SAND2019-2883C

# Hyperspectral vegetation identification at a legacy underground nuclear explosion test site

Brian J. Redman[a], John D. van der Laan[a], Dylan Z. Anderson[a], Julia M. Craven[a], Elizabeth D. Miller[b], Adam D. Collins[b], Erika M. Swanson[b], and Emily S. Schultz-Fellenz[b]

[a]Sandia National Laboratories, 1515 Eubank Blvd SE, Albuquerque, NM 87123, United States
[b]Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM 87545, United States

## ABSTRACT

The detection, location, and identification of suspected underground nuclear explosions (UNEs) are global security priorities. Vegetation disturbances may provide complementary signatures that can confirm or build on the observables produced by prompt sensing techniques such as seismic or radionuclide monitoring networks. For instance, the emergence of non-native species in an area may be indicative of anthropogenic activity or changes in vegetation health may reflect changes in the water table resulting from an underground explosion. Previously, we have collected high spatial resolution (10 cm) hyperspectral data from an unmanned aerial system at a legacy underground nuclear explosion test site and its surrounds. These data consist of visible and near-infrared wavebands over 4.3 $km^2$ of high desert terrain along with high spatial resolution (2.5 cm) RGB context imagery. In this work, we employ various spectral detection and classification algorithms to identify and map vegetative species in an area of interest containing the legacy test site. A frequentist framework for fusing multiple spectral detections across various reference spectra captured at different times and sampled from multiple locations is employed. The spatial distribution of vegetation species is compared to the location of the underground nuclear explosion. We find a difference in species abundance within a 130 m radius of the center of the test site.

**Keywords:** hyperspectral imagery, unmanned aerial systems, remote sensing, support vector machine, adaptive cosine estimator

## 1. INTRODUCTION

The detection, location, and characterization of suspected underground nuclear explosions (UNEs) are global security priorities. Current techniques utilized by the international nuclear explosion monitoring community rely on sensors recording prompt data, such as seismic or radionuclide monitoring networks.[1] While these prompt techniques can generally detect and locate suspected underground nuclear tests, they cannot leverage additional event signatures manifested through non-prompt observables. These non-prompt observables are critical for refining event location and characterization. For example, post-event surface fractures can be important access points for in-situ verification techniques and cultural artifacts such as diagnostic and instrument cabling can refine event locations to specific areas of interest. In this work, we employ machine learning and statistical classification algorithms against previously collected hyperspectral imagery[2] to identify and map vegetative species in an area containing a legacy underground nuclear test site. We find that spatial species distributions clearly demarcate the underground nuclear tests in our area of interest.

Vegetation around a suspected UNE provides a unique carrier for a number of potential signatures. Short and long term changes to the subsurface and water table may manifest as changes in vegetative health, water content, or growth levels. The presence of non-native or pioneer stage species[3] can reveal areas of anthropogenic activity long after specific artifacts or thermal signatures have decayed or been sanitized. Vegetation can capture and hold on to surface and subsurface particulates for periods of days to decades. Additionally, as a surface observable connected to the subsurface, vegetation represents an appealing target for UNE signature exploitation as it can be measured, mapped, and exploited using remote sensing techniques. To this end, optical remote sensing

Further author information: (Send correspondence to B.J.R.)
B.J.R.: E-mail: bjredma@sandia.gov, Telephone: 1 505 284 6468

is particularly attractive as it is unobtrusive, can search large areas, and can resolve surface features such as vegetation in fine detail. Hyperspectral imaging (HSI) provides a powerful signature for characterizing vegetative signatures.

HSI is commonly used for remote material identification. HSI sensors collect optical intensity measurements for hundreds of contiguous bands over a spectral range commonly including the visible and near infrared (VNIR) wavebands. Illumination models are used to convert the intensity measurements of each pixel to reflectance profiles which are indicative of material type. In contrast to multispectral data, HSI bands are not optimized for a specific application or set of materials. This allows HSI to be used to identify a broader range of materials and applications. Additionally, HSI data allow for differentiation of materials that have similar spectral signatures such as healthy and diseased crops.[4] HSI has been used in the past for identification of surface mineralogy,[5] nondestructive food analysis,[6] medical applications,[7] and vegetation mapping.[8]

This paper overviews HSI data collection and signature mapping in Section 2, the methods we employ for vegetation mapping in Section 3, our results against legacy UNE test sites in Section 4, and concludes in Section 5.

## 2. BACKGROUND

Previously, we have collected high spatial resolution (10 cm ground sample distance) hyperspectral imagery from an unmanned aerial system (UAS) at a legacy underground nuclear explosion test site and its surrounds.[2] These data consist of VNIR wavebands over 4.3 $km^2$ of high desert terrain at the Nevada National Security Site (NNSS). The NNSS is a declared legacy nuclear explosion test range in southern Nevada, and is the site of over 800 underground nuclear tests in a variety of geologies (e.g. tuffs, alluvium, and granite) and emplacements (tunnels and boreholes).[9] Although the United States has placed a moratorium on UNE testing since 1992, the NNSS remains a laboratory for measuring and characterizing non-prompt and persistent signatures and observables. An overview of the HSI data collection is shown in Fig. 1.

The area of interest is composed of Pinyon-Juniper woodland, dominated by stands of evergreen junipers and pinyon pines and brush composed of rabbitbrush, sagebrush, and saltbrush. The scene is also punctuated by various species of wild flowers and grasses. Photos taken from on the ground highlighting the vegetation are shown in Fig. 2. Growth ring counting indicates that the climax community Pinyon-Juniper stands are in excess of 50 years of age. The UNE tests in the area of interest (AOI) were conducted from 1987-1991, meaning that much of the climax vegetation has been around since before the tests. Leading up to the tests, large areas around the boreholes were cleared of vegetation for equipment, vehicles, etc. This is reflected by primary and secondary stage succession vegetative communities that have not fully recovered to the climax stage even 30 years post test.

Optical remote sensing for vegetation mapping has broad applications in agriculture and ecology. Numerous airborne and space based multi and hyperspectral systems, such as Landsat, QuickBird, or AVIRIS,[8] have been applied to vegetative mapping applications. However, the long standoff distance for these systems results in course spatial resolution, ranging from meters to kilometers ground sample distance. Recently, unmanned aerial systems (UAS) have emerged as a viable platform for optical remote sensing. UAS can fly much lower than traditional systems, thereby providing very high spatial resolution measurements. The high spatial resolution improves the ability to spatially resolve targets, but it also reduces the need to use mixture models because there are fewer types of materials in each pixel. This high resolution along with the low cost of operation over small areas[10] makes UAS a prime platform for precision agriculture applications.[11]

Workflows for vegetation mapping from spectral remote sensing often make use of various vegetation indices.[12] These indices, typically differences or ratios between spectral bands, represent scores that are correlated with properties of the vegetation. The normalized difference vegetation index (NDVI) is the normalized difference between reflectance in the red and near infrared wavelengths[13] and is perhaps the most ubiquitous index in practice. NDVI is a well established method for estimating the vegetation ground cover from multi-spectral satellite data because chlorophyll is highly absorbing in the red waveband and the internal structures of leaves makes vegetation highly reflective in the near infrared waveband.[14,15] Given the high spatial resolution of our collected HSI, NDVI is a strong indicator of the presence of vegetation on a per-pixel basis.[11] While NDVI and other indices are proven tools for estimating properties such as biomass or ground coverage, they are severly lacking

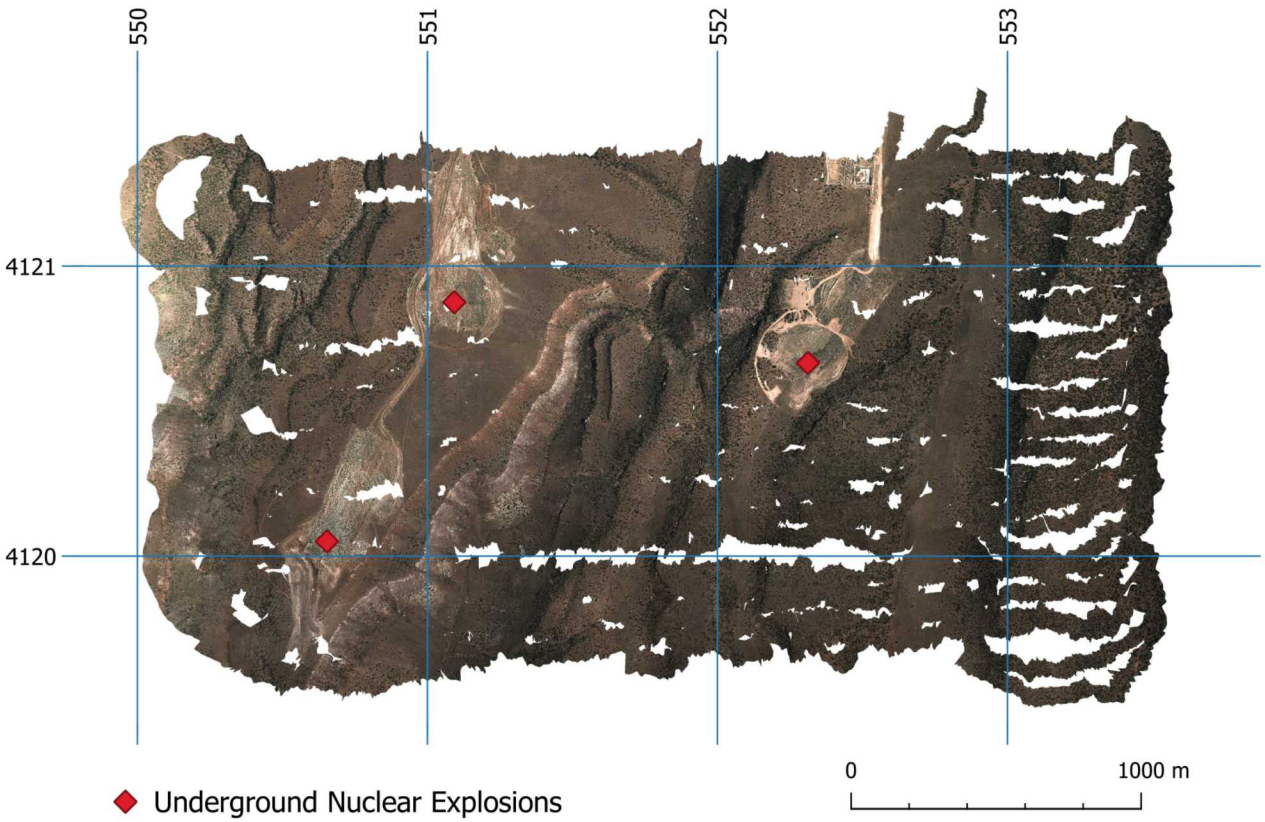


Figure 1: Pseudo-color overview of the VNIR HSI collected at the NNSS. The collection spans approximately 4.3 km$^2$. Grids represent 1 km UTM zone 11N coordinates.

in discriminating power for discerning different vegetative species. Unsupervised and supervised classification techniques, such as k-means or support vector machines (SVM), can be employed for the species classification problem. Unsupervised techniques use clustering to separate the pixels in the scene into a predetermined number of classes. Supervised techniques require training data for each class.

## 3. METHOD

We employ two competing methodologies for mapping vegetative species in our VNIR HSI scene: Adaptive Cosine Estimation (ACE) and support vector machines (SVM). Matched filtering approaches such as ACE are highly flexible and are common in hyperspectral data exploitation. While much less flexible, supervised classification approaches such as SVM bring to bear much greater discriminative power. In this section, we will overview both methodologies.

The ACE workflow makes use of ground-based reference spectra captured using a handheld ASD point spectrometer to perform detections within the HSI scene. Each pixel's reflectance in the scene is compared to the reference library to compute detections, and detections from the same species are aggregated into a single score. Pixels are assigned the species with the highest score. The workflow proceeds as follows. First, scene background statistics are computed. Calculating a covariance matrix for the entire HSI scene is intractable; rather, a region around the Barnwell UNE site was selected to estimate background statistics. To prevent vegetation contamination of the background estimate (which can degrade detector performance[16]), all pixels with an NDVI greater than 0.25 were excluded from the background statistics. Next the background estimate is used to demean and whiten (via principal component analysis rotation) the scene. The library of reference

Figure 2: Photos taken from on the ground within the AOI highlighting the diversity of topology and vegetation.

spectra have much higher spectral resolution than the HSI scene. Reference spectra are smoothed by convolving a Gaussian with full width-half max (FWHM) of 6 nm (the same FWHM of the imager used to collect the VNIR HSI) and downsampled to the imager resolution. The background demean and whitening transform is then applied to the reference spectra. Next the ACE score for each pixel and reference spectra combination are computed. The reference library consists of multiple measurements of each species. The ACE scores are grouped by species, and are fused into a single probability using the Probabilistic Fusion framework.[17] Final species determinations are made by selecting the highest probability species from the fused scores. This workflow is shown in Fig. 3.

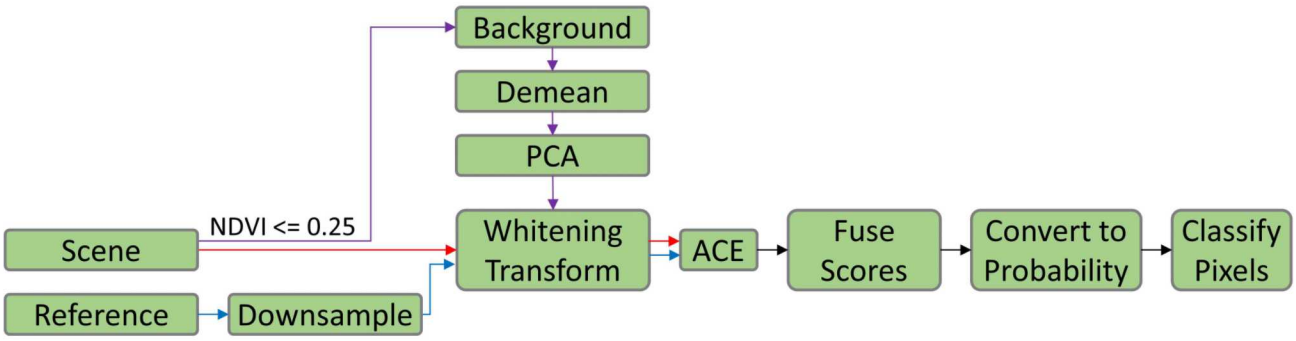


Figure 3: Overview of the detection algorithm workflow

Support vector machines (SVMs) are a type of machine learning model used to classify high dimensional data

such as HSI.[18,19] Our SVM workflow requires defining all possible classes and then providing in-scene examples of each class. The workflow proceeds as follows. First, possible classes are defined. Training an SVM to classify individual species is not feasible given the lack of labeled training data. For more tractable labeling, species are grouped into three different classes: two climax populations (old trees and old brush) that would exist on the site in absence of anthropogenic activity and a succession population (new growth) which is regrowing on the regions cleared of vegetation in support of the UNE tests. The old trees class consists of *Pinus monophylla*, commonly single-leaf pinyon, and *Juniperus osteosperma*, commonly Utah juniper. The old brush class consists of *Artemisia nova*, commonly black sagebrush, and unidentified scrub. The new growth class is dominated by brush that has grown in since the ground was cleared. This class contains *Chrysothamnus viscidiflorus*, commonly yellow rabbitbrush, *Salsola tragus*, commonly Russian thistle, and unidentified grasses. These class groupings preserve the capability to distinguish long standing climax populations (old trees and brush) for new succession and invasive growth (new growth) resulting from anthropogenic activities surround the UNE test. Next, training data is labeled as belonging to one of the three classes. For expedient per-pixel hand labeling, a region dominated predominantly by each class is selected and an NDVI threshold of 0.25 is applied to create a vegetation mask. Binary morphological opening and eroding are applied to each mask to remove features that are too large or small to belong to the predominant class for the region. Pixels remaining in each mask are labeled based on the region's predominant class and are combined to construct the labeled training dataset. In each training region, an area is selected to be a validation set. The pixels from the validation set are removed from the training set and are used as a test case for tuning hyperparameters. These labeled training regions are shown in Fig. 4.

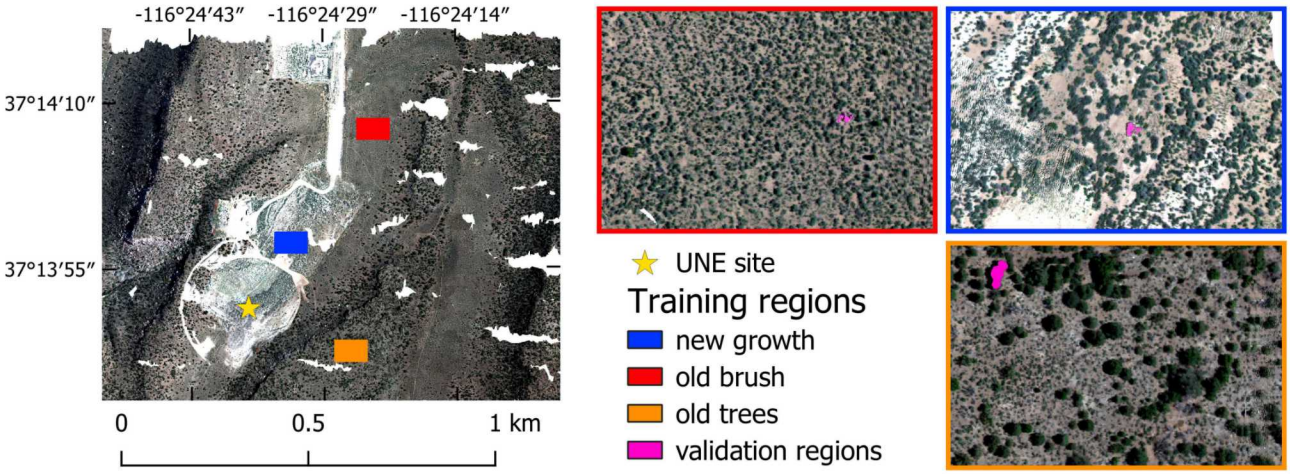


Figure 4: Regions used for selecting and labeling training data as input to the SVM workflow.

With the training data for each class assigned, the training data is next down-sampled and preprocessed for input to an SVM. The labeled training data contains over 50,000 pixels per class, far more than is necessary for training an SVM. To reduce computational cost, the training data is randomly down-sampled to 5,000 pixels per class. A standardizing transform (constructed from the training data) is applied to all spectra prior to SVM input. First each spectra is smoothed by convolving with a Gaussian function to reduce noise. The smoothed spectra are trimmed to the wavelength range of 440 nm to 900 nm to eliminate noisy bands in regions of low detector response or poor signal. Each spectral band of the data is standardized by subtracting off the mean and dividing by the variance of the training data in that band. Next a principal component analysis (PCA) transform is fit to the training data and is used to reduce input dimensionality.[20] The standardized training data is used to fit an SVM, and the entirety of the dataset is passed through the now-fitted classifier. After HSI classification, the spatial morphology in the scene is exploited with segmentation and voting to produce consistent classifications for multi-pixel individual plants. To segment the image, the HSI data is converted to red-green-blue (RGB) by selecting the bands corresponding to wavelengths 669nm, 540nm, and 458nm respectively. The pixels in the resulting RGB image corresponding to NDVI less than 0.25 are set to zero to select only the vegetation. Felzenszwalb segmentation is applied to the RGB vegetation, creating an over-segmented image.[21] Adjacent

segments without significant edges between them are hierarchically merged. Each segment is assumed to be a single species, therefore the entire segment is assigned to the class of which the majority of its pixels belong. Fig. 5 shows the workflow for creating an SVM to classify the vegetation around the UNE test sites.

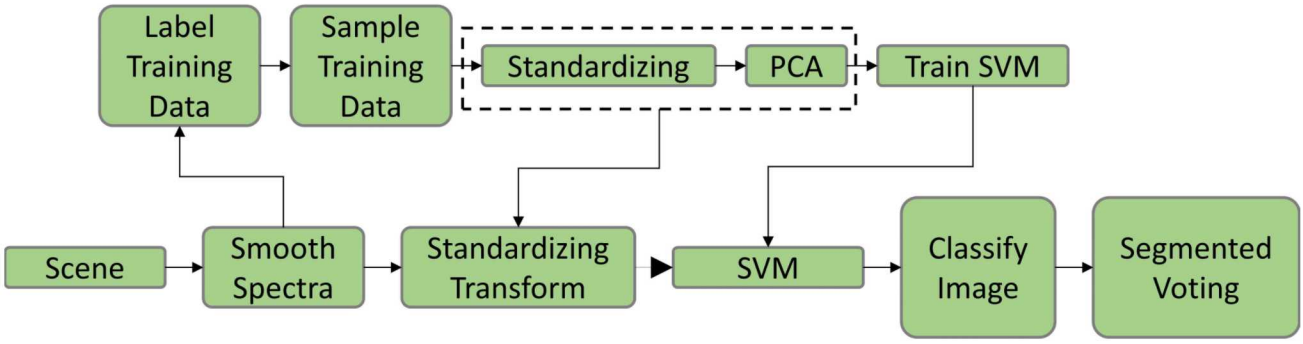


Figure 5: Overview of the SVM workflow

A hyperparameter sweep is used to determine the best SVM workflow parameters to classify the HSI data. The workflow hyperparameters are randomly varied and resulting performance is measured based on the true positive fraction of each class in the validation regions. The parameters swept over are FWHM of the smoothing Gaussian, number of components retained in the PCA transform, SVM kernel (linear, polynomial, or radial basis function), and SVM penalty. Additionally, each SVM kernel introduces a unique set of parameters that are swept over:

- Polynomial: kernel weighting $\gamma$, offset coefficient, and degree
- Radial Basis Function: kernel width $\gamma$

The best performance found uses polynomial kernel, a cost parameter of 10, a kernel coefficient of 0.1, a degree of 1, an offset of 0, a smoothing FWHM of 4 nm, and 30 PCA components. It should be noted that a polynomial kernel with the degree of 1 and the offset of 0 is equivalent to a linear kernel with the additional weighting introduced by $\gamma$. The linear kernel alone produces worse performance.

## 4. RESULTS

The ACE workflow is ineffective at classifying vegetative species. While the workflow strongly detects vegetation pixels (detection scores greater than 0.9), the detections for Pinyon are typically stronger than for all other species, even at the specific physical locations where reference spectra were measured. The classifications produced for Rabbitbrush and Pinyon are shown in Fig. 6. The majority of the Rabbitbrush classifications produced by the ACE workflow are in shaded areas where the illumination model to convert radiance to reflectance is suspect.

The high detection for all vegetation but poor discrimination performance between species indicates that the ACE workflow is emphasizing the strong red edge feature of vegetation spectra but missing the subtler between-species differences. We hypothesize that modification to data preprocessing could improve species differentiation. By disregarding all vegetation in estimating background statistics, the whitening transform strongly amplifies generic vegetation spectral features. Modifying the background removal process to remove the mean red edge profile and spectral bands that have minimal between species variation could decrease the emphasis of the red edge in the detection score. Additionally, more reference spectra samples for each species could improve the empirical distribution fits that are part of the probabilistic fusion process.

Classification results from the entirety of the HSI dataset passed through the SVM workflow are shown in Fig. 7. Larger vegetation is classified as trees, smaller vegetation in the uncleared areas is classified as old brush,

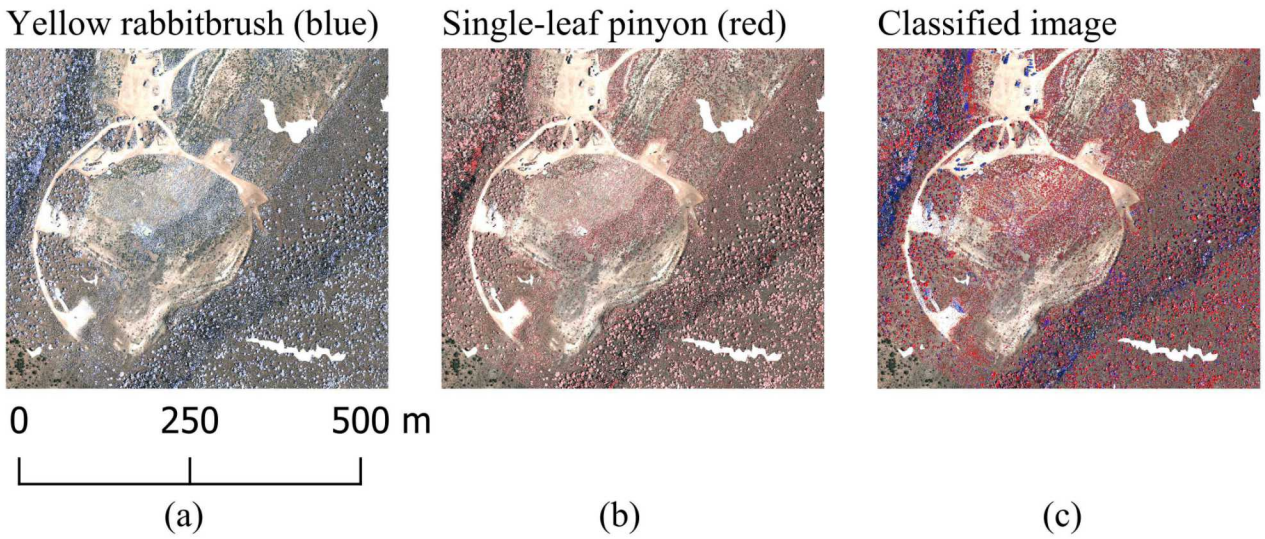


Figure 6: Detected species using the ACE workflow. Plotted are the detection scores for (a) Rabbitbrush (b) Pinyon and (c) comparison between Rabbitbrush and Pinyon highlighting the dominance of Pinyon detections throughout the scene.

and vegetation in the UNE test areas are classified as new growth. Some errant misclassifications exist, but there is strong spatial consistency throughout the scene in classifications and dominant vegetation in local regions are classified correctly. There are evident correlations and differences in vegetative type around the UNE test sites in the AOI, representing a strong remote-detectable signature.

While providing a powerful discriminant, classification approaches such as the SVM workflow have some shortcomings. They require sufficient in-scene labeled data with which to train models. Furthermore, they require specification of all possible classes up front: in this case old trees, old brush, and new growth. This is highlighted by the "old trees" along the western edge of the Barnwell test (see Fig. 8); in reality the trees are actually grass. Grass is not explicitly defined in any of the classes, so the SVM assigns the grass to the perceived most similar spectral class (in this case "old trees").

## 5. CONCLUSIONS

In this work we have demonstrated that there are remotely observable vegetation differences around UNE test site locations which can be detected and mapped using hyperspectral imaging. We showed two methodologies for classifying plant species with VNIR HSI. The ACE detection workflow was effective at detecting vegetation, but could not differentiate species type. It may be possible to improve this workflow via enhancements to the preprocessing stages. The SVM supervised machine learning workflow effectively differentiated new growth resulting from anthropogenic activity from old growth surrounding the UNE tests in the area of interest. Changes in vegetation speciation provide non-prompt, long duration observables decades after the tests. These non-prompt observables can be used alongside prompt techniques to improve and refine suspected UNE event locations and characterizations.

## 6. ACKNOWLEDGMENTS

The authors would like to thank the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development, for sponsoring this work. We would also like to thank the Underground Nuclear Explosion Signatures Experiment team, a multi-institutional and interdisciplinary group of scientists and engineers, for its technical contributions and support at the NNSS.

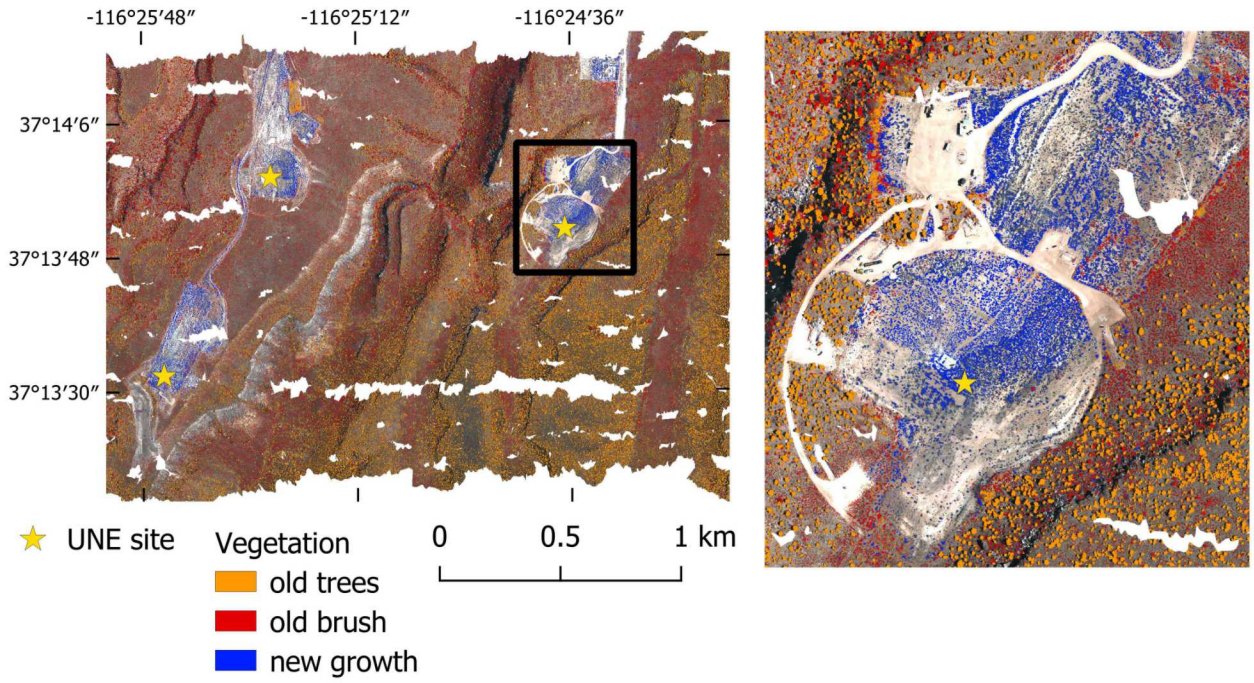


Figure 7: Vegetation classifications from the SVM workflow across the entirety of the HSI scene. New growth presents strong spatial correlation with UNE test sites in the scene.

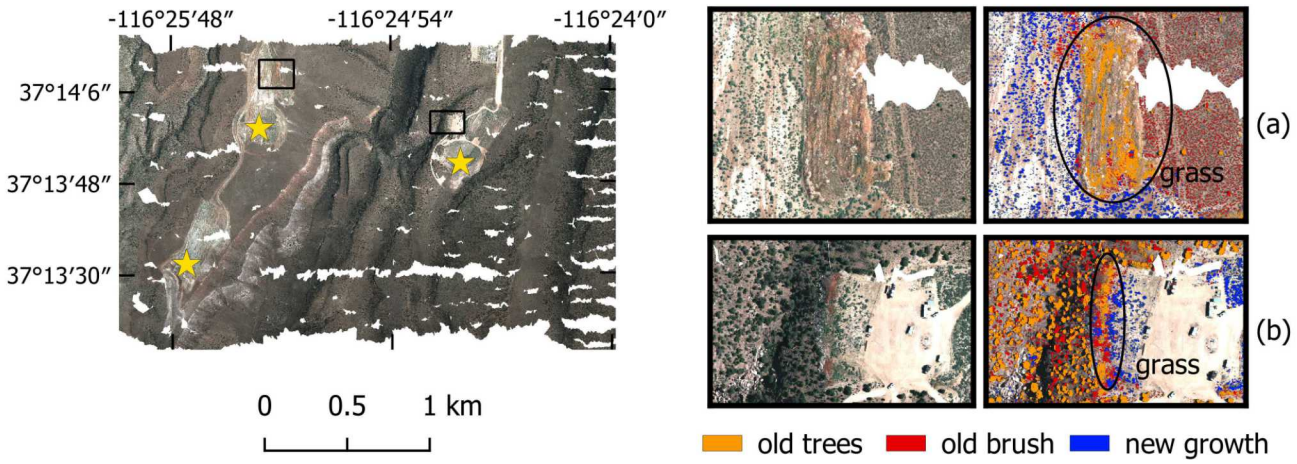


Figure 8: There were classification errors with vegetation not included in the training classes. Grass in (a) and (b) were classified as trees.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND No. 2019-TODO

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

## REFERENCES


[1] Hoffmann, W., Kebeasy, R., and Firbas, P., "Introduction to the verification regime of the comprehensive nuclear-test-ban treaty," *Physics of the Earth and Planetary Interiors* **113**(1-4), 5–9 (1999).

[2] Anderson, D., Craven, J. M., Dzur, R., Briggs, T., Lee, D. J., Miller, E., Schultz-Fellenz, E., and Vigil, S., "Using unmanned aerial systems to collect hyperspectral imagery and digital elevation models at a legacy underground nuclear explosion test site," (2018).

[3] Horn, H. S., "The ecology of secondary succession," *Annual review of ecology and systematics* **5**(1), 25–37 (1974).

[4] Mahlein, A.-K., Oerke, E.-C., Steiner, U., and Dehne, H.-W., "Recent advances in sensing plant diseases for precision crop protection," *European Journal of Plant Pathology* **133**, 197–209 (May 2012).

[5] van der Meer, F. D., van der Werff, H. M., van Ruitenbeek, F. J., Hecker, C. A., Bakker, W. H., Noomen, M. F., van der Meijde, M., Carranza, E. J. M., de Smeth, J. B., and Woldai, T., "Multi- and hyperspectral geologic remote sensing: A review," *International Journal of Applied Earth Observation and Geoinformation* **14**(1), 112 – 128 (2012).

[6] Gowen, A., O'Donnell, C., Cullen, P., Downey, G., and Frias, J., "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends in Food Science and Technology* **18**(12), 590 – 598 (2007).

[7] Guolan Lu, B. F., "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics* **19**, 19 – 19 – 24 (2014).

[8] Xie, Y., Sha, Z., and Yu, M., "Remote sensing imagery in vegetation mapping: a review," *Journal of plant ecology* **1**(1), 9–23 (2008).

[9] US Department of Energy, "United states nuclear tests, july 1945 through september 1992," (2000).

[10] Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R., and Gioli, B., "Intercomparison of uav, aircraft and satellite remote sensing platforms for precision viticulture," *Remote Sensing* **7**(3), 2971–2990 (2015).

[11] Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., and Gattelli, M., "Evaluating multispectral images and vegetation indices for precision farming applications from uav images," *Remote Sensing* **7**(4), 4026–4047 (2015).

[12] Bannari, A., Morin, D., Bonn, F., and Huete, A. R., "A review of vegetation indices," *Remote Sensing Reviews* **13**(1-2), 95–120 (1995).

[13] Jackson, R. D. and Huete, A. R., "Interpreting vegetation indices," *Preventive Veterinary Medicine* **11**(3), 185 – 200 (1991).

[14] Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L., "Overview of the radiometric and biophysical performance of the modis vegetation indices," *Remote Sensing of Environment* **83**(1), 195 – 213 (2002). The Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring.

[15] Knipling, E. B., "Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation," *Remote Sensing of Environment* **1**(3), 155 – 159 (1970).

[16] Manolakis, D., Lockwood, R., Cooley, T., and Jacobson, J., "Is there a best hyperspectral detection algorithm," (2009).

[17] Simonson, K., "Probabilistic fusion of atr results," tech. rep., Sandia National Laboratories (SNL-NM), Albuquerque, NM (1998).

[18] Mountrakis, G., Im, J., and Ogole, C., "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing* **66**(3), 247 – 259 (2011).

[19] Melgani, F. and Bruzzone, L., "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing* **42**, 1778–1790 (Aug 2004).

[20] Rodarmel, C. and Shan, J., "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Science* **62**(2), 115–122 (2002).

[21] Felzenszwalb, P. F. and Huttenlocher, D. P., "Efficient graph-based image segmentation," *International Journal of Computer Vision* **59**, 167–181 (Sep 2004).