   \text{Intensity1} & \text{Intensity2} & y_{j,k} & \ldots & y_k
   \end{align*}
   \]

   Table S1 in Supporting Information S1 shows the tabulated top five peaks from the 10-feature data set, for each mineral trained on the CNN, and the respective band assignments and vibrational modes. The top five peak positions were selected by taking a frequency distribution of the top five peak intensities in all spectra for each mineral and selecting the corresponding positions of the five peak intensities with the highest frequencies across the mineral class. While the CNN uses the entire Ricker CWT with 190 features, tabulating the top five peaks allows insights into the mechanics of how the CNN “thinks” about full spectrum matching. Not all primary peaks land in the top five for every mineral, but the Ricker CWT captures enough primary peaks to adequately influence the CNN node weights during classification. The weak peaks emphasized by the Ricker CWT provide identifying features throughout the entire spectrum, and within potentially diagnostic regions of the spectrum, instead of just being confined to grouped regions, in which the well-known vibrations of functional groups fall. These weak, often isolated features are not readily obvious to the untrained spectroscopist and may become crucial for spectral mineral identification when confronted with very noisy or highly fluorescent Raman spectra, or mixtures containing many mineral components. If the spectrum contains no identifiable peaks, the CNN algorithm has a classification option of “Dusty/Low SNR/Unknown”.

Table S1 in Supporting Information S1 shows the tabulated top five peaks from the 10-feature data set, for each mineral trained on the CNN, and the respective band assignments and vibrational modes. The top five peak positions were selected by taking a frequency distribution of the top five peak intensities in all spectra for each mineral and selecting the corresponding positions of the five peak intensities with the highest frequencies across the mineral class. While the CNN uses the entire Ricker CWT with 190 features, tabulating the top five peaks allows insights into the mechanics of how the CNN “thinks” about full spectrum matching. Not all primary peaks land in the top five for every mineral, but the Ricker CWT captures enough primary peaks to adequately influence the CNN node weights during classification. The weak peaks emphasized by the Ricker CWT provide identifying features throughout the entire spectrum, and within potentially diagnostic regions of the spectrum, instead of just being confined to grouped regions, in which the well-known vibrations of functional groups fall. These weak, often isolated features are not readily obvious to the untrained spectroscopist and may become crucial for spectral mineral identification when confronted with very noisy or highly fluorescent Raman spectra, or mixtures containing many mineral components. If the spectrum contains no identifiable peaks, the CNN algorithm has a classification option of “Dusty/Low SNR/Unknown”.

Table S1 in Supporting Information S1 shows the tabulated top five peaks from the 10-feature data set, for each mineral trained on the CNN, and the respective band assignments and vibrational modes. The top five peak positions were selected by taking a frequency distribution of the top five peak intensities in all spectra for each mineral and selecting the corresponding positions of the five peak intensities with the highest frequencies across the mineral class. While the CNN uses the entire Ricker CWT with 190 features, tabulating the top five peaks allows insights into the mechanics of how the CNN “thinks” about full spectrum matching. Not all primary peaks land in the top five for every mineral, but the Ricker CWT captures enough primary peaks to adequately influence the CNN node weights during classification. The weak peaks emphasized by the Ricker CWT provide identifying features throughout the entire spectrum, and within potentially diagnostic regions of the spectrum, instead of just being confined to grouped regions, in which the well-known vibrations of functional groups fall. These weak, often isolated features are not readily obvious to the untrained spectroscopist and may become crucial for spectral mineral identification when confronted with very noisy or highly fluorescent Raman spectra, or mixtures containing many mineral components. If the spectrum contains no identifiable peaks, the CNN algorithm has a classification option of “Dusty/Low SNR/Unknown”.

Table S1 in Supporting Information S1 shows the tabulated top five peaks from the 10-feature data set, for each mineral trained on the CNN, and the respective band assignments and vibrational modes. The top five peak positions were selected by taking a frequency distribution of the top five peak intensities in all spectra for each mineral and selecting the corresponding positions of the five peak intensities with the highest frequencies across the mineral class. While the CNN uses the entire Ricker CWT with 190 features, tabulating the top five peaks allows insights into the mechanics of how the CNN “thinks” about full spectrum matching. Not all primary peaks land in the top five for every mineral, but the Ricker CWT captures enough primary peaks to adequately influence the CNN node weights during classification. The weak peaks emphasized by the Ricker CWT provide identifying features throughout the entire spectrum, and within potentially diagnostic regions of the spectrum, instead of just being confined to grouped regions, in which the well-known vibrations of functional groups fall. These weak, often isolated features are not readily obvious to the untrained spectroscopist and may become crucial for spectral mineral identification when confronted with very noisy or highly fluorescent Raman spectra, or mixtures containing many mineral components. If the spectrum contains no identifiable peaks, the CNN algorithm has a classification option of “Dusty/Low SNR/Unknown”.

Table S1 in Supporting Information S1 shows the tabulated top five peaks from the 10-feature data set, for each mineral trained on the CNN, and the respective band assignments and vibrational modes. The top five peak positions were selected by taking a frequency distribution of the top five peak intensities in all spectra for each mineral and selecting the corresponding positions of the five peak intensities with the highest frequencies across the mineral class. While the CNN uses the entire Ricker CWT with 190 features, tabulating the top five peaks allows insights into the mechanics of how the CNN “thinks” about full spectrum matching. Not all primary peaks land in the top five for every mineral, but the Ricker CWT captures enough primary peaks to adequately influence the CNN node weights during classification. The weak peaks emphasized by the Ricker CWT provide identifying features throughout the entire spectrum, and within potentially diagnostic regions of the spectrum, instead of just being confined to grouped regions, in which the well-known vibrations of functional groups fall. These weak, often isolated features are not readily obvious to the untrained spectroscopist and may become crucial for spectral mineral identification when confronted with very noisy or highly fluorescent Raman spectra, or mixtures containing many mineral components. If the spectrum contains no identifiable peaks, the CNN algorithm has a classification option of “Dusty/Low SNR/Unknown”.
4.2. Labeling Test Models

Prior to testing our neural network with the clean gabbro, clean granite, dusty gabbro, dusty granite, and RRUFF test spectra, it was necessary to label the true mineral classifications for the spectra. Over 30,000 spectra were manually labeled by the authors over the span of 500 hr and three full passes. These true labels were used to verify the success rate of our classifying models. With each pass, the test set was run through the CNN and
the misclassified spectra often identified human mineral classification errors, effectively correcting any manual labeling mistakes.

### 4.3. Convolutional Neural Network for Raman Spectral Classification

We evaluated the performance of several supervised and unsupervised traditional classifiers and neural network classifiers using Raman spectra as inputs and compared them to convoluted neural networks. The inputs to the CNN consisted of a one-dimensional Raman spectrum fully sampled at regularly spaced wavenumber intervals. One version of the data included raw unprocessed spectra and the second version included spectra transformed through a Ricker CWT. Each kernel in the CNN was one-dimensional. A number of convolutional layers extracted features and two fully connected layers performed classification.

A CNN parameter optimization routine was implemented on the model to run through different values of all tunable hyperparameters and architecture parameters: learning rate, batch size, drop rate, number of layers, filter number and size, activation functions, optimizer functions, and number of epochs or number of training iterations. The investigated CNN parameters and resulting optimized tuned values are listed in Table 2. The parameter ranges were intentionally not limited until there was obvious performance degradation at either range extreme. Three types of CNN architectures were tested: bilayer, trilayer, and quadlayer, across a range of varying filter map numbers and shapes. Four different activation functions were examined during parameter tuning to optimize performance.

A CNN parameter optimization routine was implemented on the model to run through different values of all tunable hyperparameters and architecture parameters: learning rate, batch size, drop rate, number of layers, filter number and size, activation functions, optimizer functions, and number of epochs or number of training iterations. The investigated CNN parameters and resulting optimized tuned values are listed in Table 2. The parameter ranges were intentionally not limited until there was obvious performance degradation at either range extreme. Three types of CNN architectures were tested: bilayer, trilayer, and quadlayer, across a range of varying filter map numbers and shapes. Four different activation functions were examined during parameter tuning to optimize performance.

| Parameter | Range | Tuned value CWT | Tuned value raw |
|-----------|-------|----------------|----------------|
| Number of layers (Chollet, 2018; Goodfellow et al., 2016) | Bilayer, Trilayer, Quadlayer | Trilayer (confirms Liu et al., 2017 convention) | Trilayer (confirms Liu et al., 2017 convention) |
| Filter numbers (Chollet, 2018; Goodfellow et al., 2016) | 2, 4, 8, 16, 32, 64, 238 | 16, 32, 64 | 16, 32, 64 |
| Filter shape (Chollet, 2018; Goodfellow et al., 2016) | 3 × 1, 5 × 1, 11 × 1, 21 × 1, 31 × 1 | 21 × 1, 11 × 1, 5 × 1 | 21 × 1, 11 × 1, 5 × 1 |
| Activation function (Chollet, 2018; Bishop, 2006; Goodfellow et al., 2016) | Sigmoid, TanH, LeakyReLU, Softmax | LeakyReLU, TanH, Softmax | LeakyReLU, TanH, Softmax |
| LeakyReLU Alpha | 0.1–0.5 | 0.3 | 0.3 |
| Loss function (Chollet, 2018; Bishop, 2006; Goodfellow et al., 2016) | Sparse categorical cross entropy | Sparse categorical cross entropy | Sparse categorical cross entropy |
| Optimizer function (Bishop, 2006; Witten et al., 2011; Goodfellow et al., 2016; Chollet 2018) | SGD, ADAM, RMSProp | RMSProp | RMSProp |
| Learning rate (Witten et al., 2011; Bishop, 2006) | 0.0001–0.1 | 0.001 | 0.001 |
| Batch size (Witten et al., 2011; Bishop, 2006) | 1–500 | 8–32 | 8–32 |
| Drop out rate (Chollet, 2018; Bishop, 2006; Goodfellow et al., 2016) | 0.2–0.8 | 0.55 | 0.55 |

Note. Raw and CWT spectra performed well under similar parameters. CNN, convolutional neural network.

Figure 3. Raw Raman spectrum, trimmed to 150–800 cm⁻¹ to highlight Si-O-Si regions (left column), and spectrum passed through the Ricker CWT with top five peaks and intensities extracted for independent metrics comparisons with other statistical methods (right column) for quartz, albite, hornblende, and biotite. The Ricker CWT captures all the primary peaks traditionally used to identify silicates, unless they overlap one another, such as doublets and triplets. In the case of albite, which has a doublet in the 504-509 cm⁻¹ region, the more intense peak position is captured. The location of the most intense peak still allows identification of the feldspar endmembers albite, anorthite, and microcline, which have very similar spectra and Si-O-Si doublets. The Continuous wavelet transform pushes some non-peak portions of the data into negative intensities, which allows the peak-finding algorithm to better identify Raman peaks using a signal-to-noise (SNR) ratio threshold. Consequently, this improves the speed at which our convolutional neural network model converges. An artifact of the transform is that large peaks appear at the truncated ends of the spectra, and thus the highest and lowest frequency peaks within the transform are dropped before extracting Raman peak information. This edge effect biases the uppermost and lowermost ~50 cm⁻¹ of the spectra; the spectral regions that this bias effects is, of course, entirely determined by the spectral window chosen.
the weight values generated by each neuron. The number of epochs, or learning cycles, was varied from 1 to 500 to adjust how much time the model had to converge. Learning rates were varied to control how much the model changed in response to the estimated error each time the weights updated. Varied batch sizes were explored to control the number of training spectra that passed through the CNN before its internal parameters updated. CNN regularization was examined across a range of drop out rates to avoid strong dependencies between portions of adjacent layers, reduce overfitting, and improve generalization error.

At each node in a neural network, the inputs are multiplied by randomly assigned node weights and summed together. This value is referred to as the summed activation of the node and is transformed via an activation function into an output or node activation. A LeakyReLU nonlinearity algorithm was implemented as the default activation function because it allows the network to backpropagate even for negative values. This algorithm is also able to learn more complex relationships and structures in the data without saturating and locking very small values to 0 or −1 or very large values to 1. This feature avoids producing vanishing gradients and allows for the CNN to effectively learn. Near the end of the CNN the layers are fully connected in a dense layer and weights are passed through a TanH non-linearity function. At this juncture, most of the deep trainable spectral features have passed through the LeakyReLU activation in the feature extraction layers and the CNN generally benefits from any of the standard non-linear activation techniques. The final output layer implements a Softmax activation function with the same number of outputs as our number of mineral classes on which the training was conducted. Each individual training mineral or mineral mixture was manually encoded to its own integer (0–15). Softmax operates by standardizing the outputs to a probability range between 0 and 1, with the predicted class having the highest probability (Goodfellow et al., 2016). After each of the feature extraction layers, a 1-dimensional MaxPooling operation was implemented to downsample the input representation and extract broader and more general patterns that are more robust with respect to small changes in the input. MaxPooling downsamples by taking the maximum value over a spatial window (Keras Project, 2020). Batch normalization, dropout, and early stopping were implemented to avoid overfitting the model. Given that overfitting occurs when the CNN is learning the training samples, early stopping stops training when a performance drop is detected during validation (Goodfellow et al., 2016). Figure 4 outlines the CNN architecture used for Raman spectral mineral classification. This architecture adheres to similar conventions as deployed by Liu et al. (2017) in their single mineral spectra-oriented CNN.

Outputs from the Softmax regression were fed into a sparse categorical cross entropy loss function to address the multi-class classification problem of assigning one mineral or mineral mixture identification to each spectrum.
from the multi-mineralic gabbro and granite rocks. The loss, or difference between “true” target mineral identifications and CNN predicted identifications, was measured by this function. The loss function does this by calculating the average difference between the actual and predicted probability distributions for all classes in the problem and then minimizing the score (perfect cross entropy = 0) (Goodfellow et al., 2016). Three optimizers were explored to minimize the loss: stochastic gradient descent (SGD), adaptive moment estimation (ADAM), and RMSProp. SGD incorporates frequent model parameter updates and converges quickly, but it suffers from high variance in model parameters, can overshoot global minima, and requires slow learning rate reduction (Robbins & Monro, 1951). ADAM converges quickly (although not as quickly as SGD and RMSProp) and rectifies the vanishing learning rate and high variance of SGD, but it is computationally costly (Kingma & Ba, 2015). RMSProp incorporates adaptive model parameter updates like SGD, but it does so after every batch. It converges the fastest from this group of optimizers and reduces variance, but suffers from sensitivity to initial manual learning rate selection (Hinton, Srivastava, and Swersky, 2012). This mini-batch approach capitalizes on large training set sizes, and drastically improves CNN parameter finetuning.

Appropriate neural network parameter tuning is highly sensitive to the input training and testing datasets. We explored varying CNN parameter and hyperparameter effects on successful Raman spectra mineral identification. Our investigation led to slightly different finetuned network parameters and hyperparameters for optimal performance with preprocessed CWT Raman spectra relative to raw Raman spectra. These results are tabulated in columns 3 and 4 of Table 2, and are discussed in Section 5 along with the performance metrics.

4.4. Training the CNN on MLROD

Convolutional neural networks are data-intensive models that optimally work in the 10,000+ data input range and up into the millions and billions of inputs. We collected sufficient training spectra to adequately tap into the CNN capabilities. Alongside data volume, data variability is important for optimal neural network performance. The CNN needs to be able to generalize the spectral features it has learned during training to apply that knowledge to new, never before seen spectra such as the multi-phase granitic and gabbroic rocks used in this work or the RRUFF Project single-phase Raman spectra. We emphasized variability in spectral quality while acquiring the training sets for every mineral so that the CNN had a range of excellent to poor Raman spectra to train on for every mineral class. Training with large spectral variability is crucial for in situ Raman field measurements on Mars by autonomous rovers since dusty conditions, awkward analysis angles, topographically rough surfaces, and any other adverse environmental conditions are likely to compromise the quality of spectral measurements. The mineral sample micro- and macro-topography was used to ensure the laser was not always in focus while taking automated grid point grid measurements of the mineral surface. Additionally, grain boundaries and mineral defects added noise, internal reflections, baselines and overarching variability to the spectra. Figure 5 shows a randomly selected sample of 100 microcline training spectra from MLROD. The training set’s variability ranges from spectra with high noise, high fluorescence, and weak peak intensities, to those with excellent quality, low fluorescence, and very strong peaks. In some very low SNR cases, or cases with unconventional reflection geometries, the primary and lower-intensity peaks are broader, have similar heights, or missing peaks. Such spectra are part of what make MLROD unique, since machine learning algorithms like the proposed CNN can generalize full pattern matching with consideration toward missing information or varied spectral shapes for the same mineral. If low intensity peaks occur in areas that don’t overlap with peaks from different minerals, then those features can be diagnostic for mineral classification. Naturally, a large portion of the training spectra for each mineral do have conventional Raman spectra with the appropriate first and second order peak heights, positions, and cross sections. This allows the CNN to prioritize primary peaks (along with the full spectrum shape) for mineral identification so that lower intensity peaks are not confused for primary peaks of other minerals. Future training sets will consider additional mineral mixtures, and particularly those relevant to Mars materials. This training set

Figure 5. Random sampling of 100 spectra showing within class spectral variation for microcline. Spectra range from excellent, to good, to fair, to poor quality; they also include a range of baselines. Sharp randomly spaced lines are generated by cosmic rays.
enhancement will address situations where more complex mineral mixtures have peaks that overlap with primary and secondary peaks of minerals not included in the training set mixtures. Additionally, the broad suite of random cosmic rays across thousands of spectra effectively teaches a CNN to disregard whatever looks like a cosmic ray (very sharp peak and high local signal: noise) regardless of its occurrence frequency, and to focus on and identify the Raman spectrum. This makes the MLROD training set neutral to raw test sets with varying cosmic ray occurrences. Additionally, the training sets have sufficient data point density to interpolate up to ∼5 cm⁻¹ bins without signal loss in order to match lower resolution data from instruments such as the SuperCam spectrometer.

The Raman instrument’s laser interrogation area is 50 µm in diameter and the average gabbro and granite crystal sizes were >1 mm in diameter, allowing us to engage in Raman point counting techniques with few encounters of binary or ternary mineral mixtures in the Raman spectra. Three binary mixtures, one containing 50% quartz and 50% albite, one containing 50% forsterite and 50% augite, and one containing 50% forsterite and 50% augite were included in the training set to encompass the uncommon (for these sample geometries), but geologically critical, case that mineral mixtures were encountered within the spectral database. Practically, extrusive or finer grained rocks are more likely to encompass mixtures, and our mixture data allowed us to demonstrate that the CNN algorithm can readily recognize mixtures of multiple mineralogical components. This represents a key capability, and this demonstration clearly expands the application of CNN approaches from single-mineral identification (e.g., Liu et al., 2017) to polynmineratic rock characterization.

Rapid gradient descent within the loss function is necessary for rapid spectral training. We implemented the RMSprop optimizer to step through the gradients of our sparse categorical cross entropy loss function. This is a flexible learning rate method that adapts the learning rate to the parameter modifications: smaller updates (low learning rate) for parameters associated with frequently occurring features, and larger updates (high learning rates) for parameters associated with infrequent features. RMSprop is not privy to aggressive monotonically decreasing learning rates, works well with sparse datasets, and complements our CNN mini-batch approach by employing mini-batch SGD. The optimizer deals with vanishing or exploding gradients by using a moving average of squared gradients to normalize the gradient. This normalization balances the step size (momentum), decreasing the step for larger gradients to avoid exploding, and increasing the step for small gradients to avoid vanishing. Essentially, the adaptive learning rate becomes a hyperparameter that optimizes over time (Hinton et al., 2012).

5. Results and Discussion

We evaluated the CNN against the evaluation set to finetune the hyperparameters. Initial trials yielded an extremely fast training learning rate that took 1 and 2 epochs to reach 98% and 99% accuracy with a lagging validation set. Here, an epoch is defined in the neural network framework (by convention) as having the entire data set moving forward and backward through the network a single time. Several runs with increasing epoch sizes were tested until the training and validation set loss and accuracy curves converge near epoch 145 as seen on Figure 6a. This is a classic case of overfitting and a common issue with convoluted neural networks. In this case, overfitting was the result of having an extremely high batch parameter of 40,000. The training spectra would thus iterate 40,000 times through the CNN for every mineral, in every epoch.

Reducing the batch parameter to 8–32 and taking the mini-batch approach produced adequate convergence of our loss and accuracy curves, implying that the network is correctly classifying the validation set as quickly and accurately as it has learned the training data. Since the validation set is a subset of the training data, its features should be familiar, even though the CNN has not specifically trained on it. Figure 6b illustrates the updated learning curves. Notably, accuracy scores still converge at 98% and 99% within 2 epochs. Given the low dimensionality of 1-D spectra and the complexity of convoluted neural networks, this rapid convergence is expected. The best classification results were achieved by stopping training at 5 and 6 epochs—longer training caused generalization on the test set to begin to suffer.

5.1. CNN Classification Performance on Raman Spectra

The CNN was tested using Raman spectra of a roughly polished gabbro and granite slabs and of the same slabs covered in ∼50% (by coverage area) Hawaiian basaltic dust, which is likely to be similar to dust found on Mars. We also tested the CNN with an independently acquired set of Raman spectra from the RRUFF Project Raman
database. The RRUFF set included spectra from quartz, albite, muscovite, forsterite, augite, enstatite, and microcline. The granite test set included spectra from quartz, albite, hornblende, biotite, muscovite, and microcline. The gabbro test set included spectra from forsterite, augite, enstatite, anorthite, gypsum, and calcite. Together, all three test sets included every mineral in the training set of silicate minerals. Dusty Raman spectra where the underlying material was only partially obscured exhibited visually detectable peaks and broad bands at diminished intensities. The overall effect of this fine veneer of 50–150 μm-sized grains of dust does not invoke Raleigh or Mie scattering and only contributes optical scattering and eclipsing of laser light from interacting with the mineral surface. Hence, the predominant effects of dust coverage are to reduce the SNR, and this thus provides a limit of detection test for the fidelities of different spectral identification techniques.

We tested our proposed CNN mineral classifier on multi-mineralic gabbro and granite rock slabs and on the RRUFF mineral database and compared the success scores for various supervised and unsupervised classification methods across an average of 100 runs each. Regression, instance-based, decision tree, artificial neural network, and deep learning algorithms were employed to varying levels of success. The CNN outperformed the other methods and consistently scored >80%. Table 3 reports the weighted F1 scores across all mineral classes for 0% and 50% dust covered granite. Individual per class metrics are available in the Supplement section and highlight successful classification across all minerals. Reported F1 scores are weighted due to the natural test set imbalance resulting from randomized rock measurements.

A limitation to the current CNN performance lies in the difficulty associated with classifying intermediate mineral compositions situated within mineral group endmembers series such as in the plagioclase albite-anorthite or microcline-albite solid-solution series, unless a corresponding mineral training set that incorporates intermediate compositions is explicitly included. This is not a significant limitation with regards to CNN algorithm usage as an aid to scientist spectral interpretation, but is an aspect to recognize and address in future software iterations to better achieve onboard robotic autonomy. With the current training set comprising endmembers, the CNN typically forces intermediate compositions onto one endmember classification or the other. This forcing is

Figure 6. (a) An example of an overfitting convolutional neural network (CNN). Loss and accuracy curves do not converge until epoch 140+. The training set converges at 99% accuracy within 2 and 3 epochs while the validation set lags behind and does not fully converge until epoch 145. (b) Current working CNN learning curves converge quickly and accurately.
expected to typically converge on the more abundant endmember within the solid solution. By including training sets for intermediate compositions, successful CNN metrics should be attainable: as such, future improvements to CNN rock and mineral identification algorithms will likely require a comprehensive data set for common solid solution series, including olivines, pyroxenes, amphiboles, and micas in addition to feldspars. The success with feldspar endmember identification indicates that this is simply a matter of enhanced data inclusion within the training set, rather than a limitation on the algorithm itself. Our expectation is that accurate characterization of subtly different members of solid solution series may require reduction of the width of our bin size during the wavelet transform step, or exclusive use of raw data inputs. This expectation is generated by the comparatively small shifts in peak locations or splittings that occur across many solid solutions series (e.g., Bersani et al., 2018; Huang et al., 2000; Wang et al., 2001).

Importantly, if a spectrum contains no identifiable peaks, MLROD has spectrally passive data as part of its training set and the algorithm classifies the unknown test spectrum as a “Dusty/Low SNR/Unknown” option for feature-less spectra. The CNN algorithm will generally respond to minerals or mineral groups not included in the training sets by matching to the closest primary peaks, or the shape of the closest matching group of peaks within a particular region of the spectrum. For example, clays such as montmorillonite would default to mica or muscovite (matching the O-Al-O stretching region at ~650-700 cm⁻¹). Serpentines such as antigorite would default to either enstatite or augite (matching Si-O-Si vibrations at ~660–690 cm⁻¹). Serpentines could also default to muscovite if mixed with clays, and in rarer instances where the ~1,045 cm⁻¹ peak is of very high intensity, the CNN might default to calcite (attempting to match the CO₂ symmetric stretching vibration at 1,088 cm⁻¹) or gypsum (matching the SO₄ symmetric stretch at 1,008 cm⁻¹). Because of the generally modest variations in vibrational frequencies of the sulfate unit in different crystallographic environments, sulfates should, with the current training set, default to gypsum (matching the strong SO₄ vibration at 1,008 cm⁻¹) regardless of their hydration state since the MLROD training set does not currently span the 3,000–3,800 cm⁻¹ OH stretching region, nor does it encompass variably hydrated sulfates.

We further tested the influence of training set size per mineral class on success scores to determine a lower limit training set size for our particular neural network. Training sets increased by powers of two: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1,024, 2,048, 4,096, and 5,000. to the full training set minimums per class. At approximately 700 spectra per mineral class we gain adequate success rates of >80%, and by 5,000 spectra we gain success rates of 98% and 99%. Figure 7 shows weighted F1 scores using training sets of varying sizes.

The CNN was tested under various conditions in order to assess how its success might be reduced or slowed down. When tested on the independent RRUFF database spectra, it had an F1 success score of 96%, and even correctly identified potentially mislabeled or misidentified spectra in the database. More than half of the spectra labeled as enstatite on the database were identified as solid solutions; a few appeared to lie closer to the augite compositional field. This is not wholly unanticipated, as solid solutions are ubiquitous within pyroxenes (e.g., Morimoto, 1988). Moreover, the effect on Raman spectra of the calcium substitution relative to the shift in symmetry associated with the Pbca symmetry of enstatites/pigeonites relative to the C2/c symmetry of augites is complex (Huang et al., 2000; Wang et al., 2001). The CNN correctly labeled 92% of the pyroxene endmembers. We further replaced all training spectra with 5,000 synthetic spectra with random peaks and noise. The CNN replicated excellent learning curves between the training and validation sets but as expected from the randomized synthetic spectra, failed to predict any true spectra in the granite or RRUFF test sets. Consistent CNN behavior

---

### Table 3

|                        | K-means clustering | Random forest | Linear SVC | Logistic regression | MLP | CNN |
|------------------------|-------------------|---------------|------------|---------------------|-----|-----|
| Success Granite 0% Dust Coverage | 13.92%            | 45.82%        | 54.65%     | 63.90%              | 89.00% | 99.43% |
| Success Granite 50% Dust Coverage | 14.48%            | 39.34%        | 47.16%     | 33.33%              | 64.65% | 81.18% |
| Success Gabbro 0% Dust Coverage | 12.04%            | 23.58%        | 52.00%     | 58.02%              | 87.35% | 99.40% |
| Success Gabbro 50% Dust Coverage | 10.11%            | 20.07%        | 45.94%     | 51.13%              | 64.12% | 80.39% |

*Note. Feldspars were grouped into one category. Only quartz, albite, microcline, biotite, and hornblende were identified. The proposed CNN outperformed all models even under challenging dusty conditions. The CNN column values were formatted in bold to guide the reader's attention to the results for the primary technique proposed in this study. All of the values bold or not are weighted F1 scores. CNN, convolutional neural network; MLP, multi-layer perceptron.*
despite switching of data sets, and the high success scores using RRUFF Project test spectra verify robust CNN model performance in classifying gabbro and granite mineral components and exclude the presence of any potential biases introduced by our in-house Raman datasets.

In Figure 8, we show the CNN's classification performance. Plots for quartz, feldspar, hornblende, and biotite show examples of successfully labeled spectra with wide within-class variability of poor or dusty to excellent spectra. The top three classification probabilities are presented below the top plot, with the green spectra indicating the correct mineral classification label. The CNN reports the prediction certainty by outputting prediction scores. This also enables the user of the model to assess classification quality. All of these examples show correct classifications, but in cases where misclassification occurred, dark, Fe-rich minerals with similar spectra were confused for each other, or confusion emerged for minerals sharing similar primary or secondary peaks like calcite or gypsum. Mafic minerals are typically iron-rich with low scattering cross-sections that generate weak peaks or poorer SNR spectra at fast integrations. However, hornblende was rarely misclassified as augite, and biotite was rarely misclassified as muscovite due to their distinctive Raman spectra. Feldspars were minimally misclassified within the mineral group. Occasionally, pure minerals were misclassified as mixed mineral spectra where one of the binary components was a correct match, such as between forsterite and augite and forsterite/augite mix or between quartz and albite and quartz/albite mix. Such misclassifications, classifications with intermediate confidence scores, and subsequent classifications that fall within intermediate confidence scores, also may indicate that the CNN has converged on a different member of the correct, or closely related, mineral family. This is a likely outcome for a mineral that is not present in the teaching set.

These results reflect typical challenges that a human observer would face when interpreting Raman spectra. Darker minerals exhibit greater internal absorption of common Raman scattering laser wavelengths (e.g., 530 nm), and yield smaller Raman cross sections due to reduced Raman scattering exiting the mineral volume. Additionally, lower symmetries of minerals such as amphiboles tend to contribute to smaller polarizability tensors, and hence lower Raman scattering intensities. This complicates spectral identification, as few to no notable spectral features may appear in a mixed phase spectrum containing these minerals. Fluorescence and dust coverage impose similar effects to one another on Raman spectra, and these are areas that computational analysis and particularly, a properly trained and finetuned CNN, might excel at—potentially beyond all but the most experienced human abilities.

Receiver operating characteristic (ROC) curves for the CNN, multi-layer perceptron, and random forest binary classifiers (Figure 9) show the diagnostic ability of each classifier as their discrimination thresholds are varied. The true positive rate, also known as the sensitivity or recall, shows the probability of an actual positive classification while the false positive rate, also known as fallout, shows the probability of a false alarm during classification. These curves are for ungrouped mineral classes to illustrate the individual class label performance. While most algorithms classified intense or distinctive Raman scatters that are easy to identify such as quartz, muscovite, and forsterite, the CNN was able to successfully classify mafic minerals like biotite and hornblende, and better discriminate between minerals in silicate mineral groups such as the feldspars (albite, anorthite, microcline). CNN ROC curves for the RRUFF data are shown in the Supplement section, and illustrate almost perfect classifications, with the exception of the pyroxenes. The discrepant enstatite ROC curve in the Supplement is generated by a subset of spectra that have Raman spectra that appear more similar to augite, and hence may either
be misclassified in the RRUFF database or may be compositionally more complex than their assigned mineral name implies.

Preliminary CNN runs of SuperCam Raman measurements from Mars on Sol 152 (released on the NASA Planetary Data System Geosciences Node (https://pds-geosciences.wustl.edu/missions/mars2020/supercam.htm)) indicate our CNN algorithm can garner useful insight on the presence of silicate minerals. SuperCam peaks in the Si-O-Si vibrational region contributed to quartz and plagioclase predictions. Expansion of the MLROD

Figure 8. Convolutional neural network successful classifications with average top three probability assignments. Variability in the amphibole spectra (particularly) is likely to be also affected by the different orientations of randomly positioned samples.

be misclassified in the RRUFF database or may be compositionally more complex than their assigned mineral name implies.

Preliminary CNN runs of SuperCam Raman measurements from Mars on Sol 152 (released on the NASA Planetary Data System Geosciences Node (https://pds-geosciences.wustl.edu/missions/mars2020/supercam.htm)) indicate our CNN algorithm can garner useful insight on the presence of silicate minerals. SuperCam peaks in the Si-O-Si vibrational region contributed to quartz and plagioclase predictions. Expansion of the MLROD
training sets to include clay, serpentine, and sulfate minerals and mixed minerals will further enhance the variety of mineral identifications from SuperCam data.

5.2. Implications of Raman Spectral Data on CNN Parameters and Scientific Applications

A properly tuned, general-purpose CNN with static architecture and hyperparameters, trained on Raman spectra with sufficient within-class variability to encompass the complexity of rock and mineral Raman spectra, can identify Raman spectra that have matches in the training set. Importantly, the CNN can also generalize the results that it has learned to identify minerals within (e.g.,) mixed-phase aggregates or rocks that it has not seen before. The CNN model and high-volume machine learning training sets presented in this work succeed in identifying minerals even in dusty, low SNR Raman spectra. Proper machine learning training sets are critical for effective deployment of CNN models in spectroscopic mineral analysis, and should be the first and foremost consideration before developing deep learning algorithms for spectral classification. With this in mind, base CNN performance can be further optimized depending on the test set characteristics. Several models are proposed at the interface of Raman spectroscopy, deep learning, and planetary science, but few describe the results of model exploration and development. This is useful information that can be gleaned from model error and misclassifications. Based on insights derived from our current modeling, general guidance for custom-dataset CNN model development for Raman spectra mineral classification can be derived:

- Spectra from low SNR samples or low Raman scatterers such as dark mafic minerals (biotite, hornblende, augite) benefit from increased neural network depth, preprocessing, or both. This involves increasing the number of learning filter kernels in each feature extraction layer (detector) or applying a transform like the Ricker CWT to the training and test sets, respectively. Neural network simplicity is generally preferred, thus the more extreme measure of adding feature extraction layers should be considered as a last resort for spectra from particularly challenging lithologies.
• Spectra with pronounced baseline slopes or strong fluorescence that do not mask primary peaks also benefit from increased neural network depth, preprocessing, or both. As such, it is important to include a wide range of baselines for each mineral training set.

• Spectra from samples exhibiting very closely clustered primary Raman peaks such as feldspar doublets and triplets in the 500-520 cm\(^{-1}\) Si-O-Si bending-stretching region, or with similar mineralogies and continuously varying chemical compositions, such as in feldspar or olivine solid solution series, benefit from raw spectral inputs so as to not lose spectral fidelity.

• Mineral mixture spectra also benefit from raw spectral inputs and no preprocessing, but it is important to include a wide range of abundance ratios since the amplitudes of Raman peaks for mineral mixtures may be nonlinear.

• Spectra from high SNR and very Raman-active minerals such as gypsum, calcite, quartz, and feldspars do not require increased neural network depth for successful CNN models. Datasets that involve such high-quality spectra can benefit from CNN models with detector simplifications via smaller learning filter convolutions and wider filter maps.

• Higher drop out rates lead to improved generalization for smaller training sets or training sets with minimal within-class diversity. As training set size and within-class diversity increases, the parameter importance decreases, and a balanced dropout rate of 50%–55% is sufficient.

• Larger batch sizes lead to lower asymptotic test accuracy. The largest possible batch size is equal to the total size of the smallest training set class or mineral. This lost test accuracy can be recovered from a larger batch size by increasing the learning rate. This comes at the expense of both time and computational cost. A mini-batch approach with adaptive learning rate optimizers like RMSProp is preferred for 1-dimensional Raman spectra.

• Larger batch sizes are also very sensitive to the particular spectra drawn from the data set. Such models can fluctuate between very large or very small gradient updates, leading to learning inconsistency. Using small batch sizes allows the model to reduce sensitivity to the data subset and thus make updates that are roughly the same size. Smaller batch sizes can capture and learn from within-class training set variability.

6. Conclusion

We present an open-sourced, accessible, CNN algorithm capable of classifying Raman spectra of silicate, carbonate, and sulfate minerals with outstanding performance even under challenging, low-signal conditions, such as dust-laden or poorly crystallized rocks and minerals on the surface of Mars. An accompanying machine learning-specific Raman spectra training (MLROD) set is presented and made available to the scientific community through the Open Data Repository. We expect that this CNN algorithm can be readily generalized to other mineral groups and complex solid-solution series given proper training set expansion. General CNN development guidelines are provided to adjust the model for custom-dataset-specific needs with an awareness of Raman spectral complexities and acquisition challenges.

This is among the first known Raman spectral implementation of upwards of 5,000 training spectra per mineral label and upwards of 70,000 total training spectra. CNN classification potential was adequately tapped and the classifier's success greenlights the way for increasing the complexity of Raman spectral inputs to this type of algorithm. The pattern recognition abilities of convoluted neural networks have the potential of discriminating between spectra of adjacent mineral compositions within solid solution series and mineral groups, identifying mineral mixtures, as well as differentiating between physical processes that mimic out-of-class spectral behavior, and actually identifying spectral peaks. These qualities can facilitate Raman spectral interpretation both on Earth and Mars, and hold the prospect of increasing scientific yield and potentially correcting human classification error—particularly for relatively poor Raman scatterers or low signal-to-noise spectra. Given the current NASA Mars 2020 and upcoming ESA 2022 ExoMars missions, deploying planetary surface rovers with Raman spectroscopic capabilities that have onboard CNN classifiers could facilitate autonomous rover decision-making, enabling scientists to delegate and automate increasingly complex tasks to the rover, saving time and resources.
Data Availability Statement

The Machine Learning Raman Open Data set (MLROD) supporting the conclusions in this paper can be obtained at the Open Data Repository: https://doi.org/MLROD (https://doi.org/10.48484/pwrb-r137) as well as in the NASA Astrobiology Habitable Environments Database: https://ahed.nasa.gov/datasets/f5b6051bef18c5a7e-ae6f504582 (Berlanga, Temiquel, & Williams, 2022b; https://doi.org/10.48667/7v5q-2606). The latest version of the source code and supporting scripts can be obtained on GitHub through Zenodo: https://doi.org/10.5281/zenodo.7036374 (GitHub: https://github.com/GenTeML/Spec-CNN). The code can be compiled under any local or browser-based Python environment on Mac, Windows, or Linux. Source code and datasets are copyrighted under Creative Commons BY-NC. Programming was completed using Python 3.8 with Jupyter Notebooks on the Google Collaboratory development environment. Data was processed using the Pandas and Numpy packages (Harris et al., 2020; Reback et al., 2021). Traditional statistics and spectral signal processing were done using the SciKit Learn package (Pedregosa et al., 2011). Machine learning algorithms were developed using the Keras library on top of Tensorflow (Abadi et al., 2016).

References

Abadi, M., Agarwal, A., Paul, B., Brevdo, E., Chen, Z., Craig, C., et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems.
Artur, C., Le Ru, E. C., & Echegoin, P. G. (2011). Temperature dependence of the homogeneous broadening of resonant Raman peaks measured by single-molecule surface-enhanced Raman spectroscopy. Journal of Physical Chemistry Letters, 2(23), 3002–3005. https://doi.org/10.1021/pj2017787
Back, S.-J., Park, A., Ahn, Y.-J., & Choi, J. (2014). Baseline correction using asymetrically reweighted penalized least squares smoothing. Analyst, 140(1), 250–257. https://doi.org/10.1039/c4an01061b
Berlanga, G., Acosta-Maeda, T. E., Sharma, S. K., Porter, J. N., Dera, P., Shelton, H., et al. (2019). Remote Raman spectroscopy of natural rocks. Applied Optics, 58(32), 8971. https://doi.org/10.1364/aos.58.008971
Berlanga, G., Temiquel, N., & Williams, Q. (2022a). MLROD Raman CNN (versión 1.0.0-beta) [Source Code]. Zenodo. https://doi.org/10.5281/zenodo.7036374
Berlanga, G., Temiquel, N., & Williams, Q. (2022b). Machine Learning Raman Open Dataset (MLROD) [Dataset]. Astrobiology Habitable Environment Database. https://doi.org/10.48667/YV5Q-2606
Berlanga, G., Williams, Q., & Temiquel, N. (2022). Machine Learning Raman Open Dataset (MLROD) [Dataset]. Open Data Repository. https://doi.org/10.48484/PWBR-R137
Bersani, D., Aliais, I., Tribaudino, M., Mantovani, L., Benisek, A., Carpenter, M. A., et al. (2018). Plagioclase composition by Raman spectroscopy. Journal of Raman Spectroscopy, 40(4), 684–698. https://doi.org/10.1002/jrs.5340
Bhartia, R., Beegle, L. W., DeFlores, L., Abbey, W., Hollis, J. R., Uckert, K., et al. (2021). In Perseverance’s scanning habitable environments with Raman and luminescence for organics and chemicals (SHERLOC) investigation (Vol. 217).
Bish, D., Blake, D., Vaniman, D., Sarrazin, P., Bristow, T., Achilles, C., et al. (2014). The first x-ray diffraction measurements on Mars. IUCr, 1(6), 514–522. https://doi.org/10.1107/02525514021150
Bish, D. L., Blake, D. F., Vaniman, S. J., Chipera, S. J., Morris, R. V., Ming, D. W., et al. (2014). X-ray diffraction results from Mars. Science, 1–6.
Bishop, C. M. (2006). In M. Jordan, J. Kleinberg, & B. Scholkopf (Eds.), Pattern recognition and machine learning. Springer Science+Business Media, LLC.
Bishop, J. L., ScottMurchie, L., Pieters, C. M., & Zent, A. P. (2002). A model for formation of dust, soil, and rock coatings on Mars: Physical and chemical processes on the Martian surface. Journal of Geophysical Research, 107, 1–17. https://doi.org/10.1029/2000je001581
Blackberg, J., Alrestam, E., Cochrane, C. J., Maruyama, Y., & Farmer, J. D. (2020). Miniature high-speed picosecond Raman spectrometer for identification of minerals and organics in Planetary Science. Applied Optics, 59(2), 433–444. https://doi.org/10.1364/ao.59.000433
Blake, D. F., Morris, R. V., Kocurek, G., Morrison, S. M., Downs, R. T., Bish, D., et al. (2013). Curiosity at Gale Crater, Mars: Characterization and analysis of the rocknest sand shadow. Science, 341(6153), 1–8.
Boucher, T. F., MarieOzanne, V., Carmosino, M. L., Dyr, M. D., Mahadevan, S., Breves, E. A., et al. (2015). Spectrochimica acta Part B a study of machine learning regression methods for major elemental analysis of rocks using laser-induced breakdown spectroscopy. Spectrochimica Acta Part B: Atomic Spectroscopy, 107, 1–10. https://doi.org/10.1016/j.sab.2015.02.003
Breitenfeld, L. B., Darby Dyar, M., Carey, C. J., Tague, T. J., Wang, P., Terry, M., & Parente, M. (2018). Predicting olivine composition using Raman spectroscopy through band shift and multivariate analyses. American Mineralogist, 103(11), 1–48. https://doi.org/10.2138/amu-2018-691
Carey, C., Boucher, T., Mahadevan, S., Bartholomew, P., & Dyar, M. D. (2015). Machine learning tools for mineral recognition and classification from Raman spectroscopy. Journal of Raman Spectroscopy, 46(10), 894–903. https://doi.org/10.1002/jrs.4757
Cerniti, G., Karunatilakke, S., Zhao, Y. Y. S., Meslin, P. Y., Cousin, A., Hood, D. R., & Scaleghi, R. (2020). Disambiguating the soils of Mars. Planetary and Space Science, 186, 104922. https://doi.org/10.1016/j.pss.2020.104922
Chollert, F. (2018). In T. Arritola, J. Gaines, A. Dragosavljevic, & T. Taylor (Eds.), Deep learning with python. Manning Publications.
Du, P., Kibbe, W. A., & Lin, S. M. (2006). Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. Bioinformatics, 22(17), 2059–2065. https://doi.org/10.1093/bioinformatics/btl355
Emmanuel, I., & Stanier, C. (2016). Defining big data. In ACM International Conference Proceeding Series.
Farias, E., Jens, E., Barry, N., Kempeanar, J., & Novak, K. (2020). Medieval thermal design and validation of Mars 2020 gas dust removal tool (GDRT). In 49th International Conference on Environmental Systems.
Fauré, K. A., Williford, K. H., Stack, K. M., Bhartia, R., Chen, A. L., de la Torre, M., et al. (2020). Mars 2020 mission overview. Space Science Reviews, 216(8), 142. https://doi.org/10.1007/s11214-020-00762-y
Foucher, F., Lopez-Reyes, G., Bost, N., Rull-Perez, F., Rubmann, R., & Westall, F. (2013). Effect of grain size distribution on Raman analysis and the consequences for in situ planetary mission. Journal of Raman Spectroscopy, 44(6), 916–925. https://doi.org/10.1002/jrs.4307
Robbins, H., & Monro, S. (1951). A stochastic approximation method. The Annals of Mathematical Statistics, 22(3), 400–407. https://doi.org/10.1214/aoms/1177729586

Rogers, A. D., & Hamilton, V. E. (2015). Compositional provinces of Mars from statistical analyses of TES, GRS, OMEGA, and CRISM data. Journal of Geophysical Research: Planets, 120, 62–91. https://doi.org/10.1002/2014je004690

Ruckstuhl, A. F., Jacobson, M. P., Field, R. W., & Dosh, J. A. (2001). Baseline subtraction using robust local regression estimation. Journal of Quantitative Spectroscopy and Radiative Transfer, 68(2), 179–193. https://doi.org/10.1016/s0022-4073(00)00212-2

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learner representations by back-propagating errors. Nature, 323(6088), 533–536. https://doi.org/10.1038/323533a0

Schmidt, M. E., Glynis, P. M., Bray, S. L., Bradley, N. J., Lee, R. E., Berger, J. A., et al. (2018). Dusty rocks in Gale Crater: Assessing areal coverage and separating dust and rock contributions in APXS analyses. Journal of Geophysical Research: Planets, 123, 1649–1673. https://doi.org/10.1029/2018je005553

Schulze, G., Jraaek, A., Marcia Yu, M. L., Lim, A., Turner, F. B., & Blades, M. W. (2005). Investigation of selected baseline removal techniques as candidates for automated implementation. Applied Spectroscopy, 59(5), 545–574. https://doi.org/10.1366/000370205773948958

Sharma, S., Mannmone, J., & Nicol, M. (1981). Ring configuration in vitreous silica—a Raman spectroscopy investigation. Nature, 292(5819), 140–141. https://doi.org/10.1038/292140a0

Sharma, S. K., Angel, S. M., Ghosh, M., Hubble, H. W., & Lacey, P. G. (2002). Remote pulsed laser Raman spectroscopy system for mineral analysis on planetary surfaces to 66 meters. Applied Spectroscopy, 56(6), 699–705. https://doi.org/10.1366/0003702027680077630

Sharma, S. K., & Egan, M. (2019). Raman spectroscopy: Theory and laboratory spectra of geologic materials. In J. L. Bishop, J. F. Bell, & J. E. Moersch (Eds.), Remote Compositional analysis techniques for understanding spectroscopy, mineralogy, and geochemistry of planetary surfaces (pp. 120–146). Cambridge University Press.

Sobron, P., Misra, A. K., Rull, F., & Sansano, A. (2019). Raman spectroscopy: Field measurements. In J. L. Bishop, J. F. Bell, & J. E. Moersch (Eds.), Remote compositional analysis techniques for understanding spectroscopy, mineralogy, and geochemistry of planetary surfaces (pp. 274–286). Cambridge University Press.

Stopar, J. D., Lucey, P. G., Sharma, S. K., Misra, A. K., Taylor, G. J., & Hubble, H. W. (2005). Raman efficiencies of natural rocks and minerals: Performance of a remote Raman system for planetary exploration at a distance of 10 meters. Spectrochimica Acta A, 61, 2315–2323. https://doi.org/10.1016/j.saa.2005.02.030

Tuschel, D. (2017). Why are the Raman spectra of crystalline and amorphous solids different? Spectroscopy (Santa Monica), 32(3), 26–33.

Wang, A., Freeman, J. J., & Jolliff, B. L. (2015). Understanding the Raman spectral features of phyllosilicates. Journal of Raman Spectroscopy, 46(10), 829–845. https://doi.org/10.1002/jrs.6480

Wang, A., Jolliff, B. L., & Haskin, L. A. (1999). Raman spectroscopic characterization of a highly weathered basalt: Igneous mineralogy, alteration products, and a microorganism. Journal of Geophysical Research, 104(E1), 27067–27077. https://doi.org/10.1029/1999je001071

Wang, A., Jolliff, B. L., Haskin, L. A., Kuebler, K. E., & Viskupic, K. M. (2001). Characterization and comparison of structural and compositional features of planetary quadrilateral pyroxenes by Raman spectroscopy. American Mineralogist, 86(7–8), 790–806. https://doi.org/10.2138/am.2001-0703

Wiens, R. C., Maurice, S., Robinson, S. H., Nelson, A. E., Cais, P., Bernardi, P., et al. (2021). The SuperCam instrument suite on the NASA Mars 2020 rover: Body unit and combined system tests. Space Science Reviews, 217(1).

Wiens, R. C., Maurice, S., & Rull Perez, F. (2017). The SuperCam remote sensing instrument suite for the Mars 2020 rover mission: A preview. Spectroscopy, 32(5), 50–55.

Williams, Q. (1995). Infrared, Raman and optical spectroscopy of Earth materials. In T. J. Ahrens (Ed.), Mineral Physics & crystallography: A Handbook of Physical constants (Vol. 2, pp. 291–302).

Witten, I. H., Frank, E., & Hall, M. A. (2011). Data mining practical machine learning tools and techniques (3rd ed.). Morgan Kaufmann Publishers.

Yang, C., He, Z., & Yu, W. (2009). Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. BMC Bioinformatics, 10, 1–13. https://doi.org/10.1038/1471-2105-10-4

References From the Supporting Information

Buzgar, N., & Apopei, A. I. (2009). The Raman study on certain carbonates. Analele Stiintifice de Universitatii Al Cazan din Iasi, 55(2), 97–112.

Buzgar, N., Apopei, A. I., Diaconu, V., & Buzatu, A. (2013). The composition and source of the raw material of two stone axes of Late Bronze Age from Neamt county (Romania): A Raman study. Analele Stiintifice de Universitatii Al Cazan din Iasi, 59(1), 5.

Chio, C. H., Sharma, S. K., Lucey, P. G., & Meunow, D. W. (2005). Effects of particle size and laser-induced heating on the Raman spectra of alpha quartz grains. Applied Spectroscopy, 57(7), 774–783. https://doi.org/10.1366/0003702032032210852

Freeman, J. J., Wang, A., Kuebler, K. E., Jolliff, B. L., & Haskin, L. A. (2008). Characterization of natural feldspars by Raman spectroscopy for future planetary exploration. The Canadian Mineralogist, 46(6), 1795–1818. https://doi.org/10.7349/canmin.46.6.1477

Gunasekaran, S., Anbalagan, G., & Pandi, S. (2006). Raman and infrared spectra of carbonates of calcite structure. Journal of Raman Spectroscopy, 37(9), 892–899. https://doi.org/10.1002/jrs.1518

Kolesov, B. A., & Geiger, C. A. (2004). A Raman spectroscopic study of Fe-Mg olivines. Physics and Chemistry of Minerals, 31(3), 142–154. https://doi.org/10.1007/s12013-003-0137-y

Kuebler, K. E., Jolliff, B. L., Wang, A., & Haskin, L. A. (2006). Extracting olivine (Fo-Fu) compositions from Raman spectral peak positions. Geochimica et Cosmochimica Acta, 70(24), 6201–6222. https://doi.org/10.1016/j.gca.2006.07.035

Lam, P. K., Yu, R., Lee, M. W., & Sharma, S. K. (1990). Structural distortions and vibrational modes in MgSiO3. American Mineralogist, 75, 109–119.

McKeown, D., Bell, M., & Elz, E. (1999). Vibrational analysis of the dioctahedral mica: 2M1 muscovite. American Mineralogist, 84(7–8), 1041–1048. https://doi.org/10.2138/am-1999-7-806

McKeown, D. A. (2005). Raman spectroscopy and vibrational analyses of albite: From 25°C through the melting temperature. American Mineralogist, 90(10), 1506–1517. https://doi.org/10.2138/am.2005.1726

Mouri, T., & Enami, M. (2008). Raman spectroscopic study of olivine-group minerals. Journal of Mineralogical and Petrological Sciences, 103(2), 100–104. https://doi.org/10.2465/jmps.071015

Sharma, S. K., Philpotts, J. A., & Matson, D. W. (1985). Ring distributions in alkali- and alkaline-Earth aluminosilicate framework glasses—a Raman spectroscopic study. Journal of Non-Crystalline Solids, 71(1), 403–410. https://doi.org/10.1016/0022-0201(85)90311-4
Singha, M., & Singh, L. (2018). Vibrational spectroscopic study of muscovite and biotite layered phylosilicates. *Indian Journal of Pure & Applied Physics, 54*(2), 116–122.

Uckert, K., Bhartia, R., Beegle, L. W., Monacelli, B., Asher, S. A., Burton, A. S., et al. (2021). Calibration of the SHERLOC deep ultraviolet fluorescence-Raman spectrometer on the perseverance rover. *Applied Spectroscopy, 75*(5), 763–773. https://doi.org/10.1177/00037028211013368

Wang, S. Y., Sharma, S. K., & Cooney, T. F. (1993). Micro-Raman and infrared spectral study of forsterite under high pressure. *American Mineralogist, 78*, 469–476.

White, S. N. (2008). Laser Raman spectroscopy as a technique for identification of seafloor hydrothermal and cold seep minerals. *Chemical Geology, 259*(3–4), 15589–16252. https://doi.org/10.1016/j.chemgeo.2008.11.008