Computers & Geosciences 54 (2013) 259–268

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

An automated mineral classifier using Raman spectra

Sascha T. Ishikawa [a,b,*], Virginia C. Gulick [a,c]

[a] Space Sciences Division, NASA Ames Research Center, MS 239-20, Moffett Field, CA 94035, United States
[b] University of California Santa Cruz, Santa Cruz, CA 95064, United States
[c] SETI Institute, 189 Bernardo Avenue, Mountain View, CA 94043, United States

ARTICLE INFO

Article history:
Received 2 September 2012
Received in revised form
26 January 2013
Accepted 29 January 2013
Available online 5 February 2013

Keywords:
Mineral classification
Raman spectroscopy
Machine learning
Mars
Robotic exploration
Igneous rocks

ABSTRACT

We present a robust and autonomous mineral classifier for analyzing igneous rocks. Our study shows that machine learning methods, specifically artificial neural networks, can be trained using spectral data acquired by in situ Raman spectroscopy in order to accurately distinguish among key minerals for characterizing the composition of igneous rocks. These minerals include olivine, quartz, plagioclase, potassium feldspar, mica, and several pyroxenes. On average, our classifier performed with 83 percent accuracy. Quartz and olivine, as well as the pyroxenes, were classified with 100 percent accuracy. In addition to using traditional features such as the location of spectral bands and their shapes, our automated mineral classifier was able to incorporate fluorescence patterns, which are not as easily perceived by humans, into its classification scheme. The latter was able to improve the classification accuracy and is an example of the robustness of our classifier.

1. Introduction

The characterization of geological features provides insight into the history and formation of a site. This process requires the following: (1) detailed morphological, physical, and compositional analyses of rocks and sediments within stratigraphic layers, units, outcrops, and landforms, and (2) the integration of this information into conceptual models used to unravel a region's geologic history. Unfortunately, exploring some geological sites may be difficult or even dangerous, as in the case of deep-sea hydrothermal vent studies on Earth, or in the remote exploration of the surface of Mars. As a result, these kinds of missions would benefit from assistance, or even full autonomy in the form of robotic explorers capable of performing onboard scientific analyses.

A key task in characterizing the geology of site is selecting which rocks and sediments to analyze. Our efforts to facilitate this process are towards developing technology to serve as a field assistant for geologists, as well as to equip robotic explorers with the capacity to perform unassisted compositional analyses of geologic samples.

In this paper, we provide an application of machine-learning techniques to create an automated classifier to estimate the presence of key minerals based on in situ Raman spectroscopy. We demonstrate the robustness of our method by using different sources of Raman spectral data, both from our own rock and mineral database, and from a third-party mineral database.

1.1. Raman spectroscopy

Raman spectroscopy provides a quick and non-invasive technique (Lewis and Edwards, 2001) to determine the mineral compositions of rocks and sediments. Samples are scanned with a laser without the need for preparation, such as grinding or pulverizing. Therefore, this approach enables the ability to analyze samples as they are found, without moving or altering them.

Raman spectroscopy works by directing a focused monochromatic light source at the surface of a mineral causing individual photons to scatter. Most of the photons will bounce off elastically with a negligible transfer of energy, a phenomenon known as Rayleigh scattering. However, a small portion of photons will be scattered inelastically with a corresponding transfer of energy between the photon and the mineral surface. The latter is called Raman scattering and is observed as a shift in wavelength between the incident light and the scattered light, which is recorded by a detector. The Raman shift occurs because energy from the incident photons is transformed into vibrational and rotational motions of surface molecules. Moreover, this shift in wavelength reveals important information about the molecular structure of minerals and it often provides a unique signature of the material being analyzed (Lewis and Edwards, 2001). The resulting "fingerprint" spectra enable us to automate the

* Corresponding author at: NASA Ames Research Center, Space Sciences Division, Mail Stop 239-20, Moffett Field, CA 94035, United States.
Tel.: +1 415 595 4996.
E-mail address: Sascha.T.Ishikawa@nasa.gov (S.T. Ishikawa).

0098-3004/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2013.01.011

process of classifying minerals by using machine learning algorithms that exploit the uniqueness of such signatures.

1.2. The minerals

The presence of certain key minerals reveals important compositional information for determining the geologic origin of a rock. The minerals we used in this study help to discriminate between felsic and mafic igneous rocks. Mafic igneous rocks contain higher percentages of the minerals biotite, olivine, pyroxene, amphibole, and plagioclase feldspar than their felsic counterparts. Therefore, they are enriched in iron, magnesium, calcium and sodium, but are lower in silica. In contrast, felsic igneous rocks contain higher percentages of the minerals quartz, muscovite, and potassium feldspars and are thus enriched in silica, aluminum, and potassium.

A single mineral is often a subset of a larger mineral group. Individual minerals in a group, while different in their exact structure, may share a common chemical formula. For example, microcline ($KAlSi_3O_8$) is a lower temperature polymorph of orthoclase; although it has the same chemical formula, it possesses a different structure. Both are potassium-rich alkali feldspars, or *K*-spars. Minerals in the plagioclase group have a general chemical formula (Na, Ca)$Al_{1-2}Si_{2-3}O_8$ where the amounts of aluminum and silicon vary based on which end member (Na or Ca) is present. These minerals range from a sodium-rich albite ($NaAlSi_3O_8$) through oligoclase, andesine, labradorite, bytownite, and calcium-rich anorthite ($CaAl_2Si_2O_8$). A similar relationship extends to the remaining mineral groups considered in this study. The shared characteristics among minerals carry over to their spectral signatures, which in some cases are virtually indistinguishable in terms of individual spectra. See for example, microcline versus orthoclase in Fig. 1. Based on this ambiguity, we would expect poor resolution between certain minerals in our classifier. Instead, we find that our classifier is able to distinguish between some minerals despite the remarkable similarities in their spectra. Fig. 1 illustrates the spectral profiles of all six mineral groups considered in this study. Each individual curve represents an average of ten normalized spectral observations of a mineral specimen from the training set.

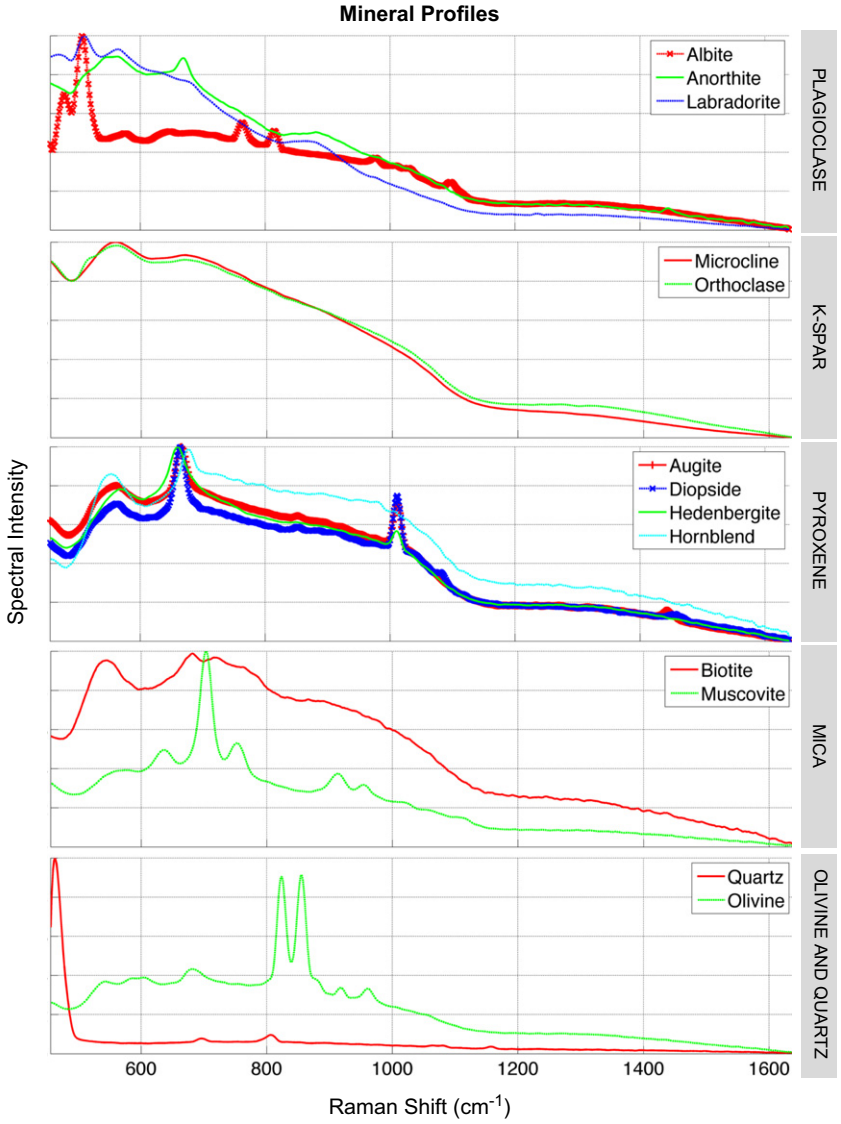

Fig. 1. Typical mineral spectra from our collection used in this study. Each plot was obtained by averaging all spectral observations for each mineral type, normalizing the spectral intensity values to the range [0,1], and smoothing with a Savitzky–Golay filter. The plots show the Raman bands associated with specific mineral groups as well as their broader fluorescence bands.

2. Previous studies

MSL's Chemistry and Mineralogy instrument (CheMin) (Blake, 2012; Blake et al., 2012) analyzes the mineral content of rocks and sediments on Mars, including olivine and pyroxene, using X-ray diffraction. Although this method produces accurate results, it requires samples to be collected and then pulverized prior to chemical analysis. CheMin can analyze up to 24 pristine samples or up to 74 samples with re-use of sample cells over the duration of the mission (Blake et al., 2012). Each analysis can take up to 10 h. However, MSL's Chemistry & Camera (ChemCam) provides rapid elemental analysis of targets within 7 m of the rover (Wiens et al., 2012). This instrument package consists of a Laser-Induced Breakdown Spectrometer (LIBS) and a Remote Micro-Imager (RMI). The chemistry is measured by LIBS and the RMI places the LIBS data into a geomorphological context.

Raman spectroscopy provides rapid, non-invasive mineralogical analyses. For example, Haskin et al. (1997) used a manual "point cloud" method for quantifying the relative abundance of minerals in rock samples using in situ Raman spectroscopy. Acquisition times were reported to be between 30 and 45 s per spectrum. This approach is appealing because it provides a method for compositional analysis of samples as they are found, without alteration. Furthermore, the mineralogical analysis can be performed an unlimited number of times, and can be automated.

Traditional methods of automatic or semi-automatic Raman analysis use properties of observed spectral bands, such as shape and peak location, by fitting functions (e.g., polynomials or splines) to the bands in order to extract relevant features, which are then used to compare spectra. While intuitive and straightforward to implement, there are several disadvantages to using the latter techniques. First, background effects such as fluorescence can often mask important spectral bands, so removing the background is usually a prerequisite to feature extraction. Unfortunately, applying the common shape-fitting background removal techniques on spectra may introduce unwanted variance to the data (Lewis and Edwards, 2001), which negatively affects classifiers. Second, the classification model becomes less effective when different materials have common or overlapping spectral bands, in which case multiple bands must be used in order to produce unambiguous classifications. In general, as the classes become more similar with respect to their chemical and spectral nature, more information will be required to further distinguish them.

An alternate and more robust approach is to use multivariate methods that consider each point in the spectrum as a separate variable. This approach has been combined with automated classification models developed in the fields of machine learning and pattern recognition. The process usually requires a transformation of the resulting high-dimensional data into a more concise representation before it can be used effectively with machine-learning algorithms.

There has been extensive work in the automated analysis of spectroscopic data using pattern recognition and machine learning techniques, which has made an impact in diverse academic fields including forensic science, archaeology, and biomedicine, as well as in commercial domains such as in the petroleum and automotive industries. To name a few examples, O'Connell et al. (2005) and Howley et al. (2006) employed machine-learning methods including decision trees and support vector machines (SVM) on Raman spectra to quantify the presence of compounds in solid mixtures of drugs and narcotics. Sigurdsson et al. (2004) trained an artificial neural network to detect cancer cells from data acquired using in vitro Raman spectroscopy of skin tissue samples. Balabin et al. (2010) surveyed a number of machine learning methods, including linear discriminate analysis (LDA), *K*-nearest neighbor, SVM, probabilistic neural networks (PNN), and multilayer perceptrons (MLP), as spectral classifiers for improving quality control in industrial gasoline production. These, and numerous similar studies, rely on various spectral pre-processing techniques that were shown to significantly improve classification accuracy, including noise reduction by smoothing, baseline removal, normalization, and dimensionality reduction.

In contrast, there are notably fewer systems developed for use in geology and mineralogy. Among them is an early method by Gazis and Roush (2001) that introduced a carbonate detector operating on a rule-based system with near-infrared reflectance spectroscopy data. A neural network classifier by Gallagher and Deacon (2002) was used to classify large numbers of mineralogical samples using X-ray spectra with applications towards mining research. Roush and Hogan (2007) and Hogan and Roush (2009) introduced a method to classify remote sensing data of terrestrial and planetary surfaces using unsupervised artificial neural networks, known as self-organizing maps (SOM). A study by Gilmore et al. (2008) used a SVM to detect sulfate minerals that are associated with the presence of water using visible/near-infrared spectroscopy. The latter was developed with the goal of providing autonomous onboard data analysis in future robotic missions to Mars.

Additionally, Pedersen (2000) addressed some technical aspects of undertaking geology from an autonomous, robotic vehicle (Apostolopoulos et al., 1999) designed to find meteorites in Antarctica. This work included a Bayesian classification model to handle evidence from multiple instruments and sensors. Computer vision methods by Gulick et al. (2001) were utilized to autonomously detect rocks and layers within rocks and outcrops for the purposes of classification, remote navigation, and hazard avoidance. Further studies by Gulick and Ishikawa (2008) and Ishikawa et al. (2010) introduced autonomous image-based color and texture analysis of rock images to distinguish between intrusive/extrusive and felsic/mafic compositions of igneous rocks. To our knowledge, there are no previous studies to determine autonomously the minerals in rocks and sediments using in situ Raman spectroscopy.

3. Experimental procedure

3.1. Gathering spectral readings using a mineral collection

We used a Concurrent Analytical Inc. Raman spectrometer with an excitation wavelength of 852 nm to measure spectra of minerals in our collection. Our instrument has a resolution of 1.8 cm^{-1}, a spectral range of 456.1–1638.8 cm^{-1}, and a spot size of $\sim$50 μm. Although this generates lower quality spectra compared to laboratory-grade instruments, it serves as a suitable analog for a portable field spectrometer that may be onboard a robotic vehicle.

Several spectra, each consisting of 765 data pairs of Raman shift and spectral intensity values, were acquired from distinct points on the surface of each mineral sample. The spectra were obtained manually under controlled conditions. Each sample was shielded from ambient light by placing it in a dark enclosure. During the acquisition process the probe was initially placed directly above the sample at a distance of approximately 1mm from the surface and slowly raised until the signal strength was at its strongest ($\sim$1–2 cm), roughly coinciding with the focal length of the laser beam. Although the spectra were obtained manually, work by Haskin et al. (1997) demonstrates the feasibility of automating such a process in the field. They report acquisition times of $\sim$30 s per spectrum for a rock fragment and refer to experiments that demonstrate their spectrometer is able to produce adequate spectra despite being significantly out of focus.

The acquired data was divided into a *training/validation set* and a separate *testing set* (see Table 1). The training/validation set contains thirteen different minerals, each with ten representative mineral spectra acquired from different points on the mineral surface. Unfortunately, due to the limited amount of mineral samples, the same could not be done for the testing set. Each mineral, however, with the exception of biotite, is represented by at least one spectral observation in the testing set, and each mineral group is guaranteed one representative mineral with a full set of at least ten spectral observations.

Table 1
Raman spectra obtained from our mineral collection.

Mineral	Group	Training set		Testing set	
		No, of samples	No, of observations	Mineral	Group
Albite	Plagioclase	1	10	1	10
Anorthite	Plagioclase	1	10	1	1
Labradorite	Plagioclase	1	10	1	1
Augite	Pyroxene	1	10	1	1
Diopside	Pyroxene	1	10	1	10
Hedenbergite	Pyroxene	1	10	1	1
Hornblende	Amphibole	1	10	2	2
Microcline	*K*-spar	1	10	1	10
Orthoclase	*K*-spar	1	10	1	10
Biotite	Mica	1	10	–	–
Muscovite	Mica	1	10	1	1
Olivine	–	1	10	1	10
Quartz	–	1	10	3	3
Total		13	130	15	60

3.1.1. Data pre-processing

We used the spectral data measured from our mineral collection to analyze the effects of various pre-processing methods on the classification accuracy. Smoothing was performed using a Savitzky and Golay (1964) filter, which applies polynomial regression to a series of data points. We chose the latter in favor of a simpler "moving average" filter because it is better at preserving the location of local maxima and minima. Background fluorescence, which can often obscure important Raman features that are used to distinguish minerals, seems to increase with the spectral response, as illustrated in the raw spectra of augite in Fig. 2. To diminish this effect, we considered derivatizing the spectra, by replacing the spectral intensity values with their first derivative. This technique enhances the sharper Raman features, but it can also increase the amount of unwanted noise requiring an additional smoothing procedure. However, and to our surprise, our classifier performed better without derivatization. This suggests that it is leveraging the unique fluorescence of some samples to obtain better classification rates.

We normalized the spectra by linearly re-scaling the intensity values to the range [0,1] in order to bring the peaks to comparable heights. Since the intensities vary based on the spectral response of the samples and the instrument they introduce a significant amount of systematic error. For example, the graph at the top of Fig. 2 shows 10 unprocessed spectra of the same augite specimen. In this case, the spectral baselines vary by approximately a factor of three. Note, however, that normalizing the spectral intensities allows for easier comparisons between spectra. Furthermore, by derivatizing the spectra and smoothing afterwards, the background fluorescence is significantly reduced and Raman bands are enhanced albeit at the cost of some added noise.

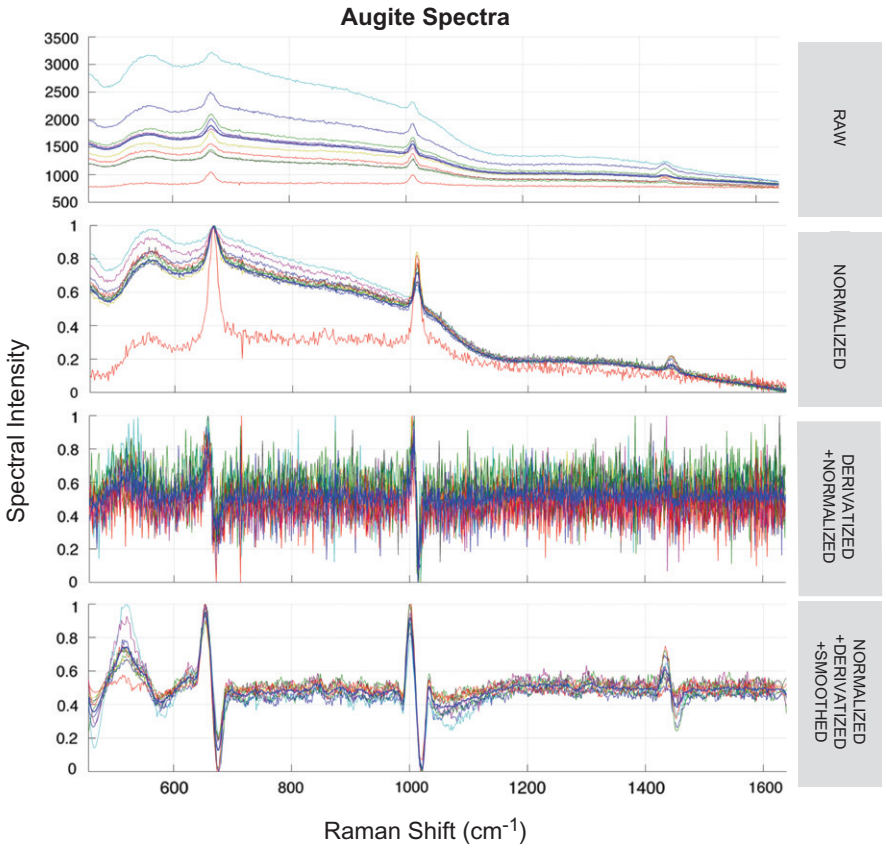

Fig. 2. Plots illustrating the effect of normalization, derivatization, and smoothing applied to the raw spectra of a single augite mineral sample. Note that the process of derivatization can amplify existing noise and should thus be followed by a smoothing procedure.

3.2. Using a third-party database of mineral Raman spectra

To further test the robustness of our algorithm, we used spectra from the RRUFF mineral database (Downs, 2006), a website that archives Raman spectra of mineral samples collected from a variety of sources. Mineral compositions have been confirmed by X-ray diffraction using unit cell parameters and powder diffraction profiles, or by chemical analysis. Using the RRUFF database, we selected a total of 167 spectra that include quartz, olivine, *K*-spars, plagioclases, micas, and pyroxenes. These are the same mineral types used in testing samples from our own collection. A total of 60 spectra, ten from each mineral group, were used in the training/validation set. The remaining 107 spectra were assigned to the testing set. Note that the testing set is significantly larger than the one we used from our own spectral database.

Multiple contributors using different instruments with varying resolutions and spectral ranges acquired the spectral measurements in the RRUFF database. The excitation wavelengths of these instruments ranged from 514 to 785 nm. Because different instruments were used, we used additional pre-processing steps to designate a common spectral range and interval. As a result, the processed spectra ranged from 272.8 to 1092.4 cm^{-1} with the smallest interval equal to 0.04816 cm^{-1}. The data points of all the spectra were equally spaced on the horizontal axis by linearly interpolating them over the common interval. To account for the variations in spectral response due to the differences in the minerals and the instruments themselves, the spectral intensity values were normalized to the range [0,1]. No further processing was applied since most spectra were already processed by smoothing and baseline correction prior to our use. Out of 167 spectra only 10 spectra were considered "raw." Regardless of the processing, compared to the spectra collected with our own spectrometer, the RRUFF spectra were qualitatively much cleaner with sharper Raman features and minimal noise and fluorescence. The minerals used from the RRUFF database are listed in Table 2.

Table 2
Spectra used from the RRUFF mineral database.

Mineral	Group	Training set No, of samples/ observations	Testing set No, of samples/ observations
Albite	Plagioclase	2	7
Andesine	Plagioclase	1	1
Annite	Plagioclase	–	1
Anorthite	Plagioclase	1	2
Bytownite	Plagioclase	1	3
Labradorite	Plagioclase	4	7
Oligoclase	Plagioclase	1	1
Augite	Pyroxene	1	4
Clinoenstatite	Pyroxene	1	1
Diopside	Pyroxene	1	14
Enstatite	Pyroxene	2	14
Ferrosilite	Pyroxene	1	2
Hedenbergite	Pyroxene	2	2
Jadeite	Pyroxene	1	1
Spodumene	Pyroxene	1	7
Anorthoclase	*K*-Spar	3	1
Mircocline	*K*-Spar	3	3
Orthoclase	*K*-Spar	4	6
Lepidolite	Mica	1	1
Muscovite	Mica	3	7
Phlogopite	Mica	2	2
Trilithionite	Mica	2	1
Zinnwaldite	Mica	2	–
Fayalite	Olivine	3	4
Forsterite	Olivine	7	10
Quartz	Quartz	10	5
Total		60	107

3.3. Dimensionality reduction

Spectra are good examples of high-dimensional data, for example, using our spectrometer we collected 190 spectra, each with 765 data points. Unfortunately, the number of observations needed to successfully train a machine-learning classifier grows exponentially with the number of data points, a phenomenon that is appropriately called the *curse of dimensionality*. However, Principal Component Analysis (PCA), a statistical method used to reduce the dimensions of large multivariate data, is frequently used to address this. Essentially, PCA exploits the redundancy of a high-dimensional coordinate system and maps it into a lower-dimensional one with more "meaningful" axes by applying linear transformations. If such transformations are possible, then we obtain a coordinate system that is able to capture most of the variance of the data with significantly fewer dimensions.

Fig. 3 shows a plot of various mineral observations mapped into a principal component (PC) feature space where the axes correspond to the top three principal components that contain 88.1 percent of the total variance in the data. This illustrates the reduction of a 765-dimensional dataset, which is impossible to visualize otherwise, to only three dimensions. In addition, the fact that the data points of similar minerals tend to cluster together in this coordinate system is a good indicator that a machine-learning approach may be successful at learning the differences.

From Fig. 3 it is evident that several minerals, including quartz, albite, muscovite, and olivine, are resolved easily by simple visual inspection. Microcline and orthoclase share a common and distinct "cloud" of points. Therefore, we can expect an accurate classification for these two potassium feldspars as a group, but poor distinction between the individual minerals themselves. We expect a less accurate classification of pyroxenes, for example, which are not tightly clustered.

4. Classifying minerals with machine learning

We used supervised machine-learning algorithms to train a mineral classifier using a set of examples labeled manually by a geologist. The term "supervised" refers to the use of training examples to generate a classification model. Once a classifier is trained it can be used autonomously. The parameters of the learning algorithm must be fine-tuned to produce the best results. Ideally, such parameter optimization is achieved using a separate validation set to predict the performance of the classification model on unseen data. Typically, 50 percent of all available data is used for training, 25 percent for validation, and the remaining 25 percent for testing. In practice, however, the amount of available data is often limited, as is true in our case. A practical alternative is to use a *k-fold cross-validation* procedure, where the training set is randomly partitioned into k subsets. Each subset is left out once and used as the validation set while the remaining $k-1$ subsets are used to train the classifier. This is repeated k times until each subset has been left out once. The advantage of this approach is that a single set can be used for both training and validation and a separate set for testing. With the training/validation results, we built a classifier by selecting parameter values that produced the highest accuracy using a ten-fold ($k=10$) cross validation procedure. Finally, we evaluated the performance of the trained classifier by how well the learned model performed on unseen data consisting of a separate but similar test set.

In this study we considered two different classification methods (1) decision trees, and (2) artificial neural networks. These classifiers were built using the multi-layer perceptron (MLP) and C4.5 algorithms, respectively, from the machine-learning library WEKA (Hall et al., 2009; Witten et al., 2011). Decision trees are useful because

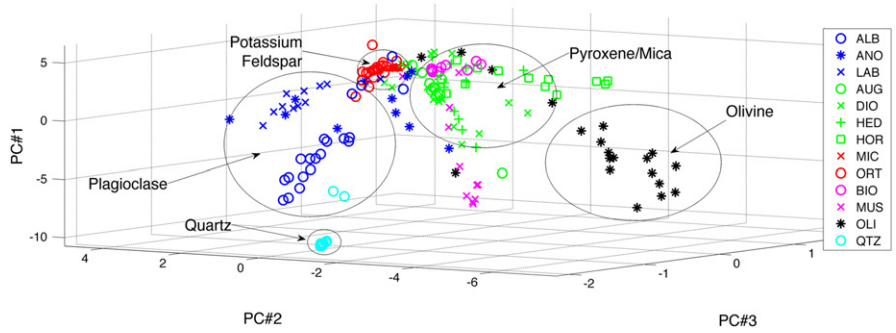

Fig. 3. Distribution of minerals from our training set using a feature space obtained by transforming high-dimensional spectral data into only three dimensions. Points of the same color belong to the same mineral group. There are 13 individual minerals types from 6 separate groups. Note that certain minerals are distributed in tight clusters, a trend that could be easily learned by a machine classifier. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

their behavior is transparent—successful training yields a set of generalized rules that specify how to use the information in the data to systematically classify similar, but unseen, data. These rules would be useful to geologists who wish to re-trace the machine's "reasoning" process in a human-readable format. Artificial neural networks, on the other hand, do not provide as much intuition as to how the classifier arrives at its decision. However, they were found to be much more robust and provided superior classification of the mineral data considered in this study. As a result, we chose to implement our mineral classifier using the MLP neural network algorithm.

4.1. Building an artificial neural network

A neural network is composed of simple models of biological neurons that are collectively capable of learning by example. Fig. 5 illustrates the resulting architecture used in our classifier consisting of a three-layer feed-forward neural network with six hidden nodes. The network is used in our study to map spectrally derived data from geologic specimens into mineral classes.

By providing training examples, a neural network is able to form meaningful generalizations on unseen data of a similar nature. We use training data, which consists of examples of minerals and their corresponding ground truth or target labels, to teach the neural network the appropriate responses that lead to successful classifications. Since each training sample has an associated target label, we are able to quantify the error between the classification and the actual value by introducing an *error metric*, which provides a measure of how far the estimated classification is from ideal.

Training is performed with the *backpropagation* (Rumelhart et al., 1986) algorithm described in detail by Marsland (2009) and Duda et al. (2000). During backpropagation, feature vectors representing each training instance are fed multiple times through the network. Each time, the resulting cascade of neural responses produces a classification vector that encodes the network's current "best guess," given the prior information learned by adapting the weights of the neural connections. After each instance is "seen," the error between the network's classification and the correct value is computed and used as feedback to nudge the appropriate weights in the right direction. The latter is achieved using a *gradient descent* approach to finding the "low points" in the topology of the error function. The process is repeated until the algorithm converges, i.e., until a local minimum of the error function has been reached. Training is successful when the provided training data can be classified within a reasonable distance to the accepted ground truth and new data can be classified using the same neural network architecture.

One well-known disadvantage of using the *backpropagation* algorithm for learning the appropriate neural weights is the possibility of multiple local minima in the error function. Since the algorithm is essentially a hill-climbing technique there is no guarantee that a global minimum has been reached. Therefore, if the algorithm gets stuck in a high-error local minimum it may fall short of the optimal amount of training required to learn classifications. To address this, we performed cross-validation ten times using different randomly assigned training and testing sets. The plots in Fig. 4 represent the average accuracy from ten separate trials with the confidence intervals representing one standard deviation of their mean. The relatively small error bars suggest that the "local minima effects" are nominal.

4.2. Optimizing the neural network parameters

To evaluate the accuracy of our classifier, we used raw spectral data from our own mineral collection to conduct a systematic study of various pre-processing parameters. The parameters we allowed to vary were the number of PCs, the amount of smoothing, the normalization, and the derivatization of the spectra. The effect of various combinations of parameters on the classification accuracy was determined using the training/validation set by using 10-fold cross-validation to predict the performance of the classifier on an unseen test set. Since the training and validation samples are drawn from the same dataset, we expect some over-fitting. Therefore, the reported accuracy will likely be higher than when using a separate testing set. We found this to be the case since the predicted training/validation accuracy was ~90 percent, but the actual testing accuracy was 83 percent. The plots in Fig. 4 show the trends in classification accuracy as a function of these parameters.

4.3. Results and discussion

4.3.1. Artificial neural networks vs. decision trees

We compared the classification performance of artificial neural networks and decision trees. Figs. 4 and 5 show the cross-validation performances, respectively. A total average classification accuracy using neural networks (MLP) and decision trees (C4.5) using five or more PCs was 88.5 and 82.0 percent, respectively. Therefore, it is clear that the cross-validation performance of neural networks were notably superior.

4.3.2. Training and validation

We used mineral spectra gathered from our own collection to run experiments using the 10-fold cross-validation procedure on the training/validation set. The experimental results show that normalizing the spectral data produces a ~12 percent increase in accuracy (see Fig. 4). This was expected because the spectral intensities change, even between different spectra of the same mineral, which makes learning harder. Furthermore, normalization facilitates comparisons between spectra of varying signal strengths so that spectra of the same mineral will look more

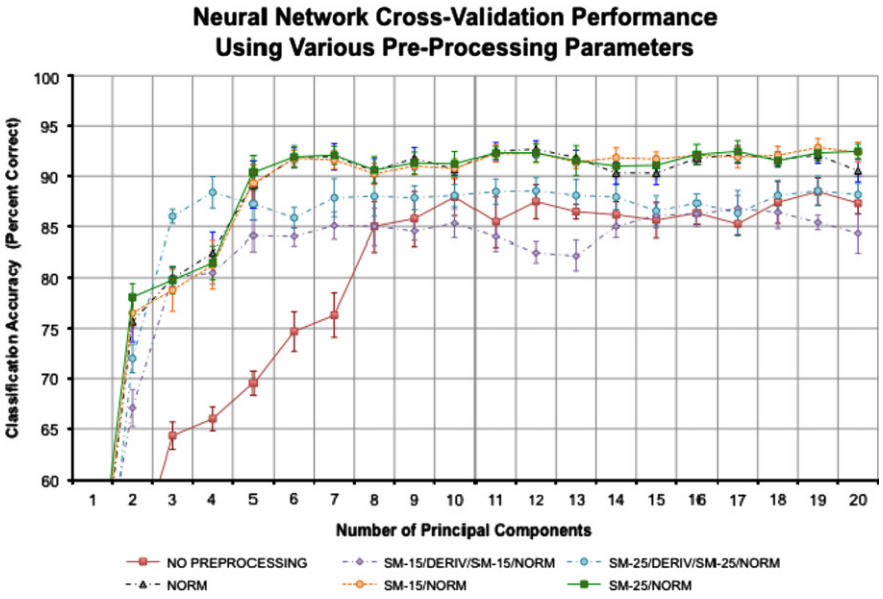

Fig. 4. Plots demonstrating the effects of various data pre-processing parameters on the classification accuracy of the neural network during cross-validation. Accuracy improves dramatically when increasing the number of principal components from 1 to 5, after which there is no significant improvement. This suggests that the spectral data can be sufficiently represented in five principal component dimensions. Accuracy is highest when the spectra are first smoothed and normalized, without derivatization.

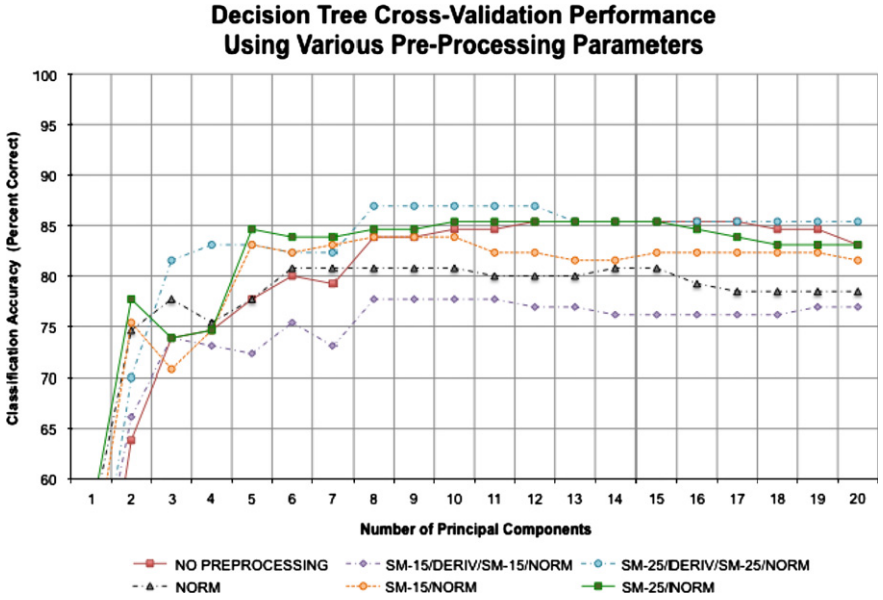

Fig. 5. Comparison of data pre-processing parameters on the classification accuracy of the decision tree during cross-validation. Similar to the neural network, accuracy improves when increasing the number of principal components up to approximately 7. Unlike the neural networks, smoothing the spectral data before PCA reduction appears to improve the classification accuracy.

similar. On the other hand, smoothing showed only marginal improvements in certain instances and taking the first derivative of the spectra did not improve the results. Given our spectral data, both smoothing and derivatization were the least effective of the pre-processing methods considered. This suggests that there is useful information within the spectral fluorescence that is being removed by derivatization.

The number of PCs used for classification has a pronounced effect on the classification accuracy when less than five PCs are used (see Fig. 4). Increasing the number of PCs past this number provides little improvement suggesting that the spectral data can be sufficiently represented in 5 PC dimensions. This is not surprising considering 88.1 percent of the spectral variance is contained in the top 3 PCs. Therefore, each additional PC is subject to diminishing returns that make it less effective towards improving the classifier.

Based on the above findings, we chose to normalize the spectra before performing a PCA procedure to reduce the data to the top five PCs. Although smoothing provided little to no improvements, we still chose to apply a 25-point Savitzky–Golay smoothing window in order to make the system more robust to noisy spectra. The cross-validation results are represented by the green curve in Fig. 4. The resulting architecture used in our classifier is illustrated in Fig. 6. It consists of a three-layer feed-forward neural network with six hidden nodes (this number was chosen simply because it is equivalent to the number of mineral

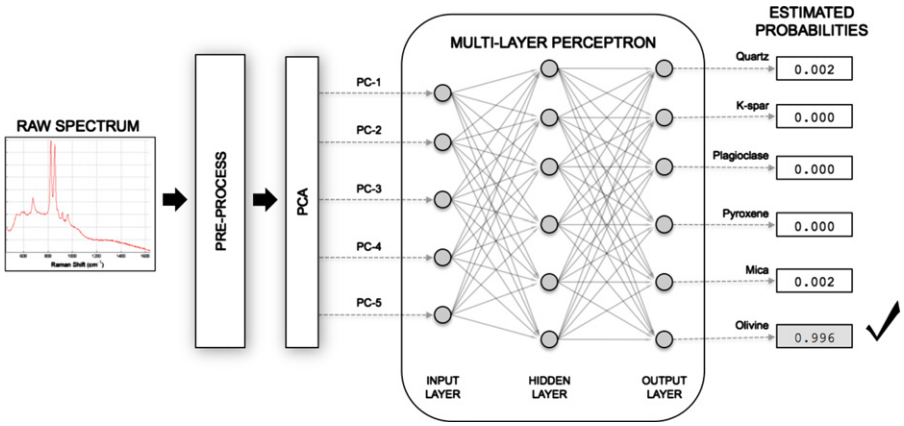

Fig. 6. Diagram of the classification process of spectral data (in this case, a sample of olivine) using a three-layer feed-forward artificial neural network trained with data on six different mineral groups.

"classes"). For the hidden nodes, a sigmoid activation function was used which determines the (nonlinear) threshold at which the neurons will "fire."

4.3.3. Testing

We evaluated our optimized classifier, described above, using a separate test set from our own spectral data consisting of 15 individual minerals with a total of 60 spectral observations. We were able to classify pure minerals by their mineral groups with an average accuracy of 83 percent. However, the classification of minerals on an individual basis was less successful with an average accuracy of only 73 percent. This is explained by three factors. First, in terms of their molecular structure, minerals in the same group vary only by their associated end members, which makes the observed Raman features often very similar. For example, diopside and hedenbergite, both in the pyroxene group, were often indistinguishable by the mineral classifier. Indeed, visual inspection of the spectra confirmed their similarity (Fig. 7). Second, fluorescence can mask subtle differences in spectra and minerals within the same group may be mistaken for one another. Third, the resolution of our spectrometer may hamper the ability to resolve some details in the spectra.

Experimental results from using our own spectral database show that the individual minerals within the pyroxene group (augite, diopside, hedenbergite, and hornblende) are especially difficult to classify. In particular, our classifier correctly identified diopside samples only 20 percent of the time because they were often mistaken for hedenbergite and vice versa. On the other hand, there was more resolution than anticipated between the potassium feldspars, microcline and orthoclase. Even though these minerals have remarkably similar spectra (see Fig. 1), our method was able to classify them with 100 and 80 percent accuracy, respectively. This was surprising because the spectra of both minerals look even more similar than those in the pyroxene group due to strong background fluorescence. In fact, we determined that it is precisely because of the fluorescence of these minerals that the classifier was able to distinguish them. Therefore, the ability of the neural network to incorporate fluorescence properties into its overall classification scheme shows its robustness as a mineral classifier.

Because of our concern with the limited sample size of our in-house collection, we also used the RRUFF (Downs, 2006) mineral database to further test the robustness of our mineral classifier. Classification of pure mineral groups, using the RRUFF database, was achieved with an average of 80.4 percent accuracy. The

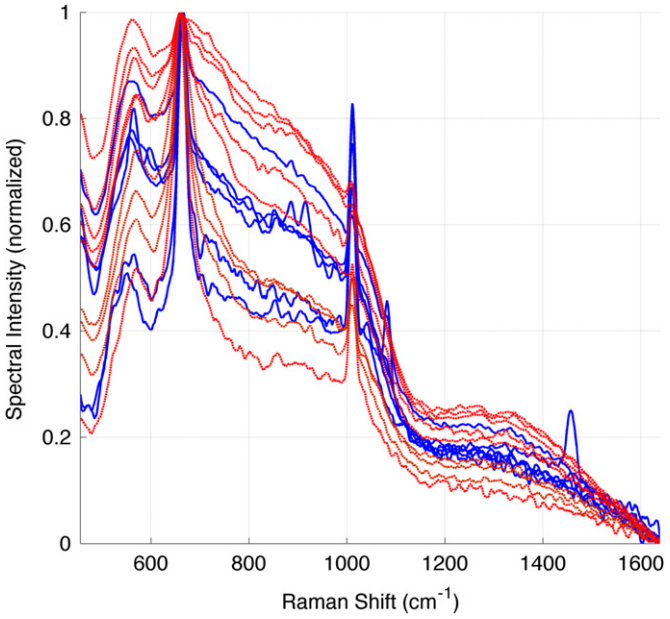

Fig. 7. Comparing Raman spectra of diopside (red) and hedenbergite (blue). Both minerals are in the pyroxene group and were indistinguishable by the mineral classifier. The spectra were normalized and smoothed using a Savitzky–Golay filter with a 4th order polynomial and a 25 point smoothing window. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

overall results are comparable to those using our own mineral collection where we obtained an average 83 percent accuracy. The minerals were similarly distributed in the PC feature space, as demonstrated with the results from our mineral collection. As expected, quartz and olivine were easy to distinguish since their points were tightly clustered in the feature space. Both minerals were classified correctly 100 percent of the time. All of the pyroxene spectra were correctly classified, which is significantly better than the 85 percent accuracy from using our own data.

One striking difference, however, is that only 18 percent of the plagioclase samples in the RRUFF testing set were classified correctly, 77 percent of which were mistaken for potassium feldspars. This is significantly lower compared to the results from using spectra from our own collection. Interestingly, potassium feldspars in the RRUFF data were classified with 90 percent accuracy, which is superior to the 80 percent accuracy from our database. We explain this as follows: (1) potassium feldspars and plagioclases have similar chemical properties that are reflected in

Table 3
Confusion matrix summarizing the classification performance of our neural network that was trained with spectral data from our mineral collection. On average, the classifications were correct 83 percent of the time.

	QTZ	KSP	PLA	PYX	MCA	OLI	(%) Correct
QTZ	3	0	0	0	0	0	100
KSP	0	16	4	0	0	0	80
PLA	0	2	10	0	0	0	83
PYX	0	1	0	12	1	0	85
MCA	0	0	0	0	1	0	100
OLI	0	1	0	1	0	8	80

Table 4
Confusion matrix summarizing the classification performance of our neural network that was trained with spectral data from the RRUFF mineral database. On average, the classifications were correct 80.4 percent of the time.

	QTZ	KSP	PLA	PYX	MCA	OLI	(%) Correct
QTZ	5	0	0	0	0	0	100
KSP	0	9	1	0	0	0	90
PLA	0	17	4	0	1	0	18
PYX	0	0	0	45	0	0	100
MCA	0	0	0	2	9	0	81
OLI	0	0	0	0	0	14	100

their spectral similarities, (2) experiments in which we used our own spectral data showed higher accuracy in discriminating *K*-spars and plagioclases due to our instrument's tendency to fluoresce in predictable patterns in the presence of *K*-spars microcline and orthoclase, and (3) although the RRUFF data was of higher resolution than our own data by a factor of $\sim$25, the spectral range was smaller by a factor of 1.44. While the higher resolution of the RRUFF data may have had a positive effect on the overall classification accuracy for non-plagioclase samples, the classification accuracy of the plagioclases was hampered by their spectral similarity to *K*-spars, the absence of predictable fluorescence patterns in the *K*-spars as observed in our particular spectrometer, the narrower spectral range of the data, or a combination thereof. On the other hand, the fact that better accuracy is attained with samples exhibiting unique fluorescence patterns demonstrates that our classifier is capable of leveraging a particular response from the instrument to improve classification. This provides further demonstration of the robustness of our method.

Tables 3 and 4 provide *confusion matrices* that summarize the classification accuracies using our own mineral archive and the RRUFF database, respectively. Confusion matrices help one visualize the accuracy of classification algorithms by providing information on how often an item was misclassified as something else. They are commonly used in machine-learning and artificial intelligence applications. In Tables 3 and 4, the rows indicate the total distribution of mineral classifications in a particular class of minerals given by the row header. Therefore, the sum of the cells in each row indicates the total number of items belonging to a given class. The sum of the shaded diagonal cells indicates the total number of correct classifications.

5. Conclusion

We demonstrated that machine-learning methods, specifically the use of artificial neural networks, could be used with Raman spectroscopy to classify key minerals that are useful indicators of igneous rock composition. The performance of our mineral classifier demonstrates the robustness of using artificial neural networks to learn complex relationships between spectral data and mineral identities. We expect that our classifier will produce accurate results on any spectral database of minerals provided the following:

- *Each category of minerals should have reasonably distinct spectral signatures*. Note, however, that we were able to demonstrate that an artificial neural network is able to learn the subtle differences, beyond simple Raman band positions and general shape, that are not as easily perceived by humans.
- *There should be sufficient training examples*. Based on our findings, we recommend that each class of minerals contain roughly 10–30 examples depending on the uniqueness and quality of the spectra.

Finally, we were able to test further the accuracy of our mineral classifier using a larger collection of minerals from the RRUFF database. This was achieved using roughly half the amount of training samples as before, probably because of the higher quality Raman spectra. Our findings serve to illustrate that autonomous mineral classification using neural networks is both accurate and robust, and can be done with as little as 160 independent mineral observations.

Acknowledgements

We thank Shawn Hart for acquiring the Raman spectra of our samples. This research was supported by a prior grant from NASA's Advanced Cross Enterprise Technology Development Program.

References

Apostolopoulos, D., Wagner, M., Whittaker, W., 1999. Technology and field demonstration results in the robotic search for antarctic meteorites. In: Proceedings of the International Conference on Field and Service Robotics, pp. 185–190.

Balabin, R.M., Safieva, R.Z., Lomakina, E.I., 2010. Gasoline classification using near infrared (NIR) spectroscopy data: comparison of multivariate techniques. Analytica Chimica Acta 671 (1-2), 27–35.

Blake, D., 2012. The development of the CheMin XRD/XRF: reflections on building a spacecraft instrument. In: Aerospace Conference, 2012 IEEE, vol., no., pp.1–8. http://dx.doi.org/10.1109/AERO.2012.6187059.

Blake, D., Vaniman, D., Achilles, C., Anderson, R., Bish, D., Bristow, T., Chen, C., Chipera, S., Crisp, J., Des Marais, D., Downs, R.T., Farmer, J., Feldman, S., Fonda, M., Gailhanou, M., Ma, H., Ming, D.W., Morris, R.V., Sarrazin, P., Stolper, E., Treiman, A., Yen, A., 2012. Characterization and Calibration of the CheMin Mineralogical Instrument on Mars Science Laboratory. Space Science Review. http://dx.doi.org/10.1007/S11214-012-9905-1.

Downs, R.T., 2006. The RRUFF project: an integrated study of the chemistry, crystallography, Raman and infrared spectroscopy of minerals. In: 19th General Meeting of the International Mineralogical Association in Kobe, Japan. O03-13.

Duda, R.O., Hart, P.E., Stork, D.G., 2000. Pattern Classification, second ed. John Wiley and Sons, New York, NY10158, USA 654 pp.

Gallagher, M., Deacon, P., 2002. Neural networks and the classification of mineralogical samples using x-ray spectra. In: Proceedings of the 9th International Conference on Neural Information Processing, vol. 5, pp. 2683–2687. http://dx.doi.org/10.1109/ICONIP.2002.1201983.

Gazis, P.R., Roush, T., 2001. Autonomous identification of carbonates using near-IR reflectance spectra during the February 1999 Marsokhod field tests. Journal of Geophysical Research 106 (E4), 7765–7773.

Gilmore, M.S., Bornstein, B., Merrill, M.D., Castaño, R., Greenwood, J.P., 2008. Generation and performance of automated jarosite mineral detectors for visible/near-infrared spectrometers at Mars. Icarus 195 (1), 169–183.

Gulick, V.C., Morris, R.L., Ruzon, M.A., Roush, T.L., 2001. Autonomous image analyses during the 1999 Marsokhod rover field test. Journal of Geophysical Research 106 (E4), 7745–7763, http://dx.doi.org/10.1029/1999JE001182.

Gulick, V.C., Ishikawa, S.T., 2008. Improving high-resolution image analysis: results for remote science exploration. In: 39th Lunar and Planetary Science Conference, pp. 2527–2528.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Ian, H., 2009. The WEKA data mining software: an update. SIGKDD Explorations 11 (1), 10–18.

Haskin, L.A., Wang, A., Rockow, K.M., Jolliff, B.L., Korotev, R.L., Viskupic, K.M., 1997. Raman spectroscopy for mineral identification and quantification for in situ planetary surface analysis: a point count method. Journal of Geophysical Research 102 (E8), http://dx.doi.org/10.1029/97JE01694 19,293–19,306.

Hogan, R., Roush, T., (2009). Mineral emittance spectra: clustering and classification using self-organizing maps. In: Aerospace Conference, 2009 IEEE. Paper no. 1456, pp.1–7, 7–14 March 2009. http://dx.doi.org/10.1109/AERO.2009.4839480.

Howley, T., Madden, M.G., O'Connell, M.-L., Ryder, A.G., 2006. The effect of principal component analysis on machine learning accuracy with high dimensional spectral data. Knowledge-Based Systems 19 (5), 363–370.

Ishikawa, S.T., Hart, S.D., Gulick, V.C., 2010. Mineral detector for igneous rocks. In: American Geophysical Union, Fall Meeting 2010. Abstract #IN51A-1137.

Lewis, I.R., Edwards, H.G.M., 2001. Handbook of Raman Spectroscopy: From the Research Laboratory to the Process Line. Marcel Dekker, Inc., New York, NY 10016, USA 1080 pp.

Marsland, S., 2009. Machine Learning: An Algorithmic Perspective. Machine Learning & Pattern Recognition Series. Chapman & Hall/CRC, Boca Raton, FL 33487, USA 406pp.

O'Connell, M.-L., Howley, T., Ryder, A.G., Leger, M.N., Madden, M.G., 2005. Classification of a target analyte in solid mixtures using principal component analysis, support vector machines and Raman spectroscopy. In: Proceedings of SPIE, the International Society for Optical Engineering. vol. 5826, pp. 340–350.

Pedersen, L., 2000. Robotic Rock Classification and Autonomous Exploration, Ph.D. Dissertation, Technical Report CMU-RI-TR-01-14. Robotics Institute, Carnegie Mellon University. November, 2000.

Roush, T.L., Hogan, R., (2007). Automated Classification of Visible and Near-Infrared Spectra Using Self-Organizing Maps. In: Aerospace Conference, 2007 IEEE. Paper no. 1456, pp.1–10, 3–10 March 2007. http://dx.doi.org/10.1109/AERO.2007.352701.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by back-propagating errors. Nature 323 (99), 533–536.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36 (8), 1627–1639, http://dx.doi.org/10.1021/AC60214A047.

Sigurdsson, S., Philipsen, P.A., Hansen, L.K., Larsen, J., Gniadecka, M., Wulf, H.C., 2004. Detection of skin cancer by classification of Raman spectra. IEEE Transactions on Biomedical Engineering 51 (10), 1784–1793.

Wiens, R.C., Maurice, S., Barraclough, B., Saccoccio, M., Barkley, W.C., Bell, J.F., Bender, S., Bernardin, J., Blaney, D., Blank, J., 70 coauthors, 2012. The ChemCam Instrument Suite on the Mars Science Laboratory (MSL) Rover: Body Unit and Combined System Tests. Space Science Reviews, http://dx.doi.org/10.1007/S11214-012-9902-4.

Witten, I.H., Frank, E., Hall, M.A., 2011. Data Mining: Practical Machine Learning Tools and Techniques, third ed. Morgan Kaufmann Publishers, Burlington, MA 01803, USA 664 pp.