HYPERSPECTRAL SIGNATURE DETECTION OF LOW ABUNDANCE INTIMATE MIXTURES IN MICROSCENES USING CONVOLUTIONAL NEURAL NETWORKS

by

Kevin P. Christiansen

A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Science
Geoinformatics and Geospatial Intelligence

Committee:

_____ Dr. Arie Croitoru, Thesis Chair

_____ Dr. Ronald G. Resmini, Committee Member

_____ Dr. Konrad Wessels, Committee Member

_____ Dr. Dieter Pfoser, Department Chairperson

_____ Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science

_____ Dr. Ali Andalibi, Interim Dean, College of Science

Date: _____ Summer Semester 2019

George Mason University

Fairfax, VA

Hyperspectral Signature Detection of Low Abundance Intimate Mixtures in Microscenes using Convolutional Neural Networks

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at George Mason University

by

Kevin P. Christiansen
Bachelor of Science
Rochester Institute of Technology, 2012

Director: Arie Croitoru, Professor
Department of Geography and Geoinformation Science

Summer Semester 2019
George Mason University
Fairfax, VA

Copyright 2019 Kevin P. Christiansen
All Rights Reserved

ACKNOWLEDGEMENTS

I would like to thank all my teachers who inspired me to explore science at an early age and enabled my career in science. I would also like to acknowledge Dr. Ronald G. Resmini, Dr. Arie Croitoru, and Dr. Konrad Wessels for their expert guidance and valuable input regarding this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

Adaptive Resonance Theory ..... ART
Central Processing Unit ..... CPU
Convolutional Neural Network ..... CNN
Deep Learning ..... DL
Digital Number ..... DN
Double Data Rate 4 ..... DDR4
Dynamic Random-Access Memory ..... DRAM
Environment for Visualizing Images (software) ..... ENVI
Gigabyte ..... GB
Graphical User Interface ..... GUI
Graphics Processing Unit ..... GPU
Ground Sampling Distance ..... GSD
Hyperspectral Imaging ..... HSI
Linear Mixing Model ..... LMM
Linear Spectral Unmixing ..... LSU
Matched Filter ..... MF
Megahertz ..... MHz
Millimeters ..... mm
Minimum Detectable Quantity ..... MDQ
Multi-layer Perceptron ..... MLP
Nanometer ..... nm
Near Infrared ..... NIR
Neodymium Oxide ..... Nd_2O_3
Region of Interest ..... ROI
Shortwave Infrared ..... SWIR
Signal-to-Clutter Ratio ..... SCR
Support Vector Machine ..... SVM
Terabyte ..... TB
Tera-Floating Point Operations Per Second ..... TFLOPS

ABSTRACT

HYPERSPECTRAL SIGNATURE DETECTION OF LOW ABUNDANCE INTIMATE MIXTURES IN MICROSCENES USING CONVOLUTIONAL NEURAL NETWORKS

Kevin P. Christiansen, M.S.

George Mason University, 2019

Thesis Director: Dr. Arie Croitoru

The high-confidence detection and identification of very low abundance, subpixel quantities of solid materials in nonlinear (intimate) mixtures are still significant challenges for HSI data analysis. Machine learning with convolutional neural networks (CNN) has proven to be an accurate means of identifying and typing features for various forms of data including estimating nonlinear functions and detecting features in one-dimensional data series. The application of CNNs to low abundance intimate mixtures, could improve minimum detectable quantities (MDQs) compared to current algorithms by processing spectra such that subtle features are enhanced and more discriminable. To test this, microscenes of three different intimate mixtures at varying abundance ratios (weight percent) were generated and measured using a benchtop shortwave infrared (SWIR) hyperspectral imager. A microscene is a hyperspectral image of a small-scale human-generated landscape measured in a laboratory. The 3 mixtures measured consisted

of silicate sand + Nd_2O_3, silicate sand + powdered sugar, and soil + aspartame. Several hundred thousand labeled spectra are easily and rapidly generated in one HSI cube of a microscene for neural network training and testing. The mixture detection abilities of several processing methods were compared including deep learning (DL) CNNs in TensorFlow, shallow (non-convolutional) neural networks in both TensorFlow and MATLAB, and ENVI's matched filter (MF), support vector machine (SVM), and linear spectral unmixing (LSU) functions. The CNN models for each mixture have >98 % average validation accuracies for detecting mixtures of varying abundances, including the lowest abundances measured. Both the DL/CNN and shallow neural networks tested showed increased detection capabilities compared to some more traditional ENVI methods. The networks were able to identify the lowest weight percent of sand + Nd_2O_3, but the MF and LSU methods did not yield results that would be considered reliable detections. The results establish confidence in using CNNs as a means of detecting low abundance intimate mixtures in real-world scenarios.

INTRODUCTION

The high confidence detection and identification of very low abundance, subpixel quantities of solid materials in nonlinear (intimate) mixtures are still significant challenges for hyperspectral imagery (HSI) data analysis. Motivated by a desire to reduce minimum detectable quantities and enabled by microscene data generation methods, this research tests the ability of deep learning (DL) convolutional neural networks (CNNs) to analyze spectral signatures of intimate mixtures. The results are compared to analysis methods in ENVI as well as neural networks developed using MATLAB.

The process of spectral unmixing is important to understanding the constituent materials of a remotely sensed HSI data. The ability to detect materials mixed with low abundances aids an analyst's understanding of what is going on in a scene. For example, questions such as how much of a chemical is present in a soil or how much oil film is on water could potentially be answered using spectral unmixing techniques. Increasing the ability to detect smaller quantities and doing so accurately, is the goal of this research.

Hyperspectral Mixtures

A detector (a pixel) in a remote sensing imaging system collects the electromagnetic energy scattered from multiple materials (endmembers) within its field of view. The resulting data contains a combination of spectral components from each

endmember. An analyst working with HSI is commonly looking to detect a specific material within a pixel or would like to know the component makeup of a pixel. Both analytic processes may involve a process known as spectral unmixing.

The types of materials that might be mixed within a pixel may vary widely with the ground sampling distance (GSD) of the system. For example, imagery from hyperspectral remote sensing systems such as NASA/JPL's AVIRIS[1], with a nominal GSD of 20 meters, may have mixtures including of a variety of vegetation, soils, and manmade materials [34]. A laboratory based HSI sensor with a controllable GSD (dependent on test setup), such as the data collected for this paper, can potentially measure any combination of materials and collect the data in a controlled environment.

Understanding the component makeup of the pixels being analyzed is important due to the phenomenology of the interaction of light and materials at certain wavelengths. This paper uses spectra captured from 936 nm to 2508 nm, spanning both the near-infrared (NIR) and shortwave infrared (SWIR) wavelength regimes. NIR/SWIR wavelengths are typically reflective as opposed to emissive, meaning photons detected are dependent on reflected sunlight, surface reflectivity, illumination geometry, and other complex phenomena. This paper does not consider possible thermal/emissive SWIR sources and focuses on the mixing effects and modeling of reflective phenomena.

In HSI analysis, a mixture model can describe the way endmembers are combined within the pixels of the image data and estimates their relative abundances. Spectral unmixing provides a comprehensive and quantitative mapping of the elementary

[1] https://aviris.jpl.nasa.gov/aviris/index.html - accessed July 2019

materials that are present in the data cube [13]. Baseline processing methods rely on linear spectral mixing models that can adequately estimate the reflectance and abundance of endmembers within a single pixel. However, the complexities of material-photon interactions such as multiple scattering within or around the materials often drive a requirement for more accurate, non-linear, models for unmixing. Understanding and modeling mixing phenomenologies due to photon interactions will lead to more accurate extractions of material abundances within spectral signatures.

The amount of detail to use when defining a mixing model can be subjective. Commonly, mixing models are defined as either linear or nonlinear. Linear mixing suffices when the arrangement of different materials in a scene or in a single pixel is macroscopic ("checkerboard"), and the incident light interacts with just one material prior to being received by the sensor [5]. The linear mixing model (LMM) is defined in Equation 1 where R_i is the reflectance of a pixel in wavelength band i, f_j is the fractional abundance of endmember j in the pixel, $M_{j,i}$ is the reflectance of endmember j in band i, r_i is the unmodeled reflectance for the pixel in band i, and n is the number of endmembers. The LMM assumes that all mixing occurs within an individual pixel and is due to the resolution of the instrument being greater than the size of the components being measured. Abundances in a linear mixture represent the relative area of the corresponding endmember in an imaged region [14].

$$R_i = \sum_{j=1}^{n} f_j M_{j,i} + r_i$$

Equation 1: Linear Mixing Model

Nonlinear mixing is due to the light being scattered by multiple materials in the scene and physical interactions between the light and the materials. These interactions can be due to multiple layers of macroscopic scattering (classical) or due to microscopic (intimate) interactions. Mixing at the classical level occurs when light is scattered from one or more objects, is reflected off additional objects, and eventually is measured by the sensor. Microscopic mixing occurs when two materials are homogeneously mixed [14]. Figure 1 shows a diagram from [6] showing a schematic of linear and nonlinear mixing interactions. Further information about unmixing taxonomies and more graphics showing linear and nonlinear interactions of light can be found in [21].

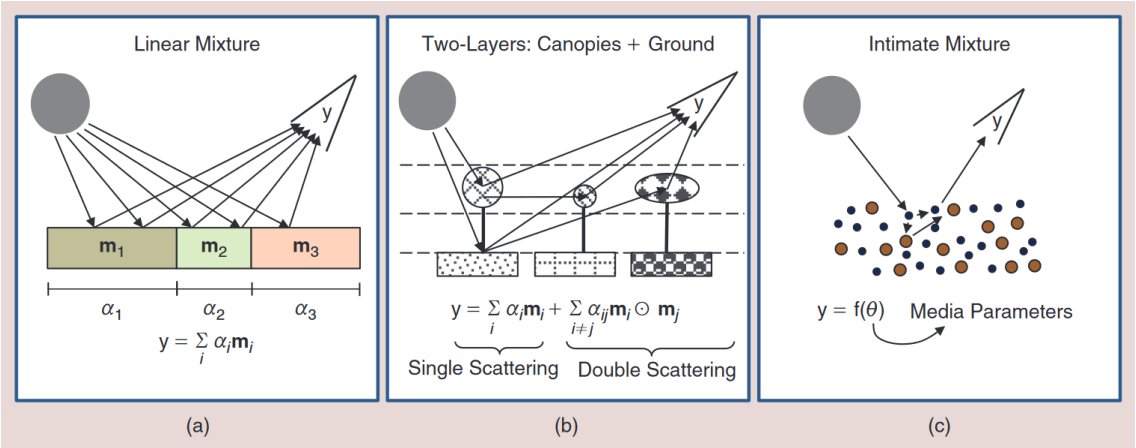


Figure 1. Schematic view of three types of spectral mixing.
(a) Linear mixing in a checkerboard type surface. (b) Nonlinear (linear plus bilinear) mixing in a two-layer media. (c) Nonlinear mixing in an intimate (particulate) media. Figure is from Bioucas-Dias et al. (2013) [6].

This research focuses on the application artificial neural networks, rather than defined analytical models, to analyze intimate mixtures. Artificial neural networks have demonstrated the capability of estimating complex nonlinear functions and several types of neural network architectures have been applied to the problem of nonlinear spectral unmixing [39]. This section will highlight some of the neural network applications to spectral unmixing.

Several studies have highlighted the benefits of neural networks applied to nonlinear mixtures and their improvement over linear unmixing techniques. Plaza et al. (2009) [33] and Kumar et al. (2012) [26] showed successful nonlinear mixture estimation using a multi-layer perceptron (MLP) networks. MLP networks are a relatively basic form of network that uses nonlinear functions to transform inputs in order to categorize the data. Improving on MLP networks, Carpenter et al. (1999) [9] developed a neural network-based mixture estimation method that outperforms maximum likelihood

classification and linear mixture models. The network is based on adaptive resonance theory (ART) which aims to increase stability and speed compared to MLPs through self-organizing the feature mapping from input vectors. Again, building on the base of MLP networks, Palsson et al. (2018) [32] showed that an autoencoder neural network extracted endmembers with improvements over traditional unmixing methods. Autoencoder networks build on MLP networks by learning a representation of the input data, which reduces data dimensionality, increasing speed. Lastly, Ahmed et al. (2017) [1] proposed using artificial neural networks to determine if a linear or nonlinear mixture model should be used for a specified input. Based on the characteristics of neighboring pixels, the network determines the best combination of mixtures models to use in order to produce the most accurate results. These studies build a foundation for further research in the area of nonlinear unmixing using neural networks.

Other types of neural networks besides MLPs have been applied to the spectral unmixing problem with the goal to estimate accurate abundances of materials within a mixture. Licciardi and Del Frate (2011) [28] used autoassociative neural networks to extract accurate pixel abundance estimations from the hyperspectral data. Guilfoyle et al. (2001) [18] use a radial basis function neural network to perform abundance estimation on nonlinear binary mixtures. The radial basis function network concept has been applied to various time series problems. Time series data often have common data structures and features to hyperspectral data so detection algorithms such as this network can be shared between the two modalities.

The research presented in this paper builds on the neural network applications discussed by applying yet another type of network (CNNs) to the low abundance intimate mixture detection problem. Of the neural networks that have been applied to the spectral unmixing problem, few, if any, have explored the use of convolutions in the hidden layers of the network [40]. The idea is that using CNNs may increase the confidence of detection for an analyst through processing subtle spectral features such that they're more discernable from the background.

Convolutional Neural Networks

Machine learning and CNNs are active areas of research for applications in computer vision / image processing. In the last decade, advances in DL/CNN methods and application programming interfaces (APIs), such as Google's TensorFlow (used in this research), have made it easier for scientific researchers to apply such processing to their problems. There are a few, now famous, networks that have paved the way for methods that are used in research discussed in this paper. AlexNet [24], VGG Net [37], ResNet [19], and the fully connected network (FCN) [29] are all examples of such networks. Although these networks weren't used "out of the box" in this research, some features of these networks that result in lower error rates were used.

Machine learning with CNNs has been widely proven to be an accurate means of identifying and typing features for various forms of data including estimating nonlinear functions and detecting features in one-dimensional data series [16]. Typical use, however, are applications to spatially unique objects in two-dimensional images with red,

green, and blue color channels [25]. In remote sensing image data, materials that need to be detected, identified, and characterized can appear spatially similar but spectrally unique. Specifically, in HSI data, that uniqueness takes the form of spectral absorption features when looking at the spectral dimension of a pixel within the HSI cube (3rd dimension or 3D). 3D CNNs have been applied to HSI data to detect material classes, considering both spatial and spectral information in one operation [11]. However, the 3D operation may not be desirable for a problem such as low-abundance spectral unmixing in spatially homogeneous materials. CNNs use kernels to transform the input data into unique classes in order to categorize them. Other kernel-based methods for classifying hyperspectral images, have been used to generalize linear algorithms to nonlinear data [8, 27].

HSI signatures in the spectral dimension are most often unique between material types (hence the colloquial term "signature.") The signatures are due to varying absorption of photons as a function of wavelengths. In nonlinear mixtures, mass fraction, density, particle size, and single scattering albedo will affect the size and shape of the absorption features. One method of analyzing spectra is derivative spectroscopy; a way of identifying and characterizing spectral features based on first- or higher-order derivatives [31]. One example of this method is shown in Figure 2 from a soil discrimination study by Bannari et al. (2018) [7]. The figure shows the reflectance spectra collected alongside the same spectra with first derivative processing done in order to be able to distinguish similar soils that would not otherwise be distinguishable.

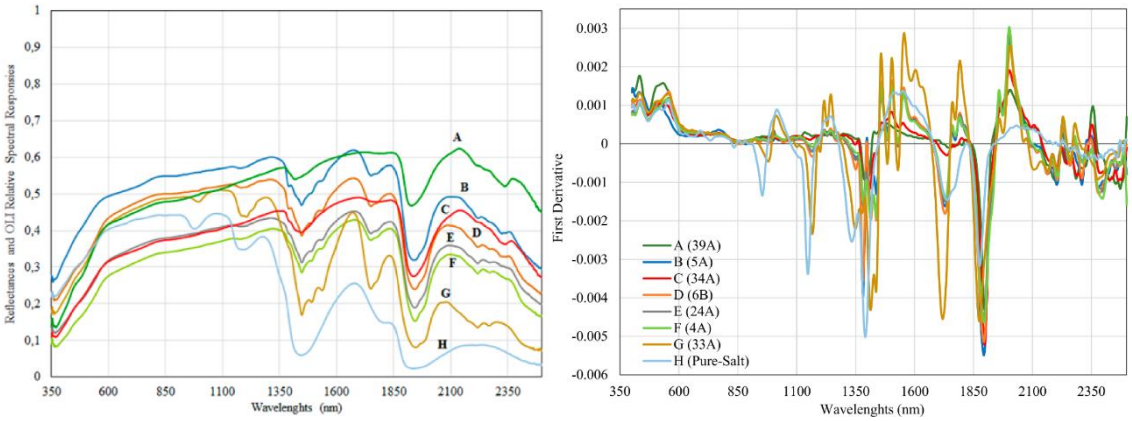


Figure 2. Example reflectance spectra and continuum removed first derivative spectra. Plots are from research by Bannari et al. (2018) [7].

A properly designed kernel convolved with spectra will lead to an analogous result to that of derivative spectroscopy, aiding the categorization of spectral components. The flexibility in CNN structure and number of kernels that could be applied within a CNN greatly expands the possibilities of derivative-like analysis. In two-dimensional CNN applications, it is commonly seen that a CNN generates weights that produce outputs like that of an image convolved with an edge detection kernel. One aspect of the research presented in this paper is how the training of a one-dimensional CNN will generate the edge-detection-like or derivative-like kernel weights such that it finds spectral uniqueness in each class of data, resulting in high detection accuracies. Considering the success of processes like derivative spectroscopy for HSI signature discrimination and the capability of convolutional processing to extract a derivative, applying CNNs to HSI data warrants investigation.

Microscenes

Microscenes are a method of generating data for the evaluation of hyperspectral systems and algorithm performance. The concept, introduced by Allen et al., (2013) [2], is to generate a small-scale setting that consists of the material constituents that would be found in a real-world scene. Applications of microscenes are typically, but not limited to, the simulation of spectral data that would be acquired during airborne or spaceborne hyperspectral remote sensing data. Benefits of microscenes include cost-effectiveness of a laboratory system vs. an air- or spaceborne remote sensing system, the customization of scene constituents, and the efficiency of data collection for analysis. Commercially available laboratory HSI sensors can acquire several hundred thousand spectra in one hyperspectral image acquisition of a Petri dish (or other suitable sample tray) thus furnishing many labeled samples. Imagery derived from microscenes contains many properties in common with non-laboratory remote sensing systems and have been shown to be useful as a proxy for overhead imagery [4, 35, 36].

Lab-based HSI data of microscenes both qualitatively and quantitively represent accurate proxies to actual remotely sensed hyperspectral data [2]. Importantly for this research, Allen et al. (2013) [2] assert that the spectral mixing interactions, among other properties, are fully represented naturally in a microscene. This is a key aspect of data quality for machine learning research that may not be accurately represented in modeled or simulated data. Additionally, the use of microscenes for machine learning training data specifically highlights two benefits related to the large amounts of data required for accuracy. First, the inexpensive and rapid collection of large amounts of data is easily

accomplished. Second, this large amount of data is easily labeled because the user has made the microscene and therefore knows exactly what materials, and quantities, are in the data. Each sample of training data doesn't have to be human or algorithmically post-processed in order the label the data correctly. Furthermore, an intimate mixture can be generated with such small weight percentages of materials that the user can't visibly or spectrally see it but is certain the materials are present due to the process of how the mixture was created. This provides an advantage that is not widely considered in other machine learning applications. It is a novel method when the goal is to determine MDQs since the scene generation can produce test cases that are below the limits of human or current, "traditional" algorithm interpretability.

METHODS

The methods discussed in this paper include the generation of microscenes for 3 unique intimate mixtures. The mixtures were then imaged with SWIR hyperspectral sensor and labelled with their respective weight percents for input into processing algorithms. 6 techniques were tested. The designs of these techniques are discussed in this section, and the results shown in the following section. Figure 3 shows a diagram of the flow of data for this research.

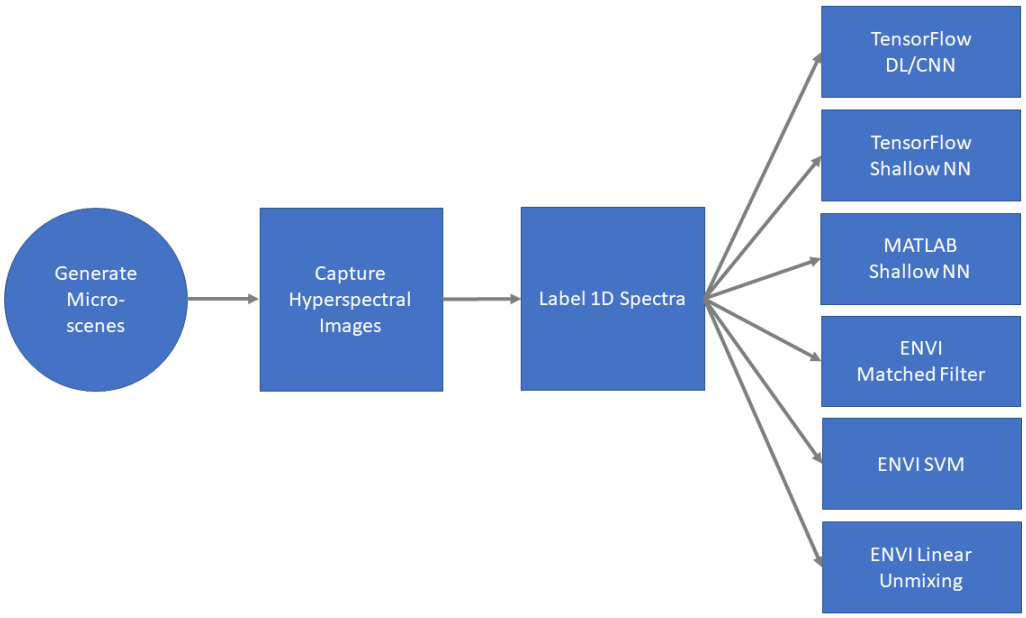


Figure 3. Data generation and processing overview.
The flow of data in this research consisted of generating the microscenes, capturing the hyperspectral image data, labelling the data, and then processing it for mixture detection using 6 different methods.

Microscenes and HSI Data Collection

Microscene Generation

This study used 150 mm diameter plastic Petri dishes in which to create microscenes. The individual Petri dishes respectively contained varying weight percents of chemicals mixed with a common base material. Table 1 outlines the 3 binary mixtures and weight percents in each microscene. Each mixture is described further in following sub-sections. A hyperspectral image of a single Petri dish generates several hundred thousand labeled spectra for analysis—as individual spectra (i.e., individual pixels extracted from the HSI data) or as n-pixel mean spectra where n can be any integer value. Single-pixel and 50-pixel average spectra were analyzed for the sand + Nd_2O_3 and 50-pixel average spectra, only, were analyzed for sand + powdered sugar and soil + aspartame.

Sand + Nd_2O_3

The sieved silicate sand comes from a commercially purchased 50-lb. bag of "sand-box" play sand. The sand fraction of the mixtures was 500 micrometers (35 mesh) and smaller. The reagent-grade Nd_2O_3, essentially a fine powder with particle size much smaller than the sand (several orders of magnitude smaller; essentially a clay), was very thoroughly mechanically mixed with the sand; it is thus intimately mixed with the sand. It is believed that due to the particle size difference and texture of the Nd_2O_3, the powder is sticking to and coating the outside of the sand particles, causing the multiple photon interactions expected in a mixture defined as intimate. Nd_2O_3 was chosen because of the

spectral features it presents in the shortwave infrared (SWIR; 900 nm to 2500 nm)[2] region of the spectrum, their location in contrast to the mean spectrum of sand, its low toxicity, and its high density. The low weight percent samples contain very low abundances (from an earth remote sensing perspective) of the chemical. Spectra of the Nd_2O_3 mixtures are shown in Figure 4.

Sand + Powdered Sugar

The powdered sugar mixture used the same sand as described in Section 2.1.1 and was chosen for many of the same reasons as the Nd_2O_3. Spectra of the powdered sugar mixtures are shown in Figure 4.

Soil + Aspartame

The soil + aspartame mixture was, in part, chosen due the finer particle size of soil vs the sand. The soil used is a clay and likely has similar particle size as the aspartame. Spectra of the aspartame mixtures are shown in Figure 4.

[2] See section "The HSI Measurements" for more details about the SWIR imaging system used in this research.

Table 1. Weight Percent of Mixtures for each Microscene

Mixture	Weight percent Reagent								
Sand + Nd_2O_3	0	N/A	N/A	0.5	1	2	3	4	5
Sand + Powdered Sugar	0	0.1	0.25	0.5	1	2	3	4	5
Soil + Aspartame	0	0.1	0.25	0.5	1	2	3	4	N/A

The HSI Measurements

The HSI data of the microscenes were acquired with a Headwall Photonic, Inc., Micro-Hyperspec® SWIR hyperspectral imaging spectrometer. The sensor was set to acquire 168 bands (of a possible 267) in the ~898 nm to ~2508 nm range. During data processing, the first four bands were discarded due to noise content; the analyses described here are thus applied to 164 bands from 936.50 nm to 2508.20 nm. A translation table moved the microscene by the sensor; illumination was provided by a quartz-tungsten-halogen lamp. Other features of the Micro-Hyperspec® SWIR and a photograph of the apparatus in the laboratory at the National Institute of Standards and Technology[3] may be found in Appendix I. Ground sample distance of the resulting image data is ~0.3 millimeters (mm).

[3] The HSI measurements used in this study were made at the National Institute of Standards and Technology (NIST) in Gaithersburg, MD.

Conversion to Reflectance

The microscene data are easily converted to reflectance by also collecting an HSI cube of a plaque of polytetrafluorethylene (PTFE), a dark current cube (a collection event with the lens cap on), and a point-spectrometer measurement of the reflectance of PTFE. Converting all raw measurements to floating point numbers, reflectance is calculated as shown in Equation 2, where DN is digital number. Conversion to reflectance is accomplished with ENVI's "Spectral Math" capability.

$$Reflectance = \left(\frac{microscene\ DN - Dark\ Current\ DN}{PTFE\ DN - Dark\ Current\ DN}\right) * PTFE\ Reflectance$$

Equation 2: Conversion to Reflectance

Spectra from the Microscenes

Two forms of the spectra are analyzed in this study: single pixel spectra and 50-pixel average spectra. Figure 4 shows a plot of a randomly selected 50-pixel average spectra from each mixture and within each plot, each weight percent. There is a noticeable change in apparent spectral features from the lowest to the highest weight percent of each mixture, however, only slight differences or no difference at all is seen among the lowest weight percents. This is highlighted in Figure 5, showing the mean spectra of 50,000 samples of the 0% mixture and the next lowest weight percent mixture. Besides some differences in total reflectance (standard deviation among 50,000 spectra is ~2% reflectance), the spectral feature differences are not identifiable by human visual

inspection. For this reason, the neural network processing methods are being studied to find their minimum abundance detection limit.

Regions of interest (ROIs) were defined for each of the microscenes. In most cases, an ROI was composed of numerous boxes placed to avoid obvious specular grains. The spectra within each ROI comprise a set of data for training and validation (for which samples are held back from the training phase). An ROI such as that shown in Figure 6 contains just over 50,000 spectra; indeed, ROIs were chosen to capture ~50,000 spectra—an arbitrary quantity assumed sufficient to form a good set of labeled measurements from which samples for training and validation may be drawn. It is evident from Figure 6 that many more spectra could be extracted from a given microscene. The spectra from an ROI are then output to an ASCII text file and can be imported to the neural network software or other tools.

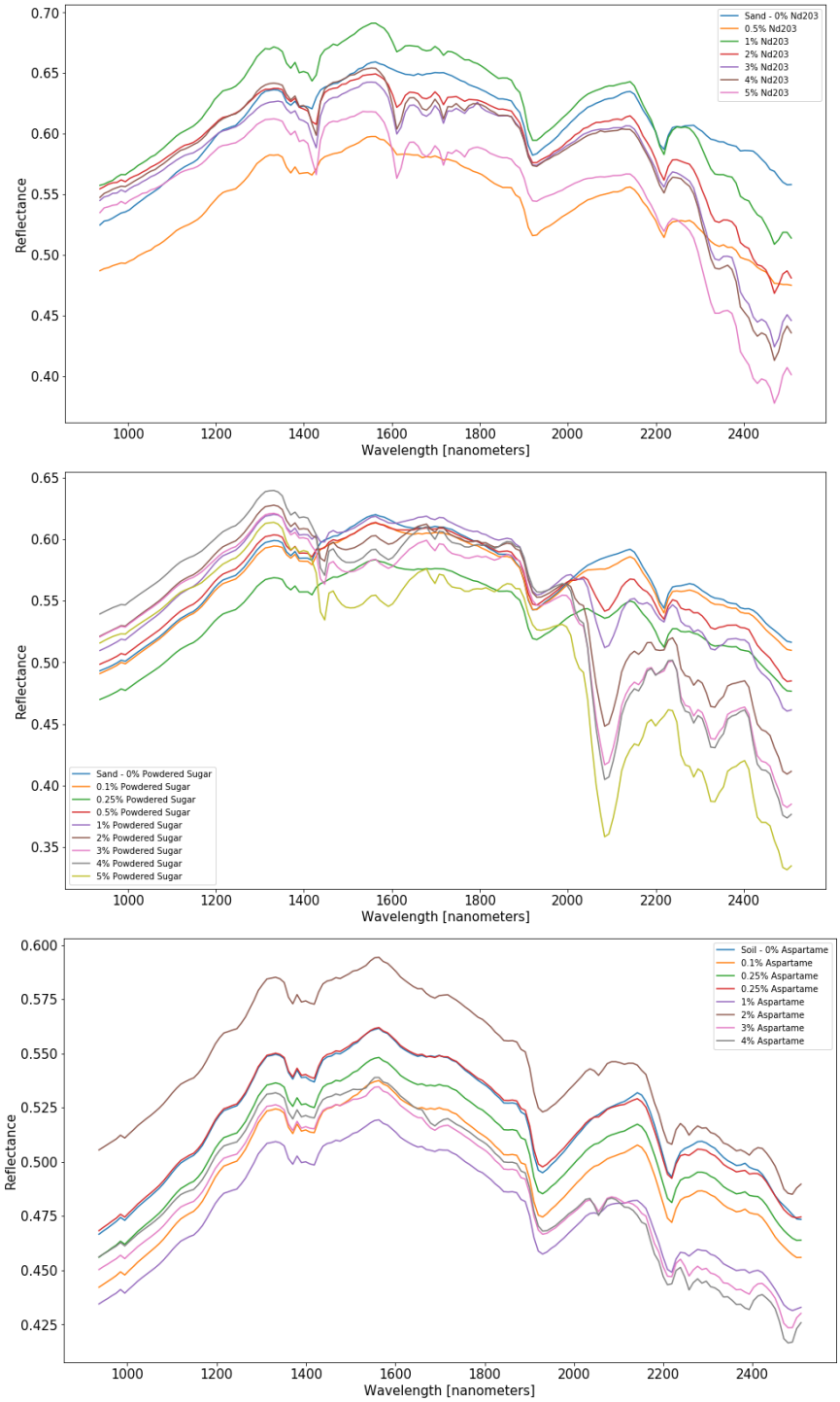


Figure 4. Example spectra from each mixture.
From top to bottom: (a) Nd_2O_3 + sand, (b) powdered sugar + sand, (c) aspartame + soil. The spectra shown are randomly selected individual spectra among the 50,000 samples used in the CNN training/validation. Each mixture's diagnostic spectral features are indications of the weight percent of the mixture.

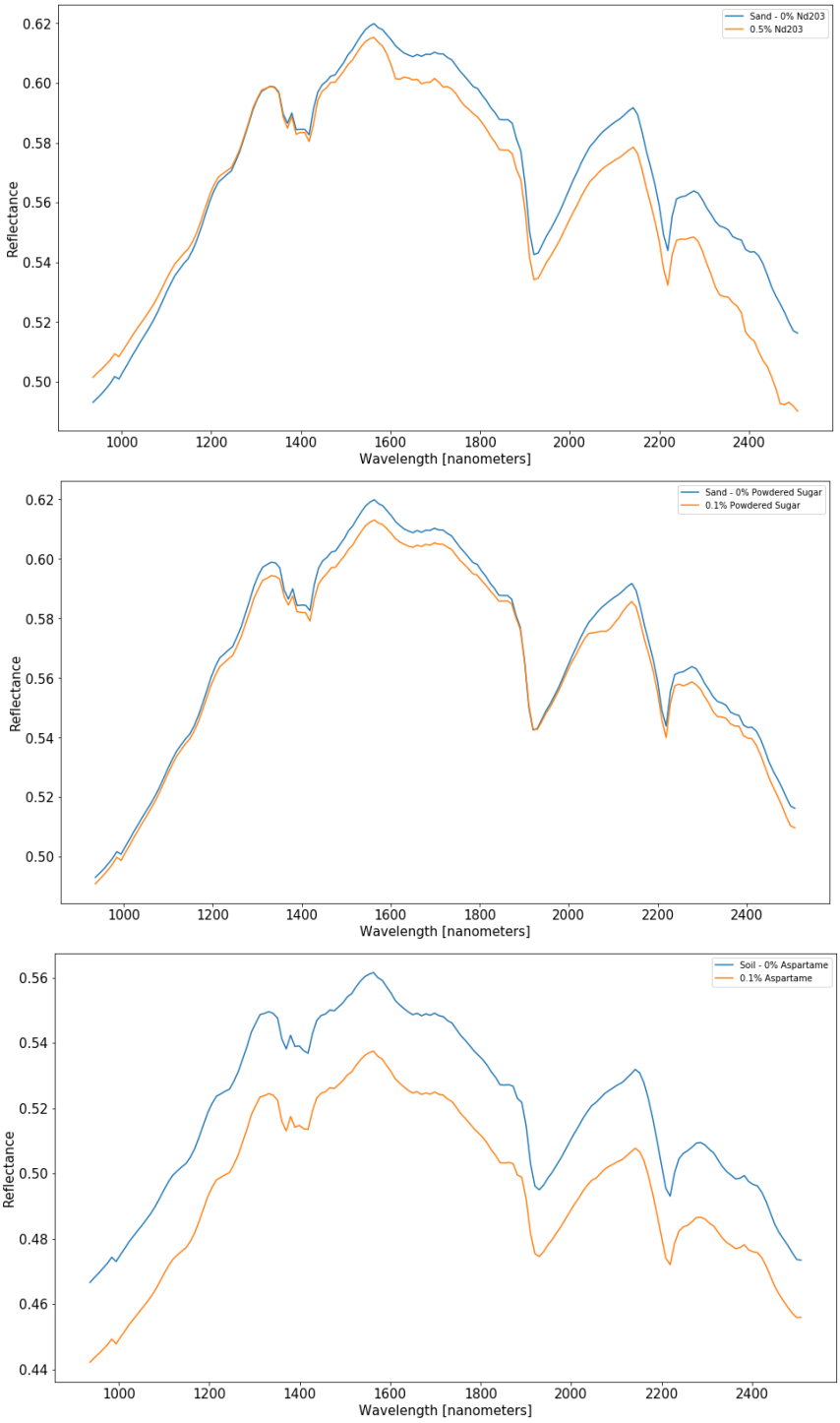


Figure 5. Example spectra from each mixture of no reagent and the lowest weight percent mixture. From top to bottom: (a) Nd_2O_3 + sand, (b) powdered sugar + sand, (c) aspartame + soil. The spectra shown are the mean spectra among the 50,000 samples used in the CNN training/validation. Each plot shows the base material (0% reagent) spectra and the lowest weight percent mixture. Note the lack of discernable features in each comparison. The standard deviation of the mean of each set of spectra is ~2% reflectance.

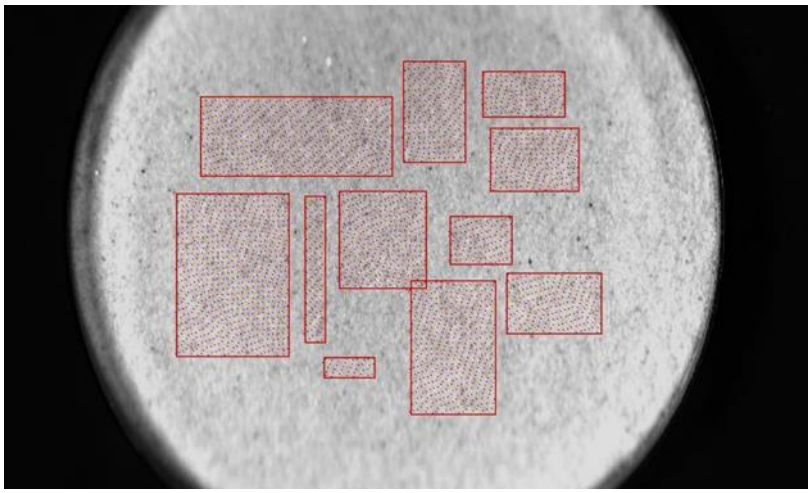


Figure 6. Microscene context image.
A grayscale image (1727.16 nm; 2 % linear stretch) of the microscene with 1.0 weight percent Nd_2O_3. The red boxes are regions of interest (ROIs) for extracting spectra. In a figure in Appendix I the microscene can be seen on the translation stage beneath the HSI sensor.

Characterizing the Sand

There is a broad range of spectral signatures in the sand. This is evident in the speckled nature of the sand in Figure 4. However, most of the grains are essentially the same as the mean spectrum. The grains are quartz, feldspars (most common minerals in the earth's crust) and a grain or so of mica and other accessory minerals. Within the Petri dish, there are also spectra due to specular reflections from various grains randomly oriented to produce such an effect, grains showing etaloning, and spectra with real sensor noise artifacts (rare, small random spikes). This spectral signature diversity is desirable because it adds a reality to the data that is also in airborne remotely sensed HSI. And it can be moderated somewhat by averaging clumps of pixels (and is indeed moderated by a point spectrometer measurement with a spot size on the order of millimeters to centimeters in diameter). The impact of light passing through the transparent plastic Petri

dish wall, evident especially at the left and right sides of the microscene in Figure 7 imparts plastic spectral features. For the generation of signatures for analysis here, these pixels are ignored as is evident from the placement of the ROI in Figure 6.

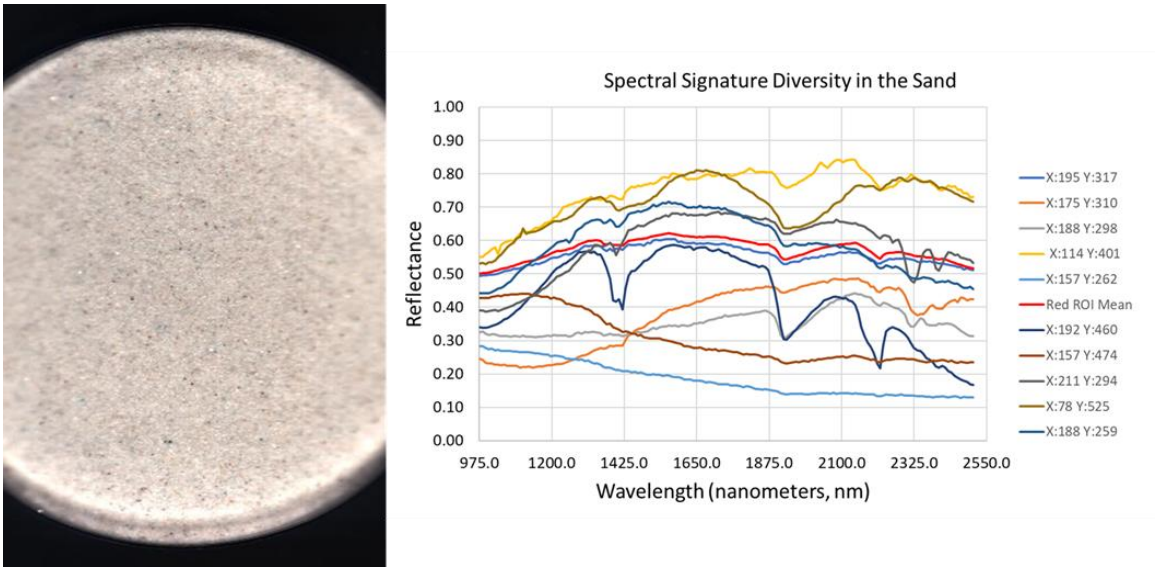


Figure 7. Signature diversity in the Sand.
(left) A SWIR false color composite image of the sand-only microscene. (right) Single-pixel spectra from the microscene at left. The red trace ('Red_ROI_Mean') is the mean spectrum of a region of interest (containing approximately 50,000 spectra) for the sand-only microscene (not shown). The plot key indicates the (sample, line) address of the pixel in the microscene.

HSI Data Analysis

Applying SVM to Sand + Nd_2O_3 Mixture

An HSI cube was formed from the 50-pixel spectra extracted from the respective ROIs in the seven sand + Nd_2O_3 microscenes. Thirty thousand spectra from each microscene are used for the cube. Half of the spectra in each class (i.e., the different weight percents) are designated as training; the remaining half are used for testing. Only one trial with the following settings (entered in ENVI's 'SVM Options' window) is

reported in this paper: the kernel type is a radial basis function; the gamma in kernel function is 0.006; the penalty parameter is 100.0; and the pyramid levels is zero as is the classification probability threshold [10].

Applying Spectral Mixture Analysis (SMA) to Sand + Nd_2O_3 Mixture

An analyst not knowing the nature of the mixture or deposit of Nd_2O_3 on (and churned/mixed within) the sand may attempt to find the reagent by a "simple" linear unmixing analysis for which spectra of the reagent and of the sand are endmembers. Proceeding in this manner, ENVI's linear spectral unmixing was applied to the individual spectra in the ROIs of the microscenes. It was run completely unconstrained using a spectrum of pure Nd_2O_3 and a mean sand spectrum (red trace in Figure 7) as endmembers. ROI mean fractions (along with mean +/- one standard deviation) of Nd_2O_3 are plotted vs. weight percent Nd_2O_3 in Figure 13.

The Matched Filter (MF)

A separate microscene was generated within which small aliquots of the 0.5, 1, 2, 3, 4, and 5 weight percent Nd_2O_3 + sand samples are emplaced along with several other materials. The scene created a realistic analog, in terms of signal-to-clutter ratio (SCR), as that presented in an airborne HSI data cube (indeed, microscenes can be built that are more stressing in terms of SCR than an airborne data set). ENVI's MF algorithm was applied using a pure Nd_2O_3 spectrum (the same one used in linear spectral unmixing) [15]. The statistics required by the algorithm (scene mean spectrum and covariance

matrix) were generated with a mask over the Nd_2O_3-containing aliquots to minimize target signal leakage that would degrade the MF performance. This trial is conducted to demonstrate that even when Nd_2O_3 is present, the MF scores can be very low (particularly in the 0.5, 1, and 2 weight percent aliquots) such that using typical data analysis practice, those pixels would most likely be deemed free of the reagent. The scores are low for two reasons: the abundance of the Nd_2O_3 is low (again, particularly in the 0.5, 1, and 2 weight percent aliquots) and it is nonlinearly mixed with the sand such that expecting a "good" match with a spectrum of regent-grade Nd_2O_3 is unrealistic.

Deep Learning Convolutional Neural Network (DL/CNN)

The DL/CNN is coded in Python using Keras with a TensorFlow-GPU backend. The computer built for processing runs Linux Ubuntu 18.04, has a 1 TB solid-state drive, 32 GB DDR4 2133 MHz DRAM, an NVIDIA GTX 1080 GPU, and an Intel Xeon E3-1270 CPU. The high-end gaming GPU is appropriate for machine learning: it has 9 TFLOPS of processing capability and typically runs TensorFlow training and validation processing in ~10 minutes for 25 epochs of ~400,000 samples. The model uses three one-dimensional (1D) convolutional layers to form a deep learning neural network. Similar structures are commonly used in processing time-series data—but here applied naturally to spectra which are also 1D data structures. Some parameters, including number of filters and kernel size, have been optimized for the input data; however, the overall structure has not yet been exhaustively studied for the application to HSI data. One

example DL/CNN model summary output from Keras is shown in Table 2. A detailed diagram of the DL/CNN layer structure is shown in Figure 8.

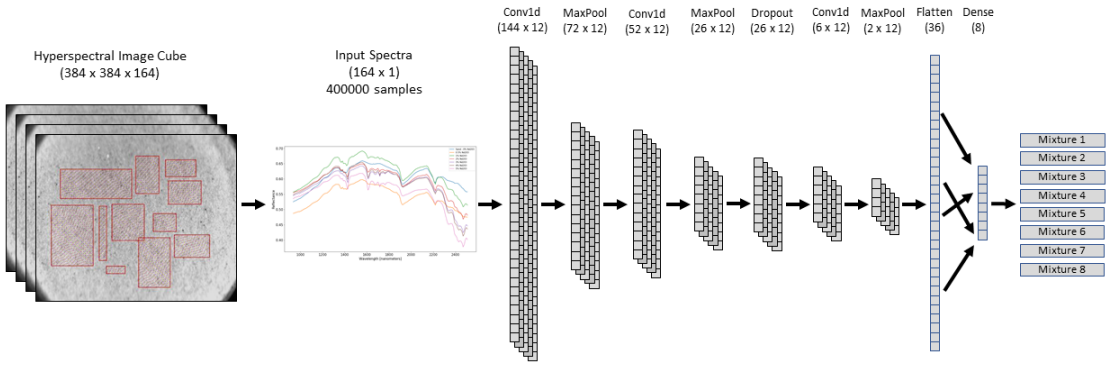


Figure 8. DL/CNN structure diagram.
1D data extracted from a 3D HSI data cube are inputs to the DL/CNN which applies a series of convolutional layers and a fully connected layer to label the weight percents of each material mixture.

Table 2. Keras DL/CNN Model Summary Output

Layer	Output Shape	Trainable Parameters
Conv1D	(None, 144, 12)	264
Activation	(None, 144, 12)	0
MaxPooling	(None, 72, 12)	0
Conv1D	(None, 52, 12)	3036
Activation	(None, 52, 12)	0
MaxPooling	(None, 26, 12)	0
Dropout	(None, 26, 12)	0
Conv1D	(None, 6, 12)	3036
Activation	(None, 6, 12)	0
MaxPooling	(None, 2, 12)	0
Flatten	(None, 36)	0
Dense	(None, 8)	296

Shallow Neural Networks

A shallow (i.e., a single hidden layer), fully connected, feed-forward neural network implemented in MATLAB (v. R2018a) was applied to the sand + Nd_2O_3 microscene spectra. Each class contained 15,000 spectra from which user-specified fractions were held back from training for testing and validation. Scaled conjugate gradient backpropagation was applied for training (though other training methods available in the MATLAB implementation were tested). Different numbers of nodes (also referred to as neurons) in the hidden layer (5, 10, and 20) were also tested. The MATLAB neural network graphical user interface (GUI) was used to implement this capability (though it can also be implemented as a series of commands at the command prompt or as a .m file).

For comparison to the MATLAB shallow neural network, a shallow, fully connected, feed-forward neural network implemented in Python using Keras with a TensorFlow-GPU backend was applied to the sand + Nd_2O_3 microscene spectra. This shallow network was branched from the same code used for the DL/CNN. The only change was that the 3 layers of CNNs with pooling were replaced by a single "Dense" layer. The focus on this paper is to explore the potential benefits of CNNs for HSI data analysis, but many other types of neural networks can also classify 1D data with high accuracy. This shallow network provides a baseline for comparison to the DL/CNN as well as a baseline for further exploration into other TensorFlow functions and model structures.

Neural Network Metrics

Two metrics were used to assess neural network performance. The first is validation accuracy. Validation accuracy is the total number of correctly categorized samples divided by the total number of samples in that category. This value is calculated from the validation dataset, a random sampling of spectra that is separated from the spectra that is used to train the network. The second metric used is F-1 score. F-1 score is the harmonic mean of precision and recall, reaching its optimal value at 1 and its worst value at 0 (Equation 3.) F-1 or F-beta (a weighted version F-1) is often used in neural network and other classification applications as a measure of accuracy. Throughout the paper, validation accuracy will be most discussed. A summary of each network's mean F-1 score can be found in the Results section in Table 4.

$$F_1 = 2 * \left(\frac{Precision * Recall}{Precision + Recall}\right)$$

Equation 3: F-1 Score

DL/CNN Parameter Optimization

Several parameters in the CNN model were optimized to improve results from a baseline validation accuracy of ~90 % to an accuracy of >98 %. The method of optimizing was to functionalize the CNN training/validation runs such that they iterate over a range of inputs and log the details of each run for later analysis. After the multiple 1D CNN structure was formed and produced a baseline of success, two arguments in the "Conv1D" Keras function were identified for optimization [20]. The "filters" and "kernel_size" arguments were iterated over six values each. These arguments were

chosen due to the possible effects they have on the kernel weights being applied to spectral absorption features. Values of "filters", the dimensionality of the output space (i.e., the number of output filters in the convolution), ranged from 6 to 192. Values of "kernel_size," the length of the 1D convolution window, ranged from 3 to 15. Figure 9 shows the validation accuracies for these ranges of values and for each material mixture. Although the results are similar, each model has different optimized values for "filters" and "kernel_size". The Optimized values are shown in Table 3.

Dropout rate, convolution activation function, and max pooling "pool_size" were also tested – though somewhat less rigorously. It was decided based on ad-hoc tuning of these parameters to use three common parameters in each network: a dropout rate of 0.5 to avoid overfitting [38], the hyperbolic tan function (tanh)[4], and pooling size of 2. These values consistently provided good results for each of the mixtures being tested and are also commonly used in various online tutorials for CNN development. The networks were trained over 25 epochs. Larger numbers of epochs (up to 100) were tested with no significant improvement to the validation accuracy. Some combinations of these values did not produce an adequate result from the training run (i.e., relu activation function, and pooling sizes > 10.) Those that did, however, produced results anywhere from 79 % to 99 % validation accuracy. Although there are many more possible layers, functions, variables, and ways to structure the CNN, the initial optimized results show successful application of this CNN to the 1D hyperspectral data.

[4] A list of Keras activation functions can be found at https://keras.io/activations/ - Accessed July 2019

Table 3. Optimized DL/CNN Parameters for each Mixture

50-Pixel Avg Mixture Model	Optimized "filters" Value	Optimized "kernel_size" Value
Sand + Nd_2O_3	12	21
Sand + Powdered Sugar	24	6
Soil + Aspartame	48	9

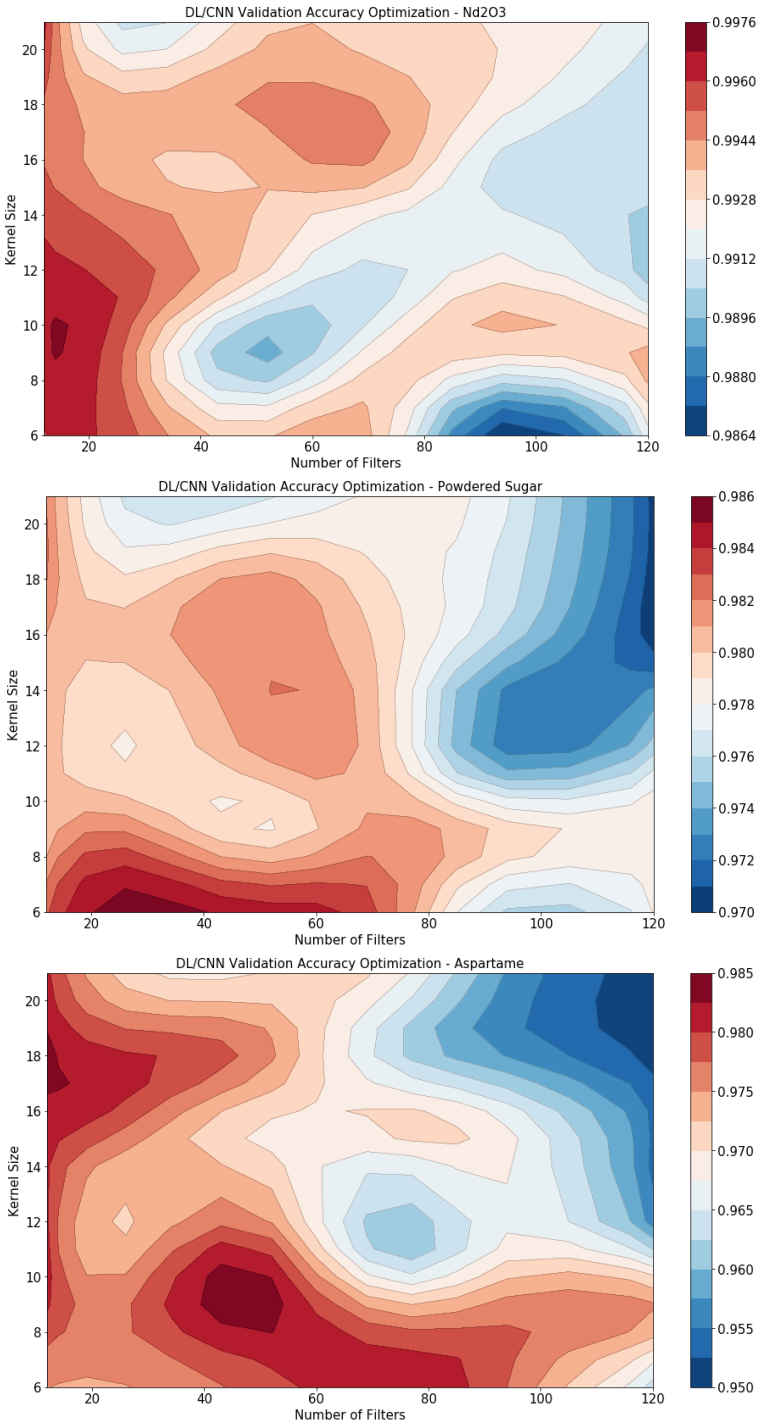


Figure 9. DL/CNN validation accuracies for each mixture as a function of kernel size and number of filters in each training/validation run.
From top to bottom: (a) Nd_2O_3+ sand, (b) Powdered Sugar + sand, (c) Aspartame + soil.

Training with a "Junk" Class

In any application of this network to a microscene or other remotely sensed scene, there will be spectra present other than those for which the CNN is trained. To ignore the materials that would be considered clutter or not material(s)-of-interest spectra, the TensorFlow networks were trained with a "junk" class. The junk class contained signatures unrelated to the sand and mixture classes. The signatures in the class were generated from the "usgs_veg.sli" and "usgs_min.sli" spectral library files bundled with ENVI [22]. These were chosen due to their availability, spectral resolution, wavelength range, and diversity of spectral features within the collections. A plot of the junk class spectra, showing these diversities, is in Figure 10. A test demonstration of making a synthetic microscene with implanted junk spectra was accomplished and demonstrates the separation of the spectra of interest and the junk spectra (See "Implementation for HSI Analysis" section). Confusion matrices of the validation accuracy of each CNN model including the junk class are shown in Figure 14. The MATLAB neural networks were not trained with a junk class; a difference to consider when comparing network accuracies.

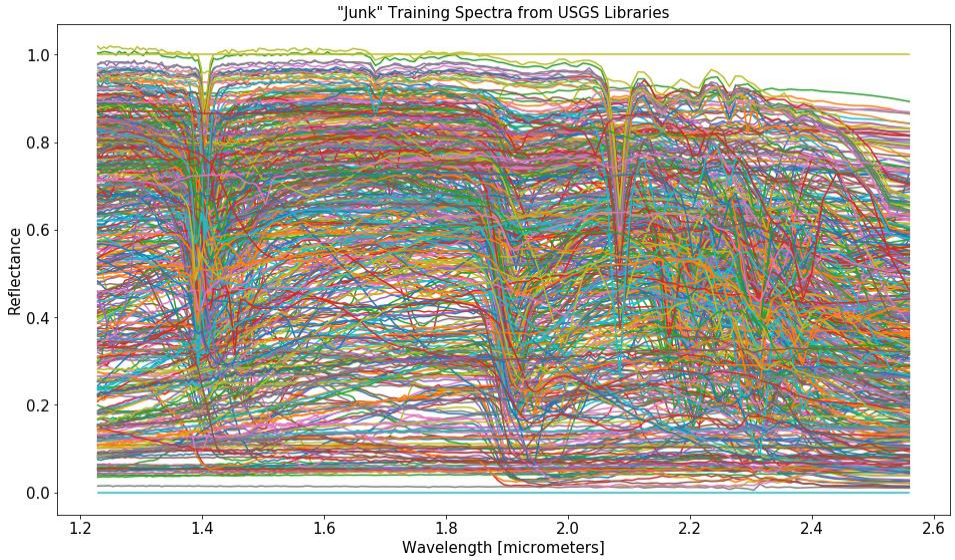


Figure 10. Plot of the junk spectra used in training the DL/CNNs.

Convolutional Layer Analysis

The convolutional functions of the neural network layers are applied to the spectra in such a way that spectral features will be transformed such that slight changes between signatures are more discernable. The hypothesis, as described earlier, is that the network's convolutions will learn derivative-like spectral-feature enhancement kernels (weights) such that unique spectral features are made discernable. To investigate whether this effect is occurring within the CNN layers, the weights of each convolutional layer are saved during each training run. Figures 11 and 12 show the pre- and post-training outputs from each of the three convolutional layers from the best performing Sand + Nd_2O_3 network. Figure 11 contains three plots that show a single spectrum convolved with the pre-trained weights in each convolutional layer. Since the network has three convolutional layers, the figure shows the first layer at the top and the third layer at the

bottom. Figure 12 shows the same spectrum convolved with the post-training weights in each convolutional layer. Note there are multiple traces in each plot because each convolution layer applies multiple “filters.”

The three post-training plots show the spectral data being transformed into unique signatures due to the kernel weights that the network established. They also show the reduction in data dimensionality mostly due to the max pooling layers. The third post-training layer shows the 6 x 12 output of an input spectrum. The signature at this step shows the data has been transformed into values near 1 and -1, which inform the network to which mixture that signature belongs. Each mixture will have a different signature at this step that the DL/CNN has learned through the training epochs.

The visualization of these layers indicates that the kernel weights are indeed transforming the data into unique signatures from each input mixture. The output of the first post-training convolution layer could be interpreted as the accentuation of spectral features, like that of derivative spectroscopy, as is hypothesized. However, the final layer doesn’t necessarily show the same resemblance, mostly due to the reduction of data dimensionality. The number of filters used, the kernel sizes, number of training samples, number of training epochs, all result in a complexity of data manipulation that could be studied further for the application of DL/CNNs for HSI data categorization.

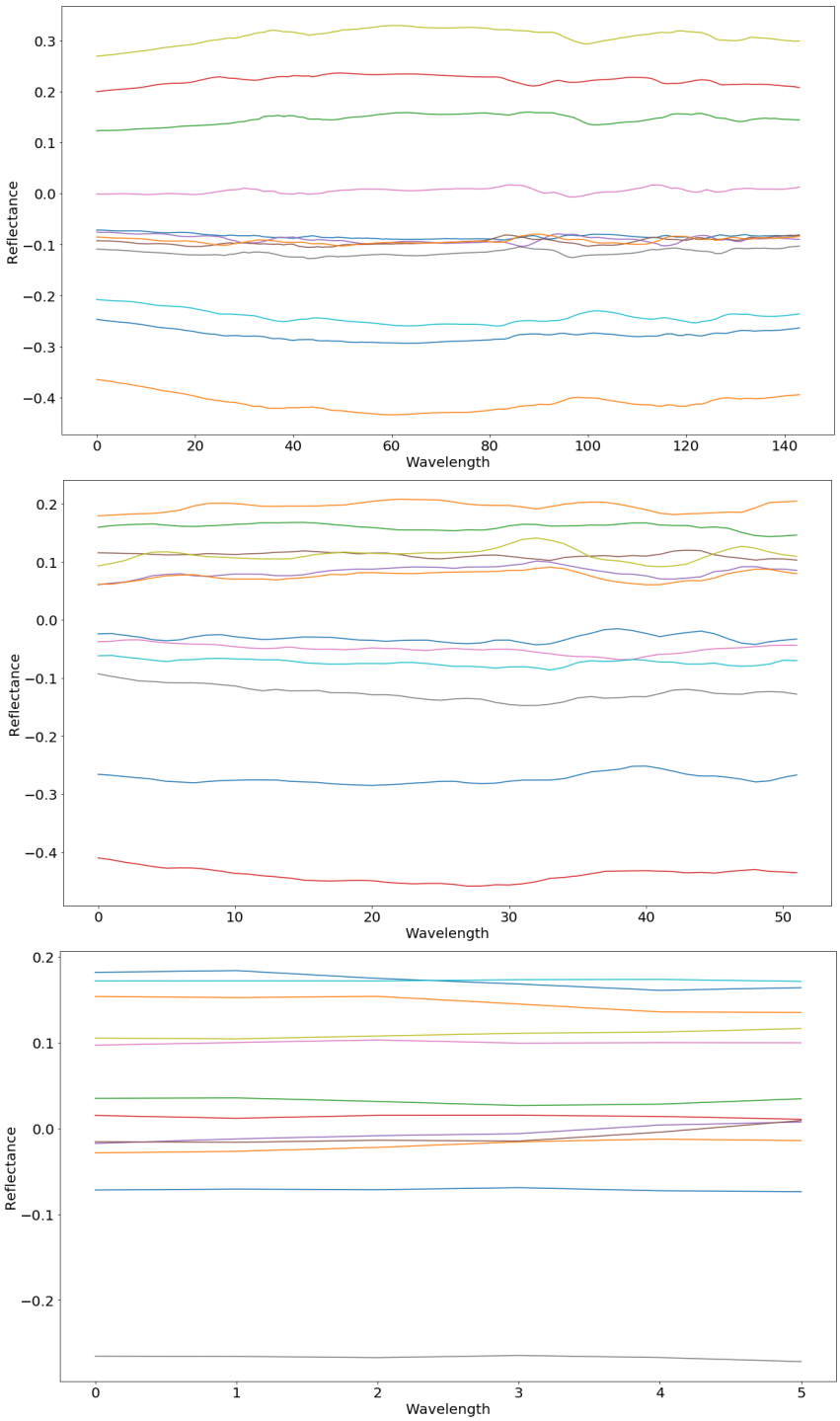


Figure 11. Pre-training convolutional layer analysis.
The three plots show a single spectrum convolved with the pre-trained weights in each of the three convolutional layers. The first layer on top, the third layer on the bottom. Note there are multiple lines in each plot, because each convolution layer applies multiple "filters." The Reflectance and Wavelength axes are representative of the dimension of those values however, the values themselves are arbitrary due to the filters applied to the spectra.

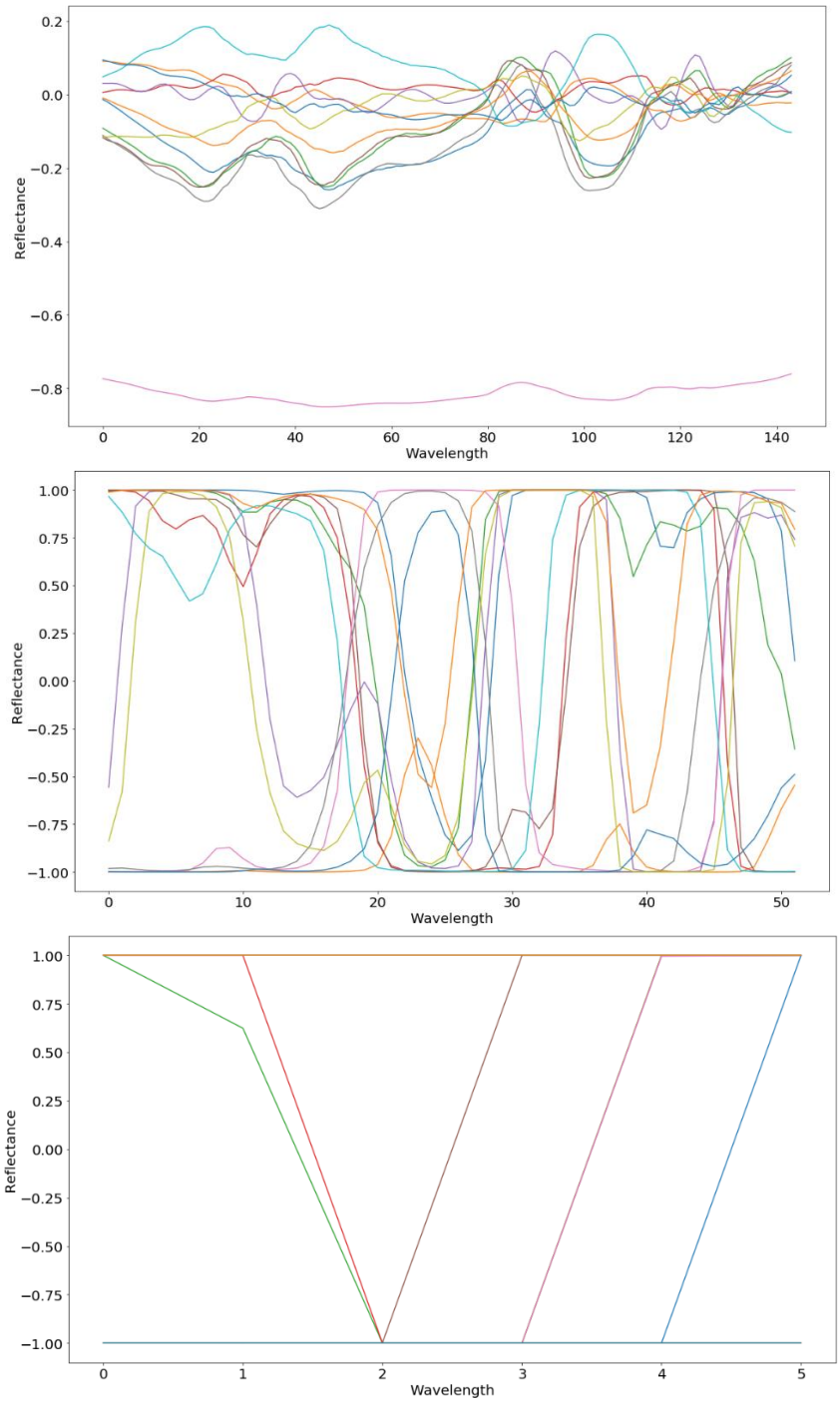


Figure 12. Post-training convolutional layer analysis.
The three plots show a single spectrum convolved with the post-trained weights in each of the three convolutional layers. The first layer on top, the third layer on the bottom. Note there are multiple lines in each plot, that is because each convolution layer applies multiple "filters." The Reflectance and Wavelength axes are representative of the dimension of those values however, the values themselves are arbitrary due to the filters applied to the spectra.

RESULTS

Baseline Analysis

The application of linear spectral unmixing, the support vector machine, and the matched filter are to present a set of results not for direct comparison to the neural networks, per se, but as examples of what more commonly applied and "traditional" algorithms yield. The "traditional" algorithms are all found in ENVI, a common, industry standard, HSI data analysis software package. An emphasis is implicitly placed on the low weight percent classes as these present the greatest challenge to current capabilities. The higher weight percent trials would, most likely, be detected with a suite of traditional or mainstay HSI analysis tools. Unless otherwise specified, the baseline ENVI analysis was done on 50-pixel average sand + Nd_2O_3 spectra.

Linear Spectral Unmixing

ENVI's "Linear Spectral Unmixing" (LSU) capability without constraints, the fractions obtained are shown in Figure 13. The nonlinear mixing is evident in the plot and in the relative magnitudes of the coefficients in the best-fit curve equation (corresponding to the blue dotted trace). The best-fit curve R^2 value is 0.9983 – an indication of the non-linearity of the estimated fractions, unrelated to the accuracy of the estimated weight fraction from SMA. For the 0.5 and 1.0 weight percent mixtures, the retrieved fractions

are very small (less than 0.1) and may be considered “noise” to an analyst. The higher weight percent cases (two to five percent) are confidently detected with LSU. No adjustments for density or particle size distribution have been applied to the retrieved fractions.

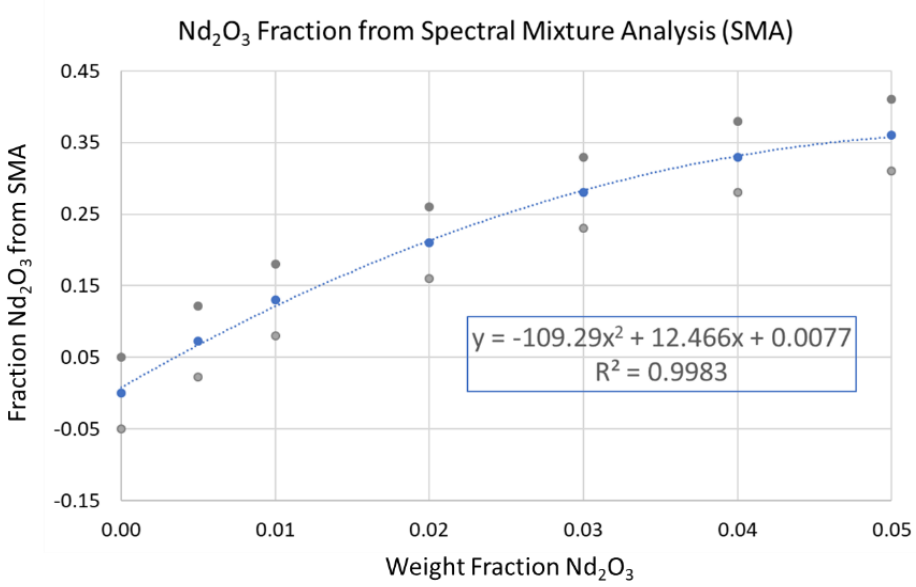


Figure 13. Spectral Mixture Analysis.
A roll-up of the fractions (vs. weight fraction [weight percent / 100.0]) in the mixtures. The gray symbols are mean +/- one standard deviation. The unmixing is completely unconstrained.

Support Vector Machine

A roll-up of ENVI’s statistics applied to the test samples is shown in Figure 14 (only the first four weight percents are shown for brevity). The SVM scores membership in a class with a value ranging from 0, meaning no membership, to 1.0, meaning membership. The SVM is also effective at separating the four lowest weight percents.

0% Nd_2O_3				
Basic Stats	Min	Max	Mean	stDev
0	0.995	1.000	1.000	0.000
0.5	0.000	0.003	0.000	0.000
1	0.000	0.001	0.000	0.000
2	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000

0.5% Nd_2O_3				
Basic Stats	Min	Max	Mean	stDev
0	0.000	0.018	0.000	0.001
0.5	0.980	1.000	1.000	0.001
1	0.000	0.006	0.000	0.000
2	0.000	0.001	0.000	0.000
3	0.000	0.001	0.000	0.000
4	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000

1% Nd_2O_3				
Basic Stats	Min	Max	Mean	stDev
0	0.000	0.001	0.000	0.000
0.5	0.000	0.056	0.000	0.001
1	0.942	1.000	1.000	0.001
2	0.000	0.001	0.000	0.000
3	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000

2% Nd_2O_3				
Basic Stats	Min	Max	Mean	stDev
0	0.000	0.000	0.000	0.000
0.5	0.000	0.000	0.000	0.000
1	0.000	0.012	0.000	0.000
2	0.987	1.000	1.000	0.001
3	0.000	0.007	0.000	0.000
4	0.000	0.005	0.000	0.000
5	0.000	0.001	0.000	0.000

Figure 14. SVM statistics for the sand + Nd_2O_3 samples.
ENVI-generated statistics for the first four classes, only, shown here for the validation spectra. (i.e., those not used in training). The class separation results are excellent. Class 2 is highlighted to emphasize that the SVM is quite capable of separating a small amount of reagent from the sand-only class.

The Matched Filter

As indicated previously, a separate microscene was created that contained small aliquots of the Nd_2O_3 + sand mixture samples. This microscene is built primarily to show that for low weight percent reagent samples, the MF target detection scores for the two lowest weight percent mixtures of Nd_2O_3 spectrum are very close to 0. On a scale from 0 to 1 (1 being a perfect match to a reference spectrum and 0 being the background,) the scores are 0.0 and 0.03 (Figure 15, third image from the left). Scores this low would likely be missed by current best practices for selecting pixels for further analysis. In Figure 15, the statistics for the application of the MF are based on 1) all pixels in the scene (third from left frame in Figure 15); and 2) those pixels containing Nd_2O_3 masked-out to avoid the target leakage effect (fourth from left frame in Figure 15).

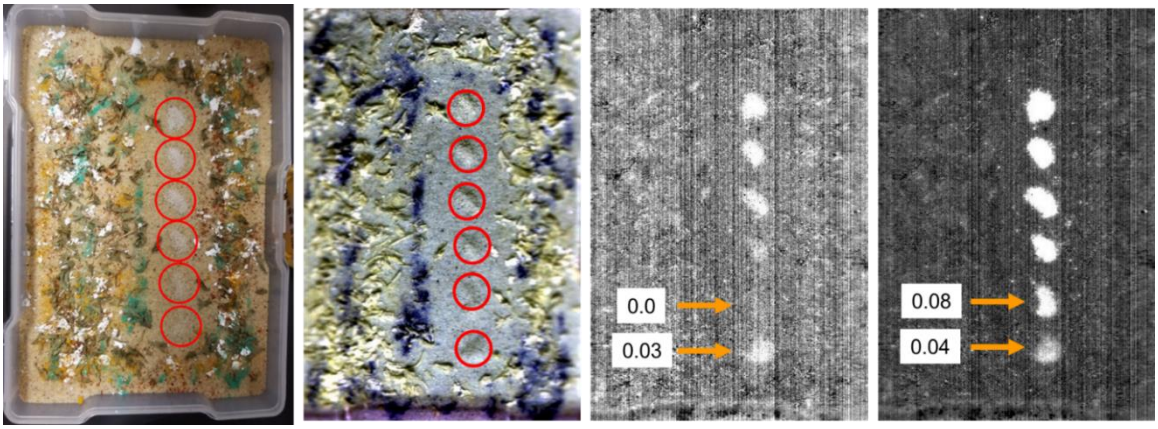


Figure 15. Matched Filter Detections.
(left to right). A digital camera photograph (i.e., normal color) of the microscene composed of sand, the six aliquots of Nd_2O_3 (red circles; top to bottom 5 weight percent to 0.5 weight percent reagent), and several other materials presenting a range of spectral signature information. A SWIR false-color composite of the microscene. A MF plane using all pixels in the microscene to generate the statistics. A MF plane generated with statistics that masked out the six Nd_2O_3-containing aliquots.

Overall, the three ENVI methods applied confidently detect the presence of higher abundance mixtures. The MF detection image in Figure 15 shows clear discrimination from the background for at least the three highest abundances. Fractions from the SMA for the 3-4 highest abundances are significant, indicating that >20% of the pixel being investigated is made up of Nd_2O_3. The MF scores and SMA fractions for the 0.5 and 1.0 weight percent mixtures that an analyst would likely consider consistent with the background of the image. The SVM, however, in addition to detecting the high abundance mixtures, was able to detect the lowest abundance mixtures.

Shallow Neural Network Results

Single-Pixel Spectra

The MATLAB NN results are shown in Figures 16 and 17. The low abundance mixtures (0, 0.5, 1, and 2 weight percent Nd_2O_3) are highlighted with a red box to show excellent separability. Performance lowers somewhat for the higher weight percent trials. This result warrants some further investigation, but current thoughts are that the Nd_2O_3 spectra are dominating, hindering the separability of the spectra after a certain weight percent is achieved. Varying the number of neurons in the single hidden layer has minimal impact on performance. Using 5 neurons was 90.7% accurate and using 10 neurons was 90.8% accurate. Figure 16 shows the results from 20 neurons which was 91.45% accurate. Again, an emphasis is directed towards 0.5 and 1 weight percent for which the scores average ≥ 98 %.

50-Pixel Average Spectra

Results obtained with the 50-pixel average spectra are better than for the single-pixel trials. The spectral signature diversity evident in Figure 7 is reduced; signal-to-noise ratio is increased by the averaging. As seen in the single-pixel results, varying the number of neurons in the hidden layer has minimal impact on performance here as for the single-pixel spectra trial above. Figure 17 shows the results when 10 neurons were used, which resulted in 100% accuracy. Using 5 neurons also resulted in 100% accuracy, and 20 neurons resulted in 99.9% accuracy.

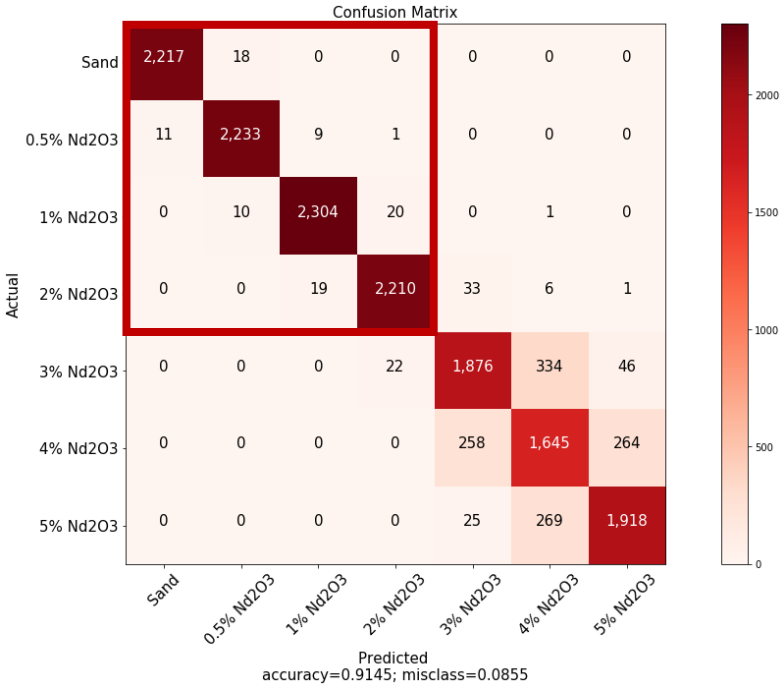


Figure 16. MATLAB Neural Network single pixel confusion matrix using 20 neurons.

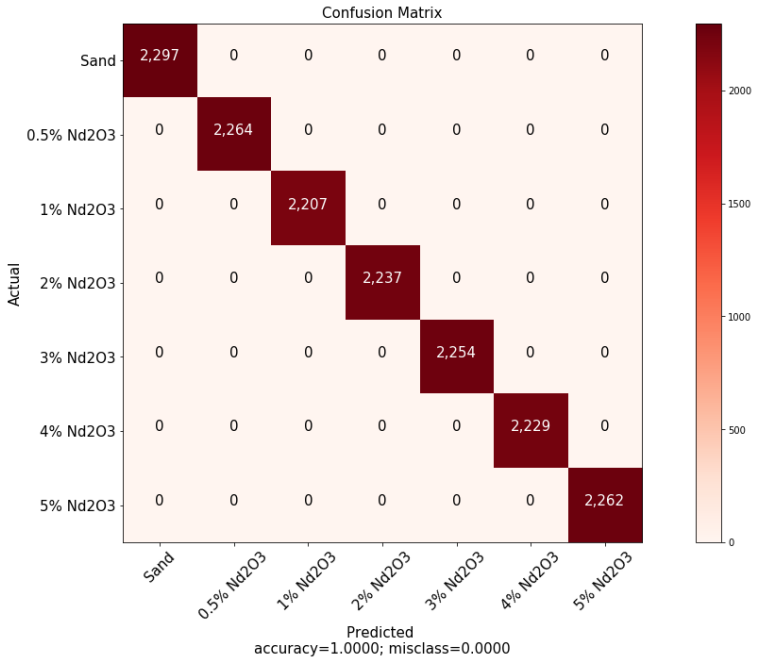
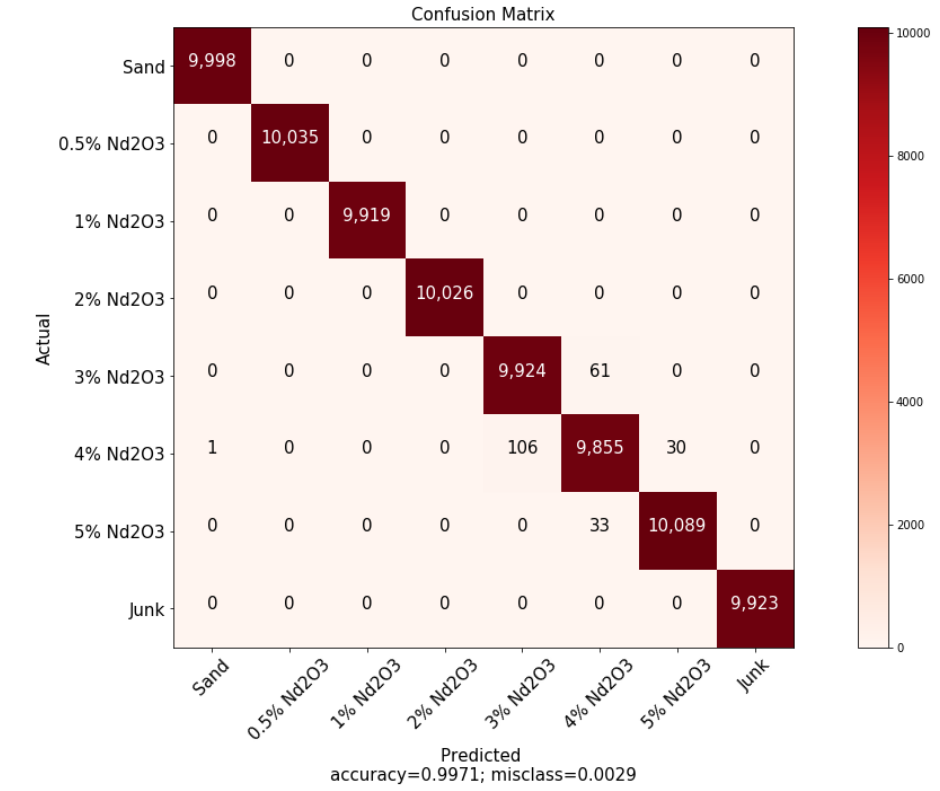


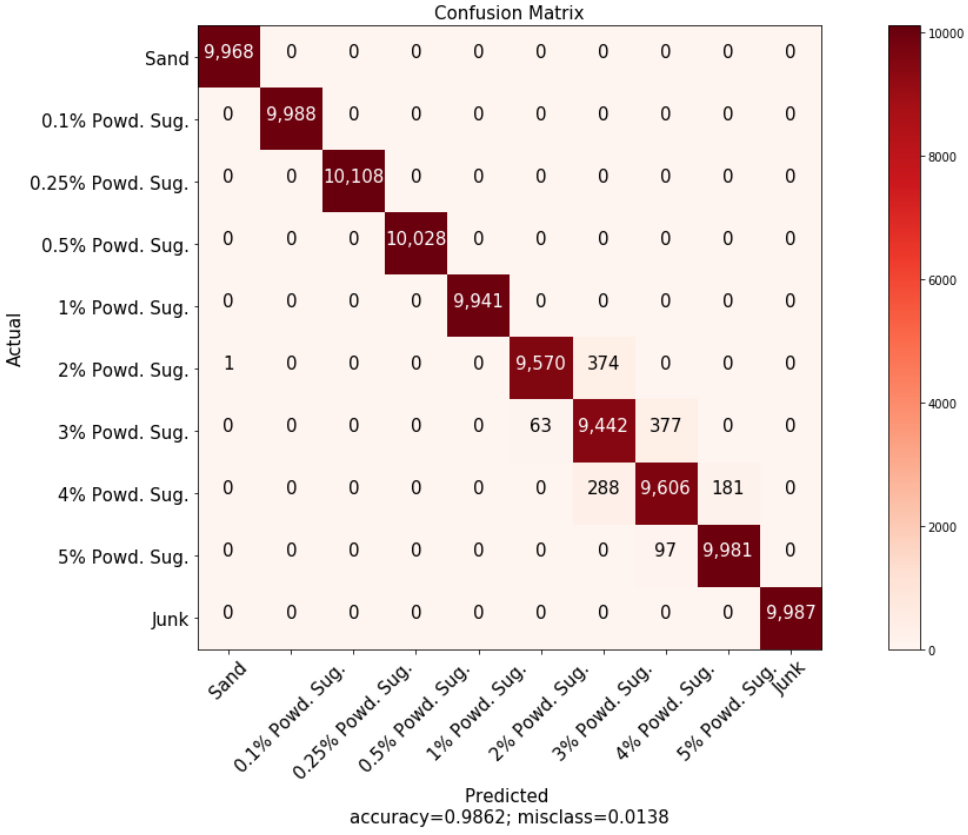
Figure 17. MATLAB Neural Network 50-pixel average confusion matrix using 10 neurons.

DL/CNN Results

DL/CNN 50-pixel Average Spectra for Three Mixture Types

Table 4 shows the overall validation results for the optimized DL/CNN of each mixture. The results show >98 % overall validation accuracy for each mixture tested. The results are encouraging that a CNN can discriminate the low abundance intimate mixtures. When visualized by validation accuracy vs. individual weight percent as shown in Figure 19, the results show two distinct outcomes dependent on the mixture composition. The sand + Nd_2O_3 and sand + powdered sugar mixtures have greater accuracies for lower weight percents than higher weight percents. This indicates that the spectral feature differences between sand and the reagent material are distinctly different but at higher weight percents may be dominated by the reagent (Nd_2O_3 or powdered sugar) spectra, hindering separability. Some of these effects can be identified by visual inspection of the Nd_2O_3 spectra in Figure 4 from 1600-1800 nm and in the powdered sugar spectra in Figure 4 from 2000-2400 nm. The aspartame mixture results are different in that the aspartame spectral features aren't as obviously visible compared to its soil background. Subtle absorption features can be seen at ~2100 nm and ~2250 nm and are much less pronounced than the ones in the other two mixtures tested. It is highly likely this is causing the detection limit to be at a higher weight percent compared to the other mixtures.


Confusion Matrix
Sand
0.5% Nd2O3
1% Nd2O3
2% Nd2O3
3% Nd2O3
4% Nd2O3
5% Nd2O3
Junk
Actual
Predicted
accuracy=0.9971; misclass=0.0029


Confusion Matrix
Sand
0.1% Powd. Sug.
0.25% Powd. Sug.
0.5% Powd. Sug.
1% Powd. Sug.
2% Powd. Sug.
3% Powd. Sug.
4% Powd. Sug.
5% Powd. Sug.
Junk
Actual
Predicted
accuracy=0.9862; misclass=0.0138

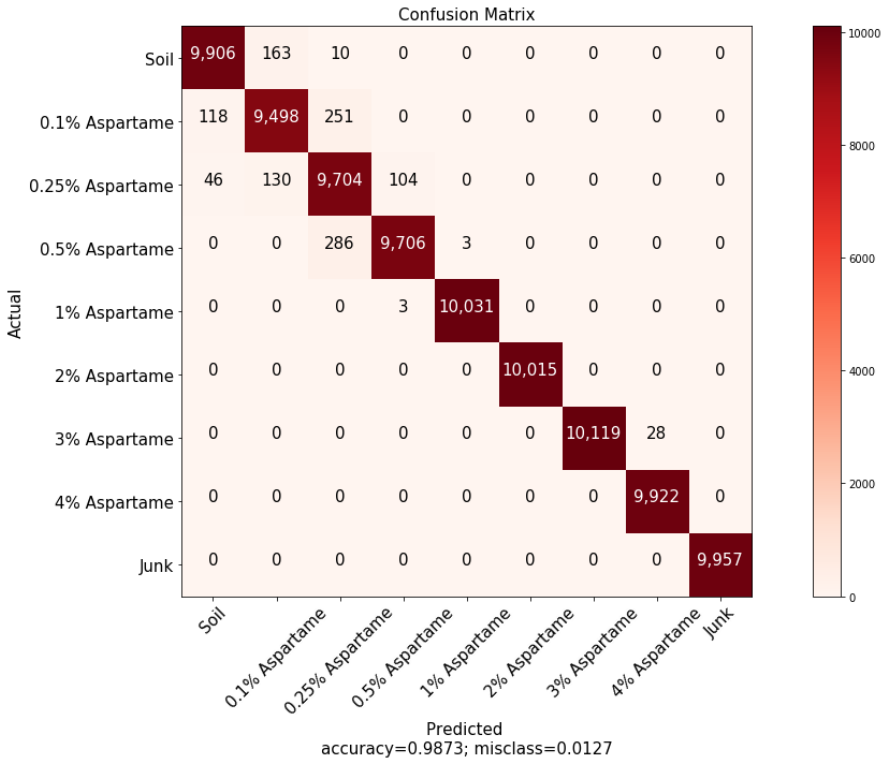


Figure 18. DL/CNN Confusion matrices of the validation result for the optimized models for each mixture. From top to bottom: (a) Nd_2O_3 + sand, (b) powdered sugar + sand, (c) aspartame + soil.

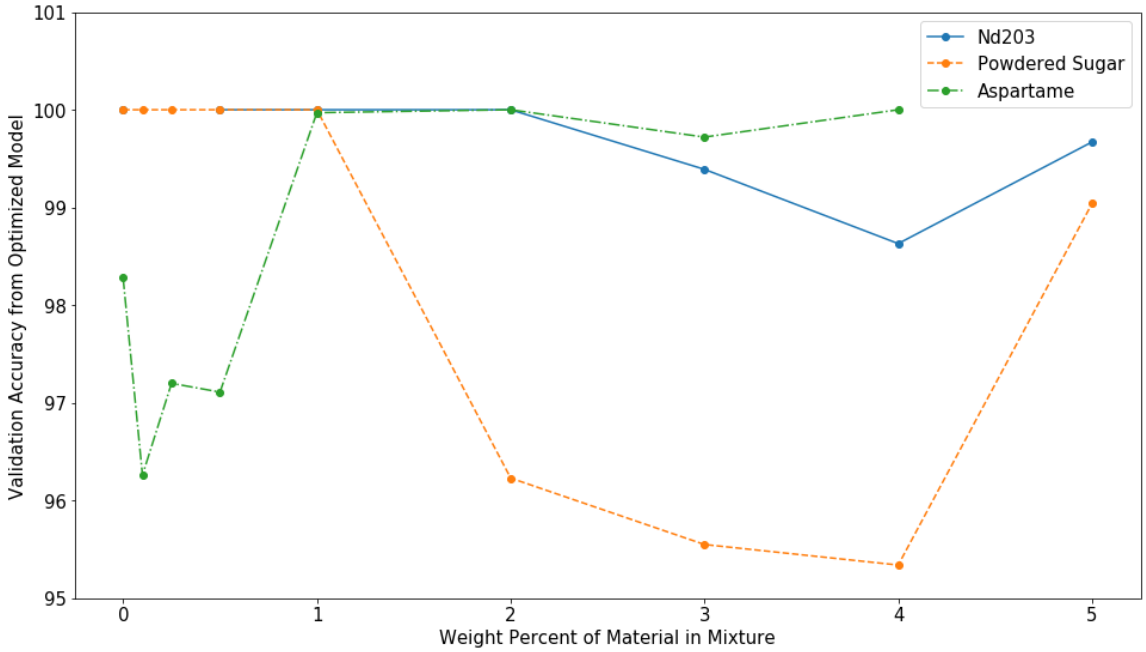


Figure 19. DL/CNN validation accuracy of each weight percent of each mixture.

DL/CNN 50-pixel Average Spectra vs. Single-Pixel Spectra

Two datasets were generated from the Nd_2O_3 + sand microscenes: one with ROIs of 50-pixel spectral averages and the other with single-pixel spectra. The overall validation accuracy is over 8% better for the 50-pixel average data (Table 4); however, the accuracies at low abundances are comparable. Figure 20 shows the drop-off in accuracy between two and three weight percent for the single pixel spectra, and subtly for the 50-pixel average spectra. This is could be due to a few factors, that should be investigated further. One possibility is that the higher weight percents mean spectra are being dominated by the Nd_2O_3 reagent spectra, therefore hindering separability of the higher wt. % classes. More specifically the sand spectra, seen in Figure 4, show very few changes in the slope of the signature – it is relatively linear compared to the Nd_2O_3. The

DL/CNN may not be able to distinguish that there is this relatively flat, featureless range within the signature of the Nd_2O_3. Another potential for the drop-off in accuracy at higher weight percents is that the depth and width of a spectral absorption feature may be very similar at >2 weight percent. Since the Nd_2O_3 spectral features may dominate at the higher weight percents, the neural network might not be able to separate them simply due to their similarity. It is curious though, how a neural net may not be able to discriminate these mixtures when band depth systematically increases with increase in reagent weight percent. More microscenes and continued research is needed to be able to add confidence to these conclusions and, ideally, identify any weight percent, from 0 to 100 %.

Grain size effect on accuracy was also considered. The variability of the grains in the mixture at single pixel scales might hinder the uniformity of the spectra at each weight percent. The effective GSD in the imagery is ~0.3 mm and the largest sand grain size is ~0.5 mm. However, the sample size of data (50,000 samples) and the mixing method (thoroughly mechanically mixed) lends confidence that we are indeed detecting spectral features generalized in each weight percent. If variability in the mixture was a major contributing factor, it would be present in the accuracy result of each weight percent.

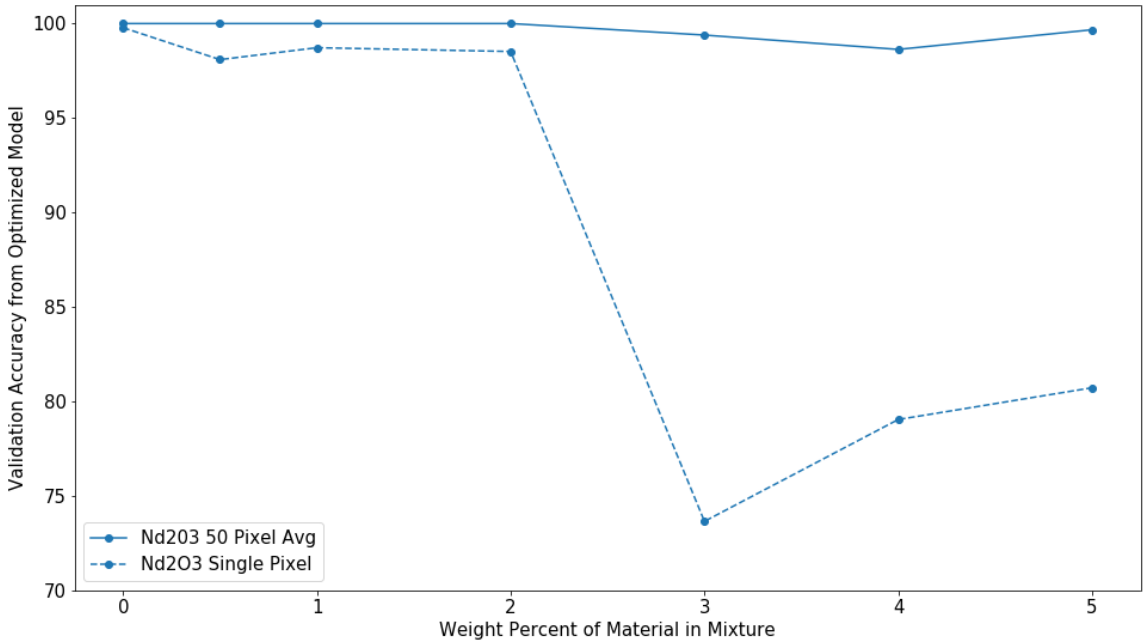


Figure 20. DL/CNN comparison of Nd_2O_3 validation accuracy for 50-pixel average spectra vs. single pixel spectra at each weight percent.
High accuracy (>98 %) for the lowest weight percents of single pixel spectra is a promising result.

DL/CNN vs. Shallow Network

A shallow neural network was implemented in Python using Keras with a TensorFlow-GPU backend and applied to the sand + Nd_2O_3 microscene spectra (both single-pixel and 50-pixel average spectra). This shallow network was branched from the same code used for the DL/CNN. The only change was that the 3 layers of CNNs with pooling were replaced by a single "Dense" layer. The TensorFlow shallow network showed decreased accuracy for both the single-pixel and 50-pixel average spectra compared to the DL/CNN. Overall validation accuracies are shown in Table 4. The 50-pixel average spectra result is very promising at 97.75 %. It is only ~2 % worse than the DL/CNN and is likely good enough to satisfy the need for detection of low abundance

Nd_2O_3. The single-pixel shallow network result is ~5% lower than the single-pixel DL/CNN, indicating it is not a data type and network type pair worth implementing.

The DL/CNN and the shallow networks both show the same trend of very good accuracies at the lower weight percent mixtures but lower accuracies at the higher weight percent mixtures. This may indicate that the difficulty with separating the higher weight percent mixtures is independent of the networks and is more likely caused by mixing phenomenologies or the microscene data itself.

Table 4. Overall Neural Network Results

Mixture	Pixel Averaging	Software	Processing Type	Overall Validation Accuracy	Lowest Weight Percent Accuracy	Mean F-1 Score
Nd_2O_3 + Sand	50 Pixels	TensorFlow	DL/CNN	99.71 %	100.0%	0.9971
Nd_2O_3 + Sand	50 Pixels	TensorFlow	Shallow	97.75 %	100.0%	0.9773
Nd_2O_3 + Sand	50 Pixels	MATLAB	Shallow	100.0%	100.0%	1.000
Nd_2O_3 + Sand	1 Pixel	TensorFlow	DL/CNN	91.09 %	98.05%	0.9112
Nd_2O_3 + Sand	1 Pixel	TensorFlow	Shallow	85.99 %	93.66%	0.8606
Nd_2O_3 + Sand	1 Pixel	MATLAB	Shallow	91.40%	99.06%	0.9150
Sand + Powdered Sugar	50 Pixels	TensorFlow	DL/CNN	98.62 %	100.0%	0.9862
Soil + Aspartame	50 Pixels	TensorFlow	DL/CNN	98.73 %	96.11%	0.9857

Implementation for HSI Analysis

Using the microscene data collected, synthetic images were generated in order to test how the algorithm would work when implemented for a user. Figure 21 shows an example of a 100x100 pixel x 164 band hyperspectral image created. The image had 9984 pixels of sand (0 weight percent Nd_2O_3), 5 pixels of each 0.5%, 1%, and 5% Nd_2O_3, as well has a vegetation spectrum in 1 pixel. The implemented prediction results when the DL/CNN was applied were 100% correct, as shown in Figure 22.

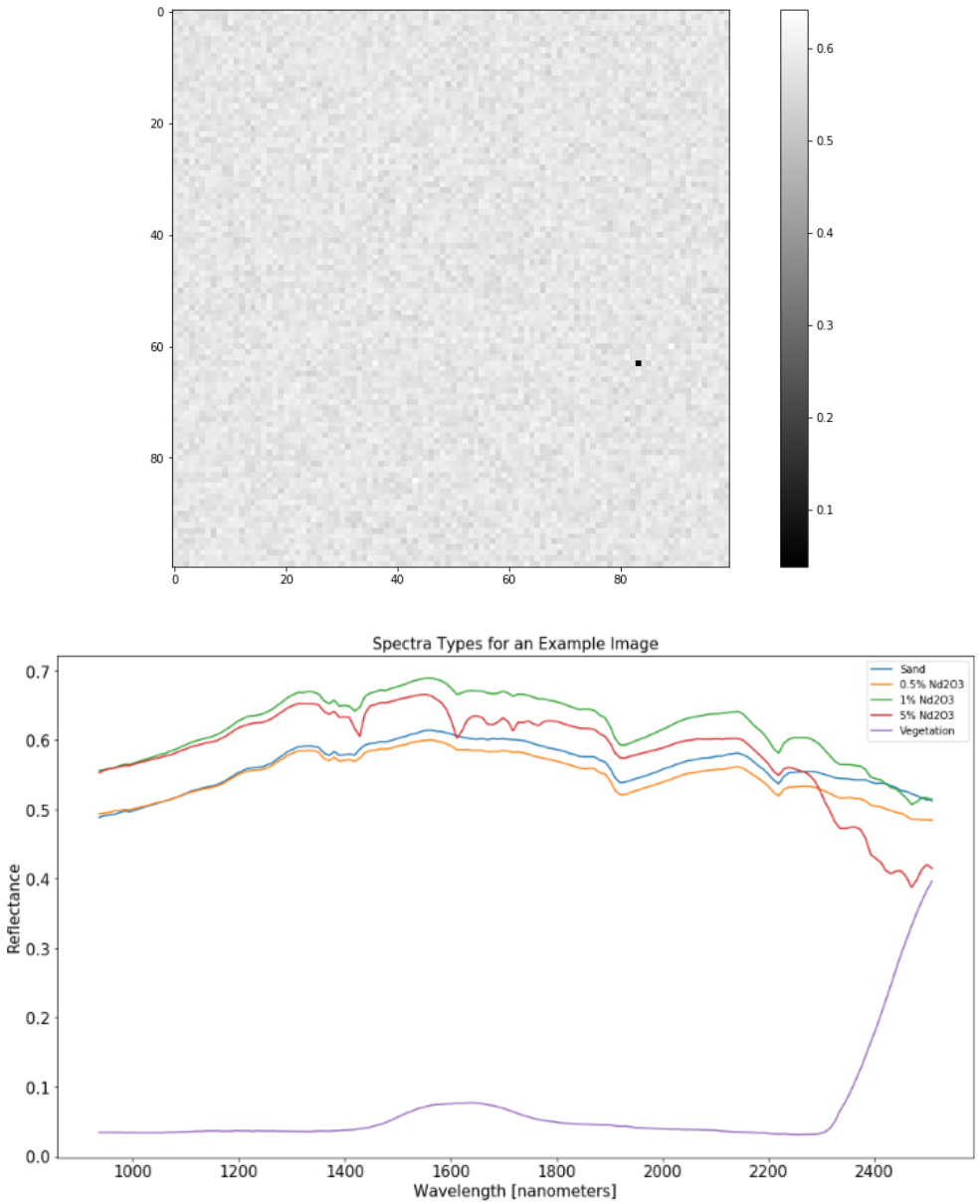


Figure 21. The synthetic image and spectra implanted in the image derived from microscene images. Note the lack of visual identification of the Nd_2O_3 mixed pixels and the obvious vegetation pixel (dark pixel in the middle-right area).

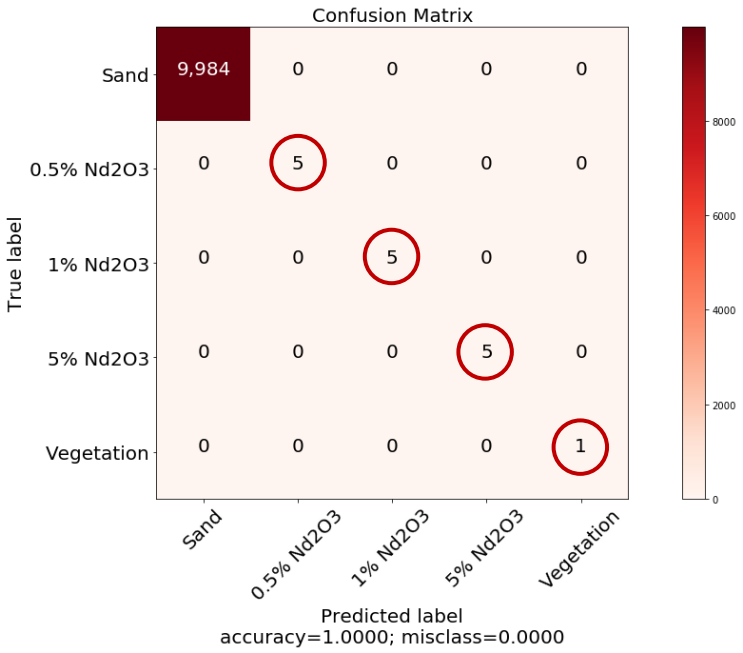


Figure 22. Results from the DL/CNN model prediction on a synthetic image from microscene data.

The implementation of these methods will also benefit from DL/CNN models that contain multiple mixtures such that iterative application of networks on the same image data aren't necessary. To test this, the three mixtures in the paper were concatenated and trained with the parameters of the optimized sand + Nd_2O_3 network. The overall validation accuracy of this concatenated DL/CNN was 97.66%; a promising result for future development of a model containing many spectra and abundances.

DISCUSSION AND CONCLUSION

The CNN results have overall validation accuracies of >98 % for each of the three mixtures studied using 50-pixel average spectra. The results demonstrate the potential of CNNs for detailed HSI analysis of low abundance intimate mixtures – especially if trained with ample data such as that generated by microscenes. A CNN in combination with results from traditional methods may form a stronger convergence-of-evidence case for the presence of a material at a low abundance. Particularly for the very low weight percent classes, this CNN application is a very promising result. Traditional methods of spectral analysis are effective at separating larger abundances including algorithms such as the matched filter (Figure 15), the adaptive cosine estimator, various approaches to nonlinear spectral unmixing, the United States Geological Survey's Tetracorder and PRISM [12, 23, 30]. However, the research shows that continued investigation is warranted for the use of trained neural networks to achieve lower MDQs and to deal effectively with nonlinear mixtures.

Several methods of unmixing and categorizing the non-linear mixtures, in addition to the CNNs, were tested in this study. Overall, the ENVI methods applied confidently detected the higher abundance mixtures (Figures 13 and 15). Additionally, the SVM was also able to detect the lowest abundance mixtures with high confidence (Figure 14). The MF and SMA methods yielded values for the 0.5 and 1.0 weight percent

mixtures that were 0 or close to 0, which are values consistent with background spectra rather than detection of the material.

Shallow, non-convolutional neural networks also showed good results for characterizing the low abundance mixtures. Using 50-pixel average spectra, the shallow networks were 100% accurate in detecting the lowest abundance Nd_2O_3 mixtures (Table 4). All neural networks tested (MATLAB shallow NN, TensorFlow DL/CNN, TensorFlow shallow NN) were able to separate low abundance classes based on the training data provided. This is an important addition to an analyst's toolbox for this type of problem where other traditional methods like the matched filter and linear unmixing did not produce adequate detections at low abundances.

The overall accuracy difference between the DL/CNN and the shallow neural networks was small (~2%). Both types of network did very well with low abundance mixtures as shown in Table 4. This outcome is an indication that less complex networks may still suffice for low abundance detection. The shallow network results are encouraging due to the simplicity and ease of entry into using tools like the MATLAB GUI for training a neural network. In some cases, the DL/CNN did end up with better overall accuracies than the shallow network (accuracies of Nd_2O_3 + Sand using both 50-pixel and single pixel data types), which indicates the convolutional transformation of the data is a better choice when creating a network for unmixing. The trade space for creating NNs and layering functions is essentially infinite so, the best functional design, whether it uses convolutions or not, is still to be created.

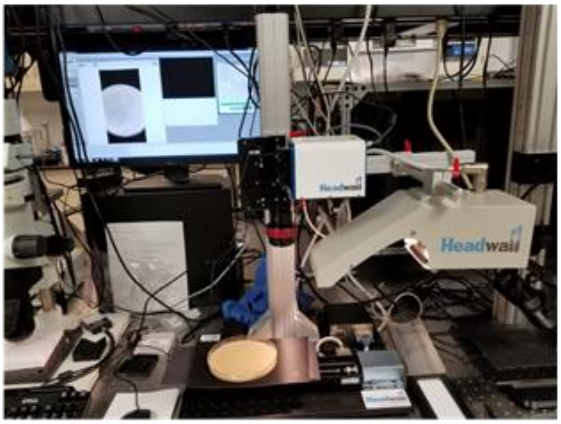
Applying the DL/CNN to multiple mixtures also revealed factors to consider in future research. A single set of DL/CNN parameters did not suffice to categorize each mixture with the same accuracy so, in order to get the highest accuracy results, parameters of the convolutions (number of filters, kernel size) were optimized for each mixture. However, data from multiple mixtures could be concatenated and trained with a single set of parameters resulting in slightly lower overall model accuracy (97.66 % vs. >98% for the 3 networks independently.) Different spectral features in the unique mixtures impact the DL/CNN's ability to separate classes with certain sized kernels and therefore the kernels need to be optimized. Ideally, an automated optimization feature for DL/CNN structure definition would be used, such as Google's AutoML approach, rather than a trial and error type method that was used in this research [17].

The high accuracy results of this study encourage future use of microscenes as a data source for training neural networks. Additional microscenes will need to be built for specific background/target-material combinations but such an activity is fast and low-cost (once a suitable laboratory sensor has been acquired). Also, spectral library construction should, in the future, be based on data from a microscene—from laboratory (and field-portable) imaging spectrometers. The spectral remote sensing community should augment the current single-pixel measurement of complex materials and mixtures with probability distributions of reflectance values on a band-by-band basis to capture, understand, and somehow utilize (e.g., with CNNs) the variability that real materials in real remote sensing scenes present.

Future work in this research includes: to explore the performance impact of different parameter settings and architectures of the DL/CNN (and of the shallow NN and the SVM), conduct a systematic study of SNR of the HSI data on performance (notable difference between 50-pixel average and single pixel results), develop confidence metrics for CNN-based material identification, build microscenes with other chemicals—to include ternary mixtures, and convert a spectrum to a two-dimensional data structure with parameterizations such as a wavelet transform and use "conventional"/computer vision 2D or 3D CNNs.

APPENDIX I

- Wavelength Range: **900 nm to 2500 nm**
- Aperture: **f/2.0**
- Entrance Slit Width: **25 µm**
- Dispersion/Pixel (nm/pixel): **6**
- FWHM Slit Image: **6.3 nm**
- Slit Length: **12 mm**
- Spectral Resolution: **12 nm**
- Spectral Bands: **168** (up to: 267)
- Spatial pixels: **384**
- Smile - Aberration-corrected: **Yes**
- Keystone - Aberration-corrected: **Yes**
- Detector: **Stirling-cooled MCT (HgCdTe)**
- Max. Frame Rate: **450 Hz**
- Pixel Pitch: **24 µm**
- Read A/D: **16-bit**
- Camera Control Interface: **Base Camera Link and RS232**
- Weight: **9.6 lbs.** (4.4 kg)
- Max. Power: **14.4 W**



The Headwall Photonics, Inc., Micro-Hyperspec® SWIR sensor parameters (left) and a digital camera photograph of the sensor in the laboratory (right). A microscene is shown on the translation stage. Note also the light source.

APPENDIX II

The Keras code for the DL/CNN:

```
model = Sequential()
conv1 = Conv1D(filters=conv1_nFilters, kernel_size=conv1_kernel_size,
input_shape=(164, 1))
model.add(conv1)
convout1 = Activation(activationFunc)
model.add(convout1)
model.add(MaxPooling1D(2))
model.add(Conv1D(filters=conv2_nFilters, kernel_size=conv2_kernel_size,
input_shape=(164, 1)))
convout2 = Activation(activationFunc)
model.add(convout2)
model.add(MaxPooling1D(2))
model.add(Dropout(droputRate))
model.add(Conv1D(filters=conv3_nFilters, kernel_size=conv3_kernel_size,
input_shape=(164, 1)))
convout3 = Activation(activationFunc)
model.add(convout3)
model.add(MaxPooling1D(2))
model.add(Flatten())
model.add(Dense(nclasses, activation='sigmoid'))
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
```

REFERENCES

[1] Ahmed, A.M., Duran, O., Zweiri, Y., Smith, M., (2017). Hybrid spectral unmixing: using artificial neural networks for linear/non-linear switching. Remote Sensing, v. 9, no. 775, 22 p., doi:10.3390/rs9080775.

[2] Allen, D.W., Resmini, R.G., Deloye, C.J., and Stevens, J.R., (2013). A microscene approach to the evaluation of hyperspectral system level performance. In Proceedings of the SPIE, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX, S.S. Shen and P.E. Lewis, eds., v. 8743, doi: http://dx.doi.org/10.1117/12.2015834, Baltimore, MD, 29 April -3 May 2013, 13 p.

[3] Baraldi, A., E. Binaghi, P. Blonda, P.A. Brivio, A. Rampini, (2001). Comparison of the multilayer perceptron with neuro-fuzzy techniques in the estimation of cover class mixture in remotely sensed data, IEEE Transactions on Geoscience and Remote Sensing 39 994–1005.

[4] Basener, W.F., and Flynn, M., (2018). Microscene evaluation using the Bhattacharyya distance. In: Proceedings of the SPIE, v. 10780, Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII, 107800R (23 October 2018); 13 p., doi:10.1117/12.2327004

[5] Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot, J., (2012). Hyperspectral unmixing overview: Geometrical, statistical,

and sparse regression-based approaches, IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing, vol. 5, no. 2, pp. 354–379.

[6] Bioucas-Dias, J., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N.M., Chanussot, J., (2013). Hyperspectral Remote Sensing Data Analysis and Future Challenges. Geoscience and Remote Sensing Magazine, IEEE. 1. 6-36. 10.1109/MGRS.2013.2244672.

[7] Bannari, A.; El-Battay, A.; Bannari, R.; Rhinane, H. Sentinel-MSI VNIR and SWIR Bands Sensitivity Analysis for Soil Salinity Discrimination in an Arid Landscape. Remote Sens. 2018, 10, 855.

[8] Camps-Valls, G. and Bruzzone, L., Kernel-based methods for hyperspectral image classification, IEEE Trans. Geosci. Remote Sens. 43(6), 1351–1362 (2005).

[9] Carpenter, G.A., Gopal, S., Macomber, S., Martens, S., and Woodcock, C.E., (1999). A neural network method for mixture estimation for vegetation mapping. Remote Sensing of Environment, v. 70, pp. 138–152.

[10] Chang, C., and Lin, C., (2011). LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1-27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[11] Chen, Y., Jiang, H., Li, C., Jia, X., and Ghamisi, P., (2016). Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, VOL. 54, NO. 10

[12] Clark, R.N., Swayze, G.A., Livo, K.E., Kokaly, R.F., Sutley, S.J., Dalton, J.B., McDougal, R.R., and Gent, C.A., (2003). Imaging spectroscopy: Earth and planetary

remote sensing with the USGS Tetracorder and expert systems. Journal of Geophysical Research, v. 108, no. E12, 5131, 44 p., doi:10.1029/2002JE001847.

[13] Dobigeon, N., Tourneret, J.Y., Richard, C., Bermudez, J., McLaughlin, S., and Hero, A.O., (2014). Nonlinear unmixing of hyperspectral images: Models and algorithms. IEEE Signal Processing Magazine, v. 31, pp. 82–94, doi:10.1109/MSP.2013.2279274.

[14] Hapke, B., (1993). Theory of Reflectance and Emittance Spectroscopy. Cambridge, UK: Cambridge Univ. Press.

[15] Harsanyi, J. C., and Chang, C. I., (1994.) Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach, IEEE Transactions on Geoscience and Remote Sensing, V. 32, pp. 779-785.

[16] Goodfellow, I., Bengio, Y., and Courville, A., (2016). Deep Learning. MIT Press, http://www.deeplearningbook.org

[17] Google. (2017). AutoML for large scale image classification and object detection. https://ai.googleblog.com/2017/11/automl-for-large-scale-image.html. (accessed on June 2019).

[18] Guilfoyle, K.J., Althouse, M.L., and Chang, C.-I., (2001). A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis neural networks. IEEE Transactions on Geoscience and Remote Sensing, v. 39, no. 8, pp. 2314–2318.

[19] He, K., Zhang, X., Ren, S., and Sun, J., (2016). Deep residual learning for image recognition. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR).

[20] Keras Documentation, https://keras.io/layers/convolutional/. Accessed January 2019

[21] Keshava, N., Kerekes, J., Manolakis, D., and Shaw, G., "Algorithm taxonomy for hyperspectral unmixing," Proc. SPIE 4049, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, (23 August 2000).

[22] Kokaly, R.F., Clark, R.N., Swayze, G.A., Livo, K.E., Hoefen, T.M., Pearson, N.C., Wise, R.A., Benzel, W.M., Lowers, H.A., Driscoll, R.L., and Klein, A.J., 2017, USGS Spectral Library Version 7: U.S. Geological Survey Data Series 1035, 61 p., https://doi.org/10.3133/ds1035.

[23] Kokaly, R.F., (2011). PRISM: Processing routines in IDL for spectroscopic measurements (installation manual and user's guide, version 1.0). U.S. Geological Survey Open-File Report 2011–1155, 432 p., available at https://pubs.usgs.gov/of/2011/1155/.

[24] Krizhevsky, A., Sutskever, I., and Hinton, G., (2012). Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems (NIPS).

[25] Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 6, May 2017, 84-90.

[26] Kumar, U., Raja, K.S., Mukhopadhyay, C., and Ramachandra, T.V., (2012). A neural network based hybrid mixture model to extract information from non-linear mixed pixels. Information, v. 3, pp. 420-441, doi:10.3390/info3030420.

[27] Kwon, H. and Nasrabadi, N.M., (2006). Kernel matched subspace detectors for hyperspectral target detection, IEEE Trans. Pattern Anal. Mach. Intell. 28(2), 178–194.

[28] Licciardi, G.A., and and Del Frate, F., (2011). Pixel unmixing in hyperspectral data by means of neural networks. IEEE Transactions on Geoscience and Remote Sensing, v. 49, no. 11, November, pp. 4163-4172, doi: 10.1109/TGRS.2011.2160950.

[29] J. Long, E. Shelhamer, and T. Darrell, (2015). Fully convolutional networks for semantic segmen-tation. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR).

[30] Manolakis, D., Marden, D., and Shaw, D. A., (2003). Hyperspectral image processing for automatic target detection applications. Lincoln Laboratory Journal, vol. 14, no. 1, pp. 79–116.

[31] Owen, A.J. (1995). Uses of Derivative Spectroscopy. Application Note, Agilent Technologies Innovating the HP Way. Available online: http://www.whoi.edu/cms/files/derivative_spectroscopy_59633940_175744.pdf (accessed on June 2019).

[32] Palsson, B., Sigurdsson, J., Sveinsson, J.R., and Ulfarsson, M.O., (2018). Hyperspectral unmixing using a neural network autoencoder. IEEE Access, v. 6, doi: 10.1109/ACCESS.2018.2818280, 11 p.

[33] Plaza, J., Plaza, A., Perez, R., and Martinez, P., (2009). On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images. Pattern Recognition, v. 42, pp. 3032-3045.

[34] Porter, W. M., and Enmark, H. E., (1987). System overview of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), in Proceedings, Society of Photo-Optical Instrumentation Engineers (SPIE), v. 834, p. 22-31

[35] Rand, R.S., Resmini, R.G., and Allen, D.W., (2017). Modeling linear and intimate mixtures of materials in hyperspectral imagery with single scattering albedo and kernel approaches. Journal of Applied Remote Sensing, v. 11, no. 1, pp. 016005-1 to 016005-28, doi: 10.1117/1.JRS.11.016005.

[36] Resmini, R.G., and Allen, D.W., (2014). The microscene: a novel method for hyperspectral image generation for algorithm development and testing. The 2014 Meeting of the Military Sensing Symposium (MSS) Specialty Groups on Passive Sensors, Battlefield Survivability & Discrimination, Detectors & Materials, Gaithersburg, MD, 8-12 Sept., 2014, 10 p.

[37] Simonyan, K. and Zisserman, A., (2015). Very deep convolutional networks for large-scale image recognition. IEEE International Conference on Learning Representation (ICLR).

[38] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research 15. 1929-1958 Submitted 11/13; Published 6/14

[39] Wilkinson, G.G., (1999). Open questions in neurocomputing for earth observation, in: Kanellopoulos, I., Wilkinson, G.G., Roli, F., Austin, J., (Eds.), Neurocomputation in remote sensing data analysis, Springer, Berlin, pp. 3–13.

[40] Zhu, X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., and Fraundorfer, F., (2017). Deep Learning in Remote Sensing: A Review. IEEE Geoscience and Remote Sensing Magazine, IN PRESS. October 12.

BIOGRAPHY

Kevin P. Christiansen graduated from Penfield High School in Penfield, New York, in 2008. He received his Bachelor of Science from the Rochester Institute of Technology in 2012 where he studied Imaging Science and Astronomy. He is an employee of The Boeing Company.