

Article

Machine-learning for mineral identification and ore estimation from hyperspectral imagery in tin-tungsten deposits

Agustin Lobo^{1,*}, Emma Garcia², Gisela Barroso³, David Martí², Jose-Luis Fernandez-Turiel¹, Jordi Ibáñez-Insa¹

¹Geosciences Barcelona, GEO3BCN-CSIC, Spain

²Lithica (SCCL), Spain

³MSc. Program on Remote Sensing and GIS, Universitat Autònoma de Barcelona, Spain

*Correspondence: Agustin.Lobo@geo3bcn.csic.es

Abstract: This study aims to assess the feasibility of delineating and identifying mineral ores from hyperspectral images of tin-tungsten mine excavation faces using machine-learning classification. We compiled a set of hand samples of minerals of interest from a tin-tungsten mine and analyzed two types of hyperspectral images: 1) images acquired with a laboratory set-up under close-to-optimal conditions; and 2) scan of a simulated mine face using a field set-up, under conditions closer to those in the gallery. We have analyzed the following minerals: cassiterite (tin ore), wolframite (tungsten ore), chalcopryite, malachite, muscovite, and quartz. Classification (Linear Discriminant Analysis, Singular Vector Machines and Random Forest) of laboratory spectra had a very high overall accuracy rate (98%), slightly lower if the 450 – 950 nm and 950 – 1780 nm ranges are considered independently, and much lower (74.5%) for simulated conventional RGB imagery. Classification accuracy for the simulation was lower than in the laboratory but still high (85%), likely a consequence of the lower spatial resolution. All three classification methods performed similarly in this case, with Random Forest producing results of slightly higher accuracy. The user's accuracy for wolframite was 85%, but cassiterite was often confused with wolframite (user's accuracy: 70%). A lumped ore category achieved 94.9% user's accuracy. Our study confirms the suitability of hyperspectral imaging to record the spatial distribution of ore mineralization in progressing tungsten-tin mine faces.

Keywords: hyperspectral imaging; machine learning; spectral geology.

1. Introduction

Building and updating 3D models to quantitatively characterize the distribution and heterogeneity of ore grade, vein thickness, orientation, and network geometry are critical for the mining industry as it provides a better understanding of the mineralization with important benefits for the management and optimization of the exploitation, thus reducing environmental impact. Typically, these 3D models are based on knowledge of the site's structural geology, with geophysical exploration contributing to provide a better understanding of the deposit. Image data acquired from the mine excavation itself could be used to confirm or update these models, monitoring ore grade and distribution. In this regard, we aim at integrating proximal ground-based conventional and hyperspectral imaging of the mine excavation face to provide the ore grade and 2D geometry of the mineralization in the section. Ideally, a 3D tomography could be derived from these 2D planes as the excavation front progresses.

Hyperspectral images produced by imaging spectrometers are 3D arrays in which each voxel holds a radiance spectrum that is processed to reflectance [1] [2]. Acquisition of single reflectance spectra with spectrometers in the visible, near-infrared, and short-wave infrared wavelength domains (400 – 2600 nm) is a relatively simple and non-invasive technique that has been used in the laboratory and the field for decades ago [3] [4]. 1999 The particular optical and electronic properties of each material result, under

illumination results in specific spectral features that are often diagnostic of given minerals and rocks [5][6][7], although the influence of variations in illumination and viewing geometry, the macro-structure of the sample, and the presence of mixtures in the field-of-view, tend to reduce the diagnostic power of reflectance spectra. Extending point readings to a hyperspectral image has two main advantages. First, the 2D result allows considering the spatial distribution of different materials in the imaged scene. Second, in case of a crude identification, being able to delimit a given uncertain material facilitates an accurate sampling for complementary techniques such as x-ray diffraction (XRD) or x-ray fluorescence (XRF).

Most hyperspectral images have been remotely acquired from airborne sensors and a few satellites, and, in the context of the mining industry, common applications of remotely-sensed hyperspectral images are mineral exploration (notably of hydrothermal alteration, see examples in [8][9]) and environmental impact [10] [11]. Currently, close-range hyperspectral images of hand samples and/or drill cores [12] [13] [14] [15], along with ground-based panoramic hyperspectral imaging of semi-vertical outcrops [16] [17] [18] [19] [20], are increasingly used as hyperspectral imaging systems become more portable and widespread. Multi- and hyperspectral image systems have also been developed for ore microscopy [21] [22] [23] [24] with the aim of achieving quantitative mineralogical analysis. In their extensive review, Krupnik and Khan (2019) [20] organize very recent articles (most of them published after 2010) reporting close-range hyperspectral imaging applications, which mainly deal with the analysis of economically-valued materials (32 studies, of which 13 focus on validation of spectral identification methods), and environmental impact (13 studies). Horizontal and oblique ground-based hyperspectral imaging of vertical structures (cliffs, road cuts, open-pit walls...) can be integrated with digital topography, which can be retrieved from either terrestrial LiDAR data [25] or Structure from Motion (SfM) processing of conventional photographs [26] [27].

An extended panoply of methods have been developed for the analysis of remotely sensed hyperspectral images, (e.g. [28]; [29] [30]), which are also being applied to close-range hyperspectral images, and few multi-scale studies have been undergone ([13] [31]. Amigo (2020) [2] compiles a number of modern methods that can be applied to both remote and close-range hyperspectral imagery. Machine learning methods, in particular supervised classification methods, have become common for the analysis of hyperspectral images [32], [33], [34]), and are starting to be applied to close-range imagery for the automatic identification and mapping of different materials within the imaged sample. Murphy et al. (2012) [16] acquired both point spectra and ground-based hyperspectral images of vertical faces in an iron-ore mine from a distance of 30 m in the visible and near-infrared (VISNIR) and short-wave infrared (SWIR). The authors compared spectral angle mapping (SAM) and support vector machines (SVM) to classify materials such as shale, manganese-shale, goethite, martite, and chert, taking shadowing effects into account. While SVM outperformed SAM under uniform illumination conditions, SAM was more robust to the presence of shadows unless shadowed training spectra were provided to the SVM classifier. Also, SAM performed better than SVM if training spectra were selected from a spectral library, instead of from within the image itself. Krupnik and Khan (2019) [20] also presented their own case studies of close-range hyperspectral panoramic imaging for the characterization of semi-vertical outcrops (faults, pits, roadcuts and quarry settings). They applied SVM and Multi-range Spectral Feature Fitting (SFF) to produce classified images of materials such as limestone, siltstone, shale, hematitic siltstone, calcite, illite, jarosite, iron oxides, and dolomite. They also obtained considerable geological information from mapping the wavelength and depth of absorption features in selected wavelength ranges, which they applied as well later [35] to analyze hyperspectral images of small, laboratory prepared mixtures of limestone minerals (calcite, dolomite, and chert), with results that were consistent to those provided by point count. The same study reports a successful mapping of mixtures in hyperspectral images of small rock chips by applying machine learning methods and using the spectra of known mixtures as references. López-Benito et al. (2020) [24] applied classification methods for the analysis of microscope hyperspectral

images in the VISNIR (350 – 1000 nm) range for the automated identification of metallic ore species, which was achieved with very high accuracy.

Deployment and operation of hyperspectral imaging systems in the field is always involved and doing so under gallery mine conditions is challenging. As a previous step to actually acquire, process and analyze hyperspectral images of a mine excavation face, we explore here the feasibility and interest of such an approach by using hand samples that feature the minerals of interest. For this purpose, we conducted the following studies:

1. Laboratory imaging spectrometry. We scanned the hand samples using hyperspectral cameras on a laboratory set-up to assess spectral separability and evaluate machine-learning classification methods under close-to-optimal conditions.
2. Simulation of hyperspectral imaging of the mine face. We scanned the whole set of hand samples with the same field set-up, illumination, and distance to object that are expected to be used in the mine gallery, evaluating machine-learning methods to identify and map the distribution of materials in the resulting image.

2. Geological settings

The studied samples were collected in the San Finx tin-tungsten mine, which is located in Lousame, A Coruña, Galicia, NW Spain. San Finx is a typical case of Late Palaeozoic, granite-related hydrothermal deposit, associated to the European metallogenic belt. Tin-tungsten ore is directly related to granites formed from Late Devonian to Permian during the Variscan or the so-called Hercynian orogeny, resulting from the continental collision of Laurussia and Gondwana [36]. The San Finx mineralization consists of a sub-vertical set of quartz lodes prevalent in the N 50° E direction, which at its eastern end shows pegmatitic features (e.g., the occurrence of K-feldspar). This lode field reaches more than 1 km in length, and the lode thickness varies between 0.5 and 1.5 m [37]. Host rocks consist of schists, migmatites and granites, while the metallogenesis consists of Sn-W-Ta-Nb-Mo-Cu-As-Au-Ag-Bi. The exploited ore minerals are cassiterite and wolframite. Other minerals occurring in the quartz lodes are arsenopyrite, pyrite, scheelite, chalcopyrite, molybdenite columbite-tantalite, muscovite and tourmaline, among others [38,39]. The San Finx deposit is declared as a Site of Geological Interest (<http://info.igme.es/ielig/LIG-Info.aspx?codigo=GM044>).

3. Materials and Methods

3.1 Laboratory imaging spectrometry

In order to achieve close-to-optimal conditions, we scanned samples from the mine with *Specim FX10* and *FX17* cameras (Table 1) consecutively mounted on the same *Specim LabScanner Setup 40 x 20* and with care not to move the samples between each scan. We used *Specim's Lumo* software for scanning, which renders hyperspectral images of radiance for the whole scanned area, and two images of, respectively, the dark (internal shutter) and white references, which we used to calculate images of spectral reflectance for each camera.

Table 1. Specifications of images acquired with Specim cameras FX10 and FX17. FWHM: full width height maximum; FOV: field of view; SNR: signal to noise ratio.

	FX10	FX17	FX10_FX17
Spectral range (nm)	397 – 1004	936 – 1720	397 – 1720
Spectral bands	448	224	628
Spectral FWHM (nm)	1.34 – 1.41	3.46 – 3.48	1.34 – 3.48
pixels/line	1024	640	1024
FOV (°)	38°	38°	
SNR	600:1	1000:1	

We co-registered the FX17 reflectance images to their FX10 counterparts based on bands corresponding to the common part of the wavelength range and stacked the resulting FX10 and FX17 co-registered reflectance images into one single reflectance image of 628 bands for each sample and spatial resolution of 0.224 mm/pixel. In order to perform reflectance calculation, co-registration, band stacking, and cropping, we developed in-house software consisting of R [40] scripts (packages raster [41], rgdal [42], RStoolbox [43], gdalUtilities [44] and link2GI [45], which call commands of GDAL [46], OTB [47] and ptools [48]. The co-registration processing proceeded in two steps. First, we run a first-order polynomial warp with homologous points automatically extracted by the SIFT method (HomologousPointExtraction in OTB; [49]). Second, the co-registration was refined by calculating subtle local X and Y shifts that optimized local correlation between the corresponding bands of both cameras (FineRegistration in OTB). Geometric transformations were calculated using corresponding bands of the same wavelength in both cameras and then applied to all bands of FX17 (gridBasedImageResampling in OTB).

Hyperspectral images were displayed and sets of pixel spectra for each mineral were interactively extracted with the EnMap Box plugin [50] of QGIS [51], which required combining photo-interpretation and direct visual inspection of the samples. Finally, we imported the spectra into R, where we made comparative graphics along with reference spectra extracted from spectral libraries by USGS [52] and JPL [53] [54] and run statistical analysis. We calculated matrices of dissimilarities using the Jeffries-Matutsita metric [55] to measure the spectral resolving power of each camera (FX10, FX17, and combined FX10_FX17). We also simulated the RGB values of a conventional camera (Canon 60D) from the FX10 values using camera spectral sensitivity curves [56] and calculated the RGB dissimilarity matrix with these values as well to evaluate the power of conventional photos to identify the minerals of interest. Finally, we run three different classification methods: Linear Discriminant Analysis (LDA), Singular Vector Machines (SVM), and Random Forest (RF) for each of the four datasets (FX10, FX17, combined FX10_FX17, and Canon 60D), which basics are briefly introduced in section 3.3. To this end, we used R packages MASS [57], e1071 [58], RandomForest [59], and caret [60].

3.2 Ground-based panoramic hyperspectral imaging of simulated mine face

In order to prepare for the deployment of the hyperspectral equipment and analyze hyperspectral imagery acquired under conditions closer to those in the mine, we assembled a panel of 52 cm x 57 cm with the samples, along with a white reference (Sphere Optics SG 3141 95%), a Multi-Component Wavelength Standard reference (Labsphere WCS-MC-020) and a color chart (Figure 2). The image was acquired in a dark room by a Specim FX17 camera mounted on a Specim Rotary scanner on top of a tripod at a distance of 2.5 m from the lens to the panel. This distance was selected as a realistic choice in the gallery. Illumination was provided by a Fresnel FilmGear 650w 3200k studio light with a diffuser behind and above the camera. All systems were powered by batteries as they would be under gallery mine conditions. Image resolution was 2.6 mm/pixel and covered

an area of 166.4 cm (height) x 106.1 cm (width), from which the subscene corresponding to the samples in the panel was extracted.

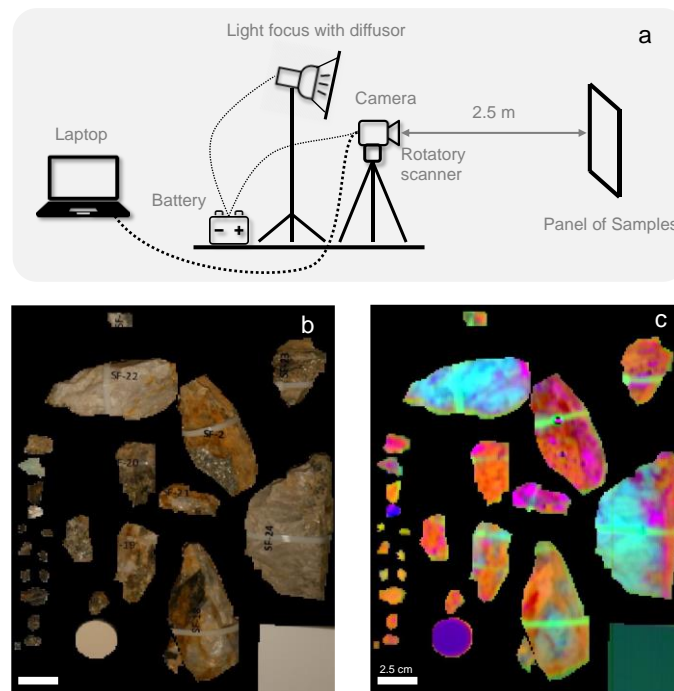


Figure 1.a. Field set-up for the simulation of gallery conditions; b. Conventional photograph of the panel of samples. c. Color composite of the first three principal components of the hyperspectral image.

Despite the diffuser, some illumination unevenness could be appreciated, which we solved by applying a Rolling Ball Background Subtraction [61] in Fiji [62] to the first Principal Component and then applying the inverse transformation. As in the case of laboratory images, the resulting hyperspectral image was displayed in QGIS, and sets of pixel spectra for each mineral and other targets (Table S1) were extracted by interactively digitizing training and validation polygons, a task that required direct visual inspection of the samples as well. Finally, we input the file of polygons along with the hyperspectral image to the classifiers (LDA, SVM, and RF), and their respective accuracy assessments. Each classifier produced a predicted map of the panel materials according to the pre-defined set of categories (Table S1). We used tools in EnMap Box [30], which are based on scikit-learning, for SVM and RF classification, while R packages MASS [57] and raster [41] were used for LDA, as in the case of the laboratory spectra. As combining results of different methods produces more robust results, we also combined the classifications by selecting, for each pixel, the class that had been predicted by most methods. For those pixels in which each method was predicting a different class, we selected the class predicted by the classifier with the highest accuracy.

3.3 Classification processing

3.3.1 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis, like Principal Component Analysis (PCA), is often used as a dimensionality reduction technique before machine learning applications. LDA seeks to project the dataset in a space of fewer dimensions with maximum separability among classes, by maximizing the relationship between the within-class variance and the among classes variance. Unlike PCA, which is an unsupervised algorithm (it does not require training polygons for each class as it maximizes total variance for each principal

component), LDA is supervised: it transforms the original space into components that linearly maximize the separation among classes, based on estimates of the mean and covariance matrix of each class, which are calculated on a training set of labelled observations [55] [63]. Classification is performed by applying a Bayes rule assuming multi-variate Gaussian distributions of common variance, which simplifies to the nearest centroid rule modulo the prior class probabilities in the transformed space, as the covariance matrix becomes the identity matrix. For the present work, we performed LDA using R with package MASS [57] for both the laboratory spectra and the hyperspectral image of the panel.

3.3.2 Support Vector Machines (SVM)

As comprehensively explained in [63], SVM are a generalization of the Maximal Margin Classifier (MMC). Given a p -dimensional space of descriptors and a simple binary (2 classes) problem with n training observations (the method can be generalized to >2 classes), the MMC seeks an $n-1$ hyperplane that linearly separates the two classes with the maximum margin (that is: with the maximum distance from the nearest point of each class to the hyperplane). Maximizing the margin increases the chances of having a correct classification of the rest of the observations (those not included in the training set). On some occasions, the distribution of the observations is such that the correct separation of the 2 classes according to MMC implies that the maximal margin hyperplane still lies very close to some observations and the distance between margins is very narrow. In these cases, a more tolerant approach defining a wider margin even at the expense of some errors in the classification of the training set would result in fewer errors with newer observations. This method is named Support Vectors Classification (SVC) by [63] and linearSVC in the scikit-learn library [64]. The degree of tolerance is based on the relative distances from observations to their class margin: observations at the correct side of their margin or on the margin itself are at a distance 0; observations beyond the margin of their respective class but within the limit of the hyperplane are at distances $0 < e < 1$, while those at the wrong side of the hyperplane are at distances > 1 . The user sets the total tolerance (named "regularization" in the SVM jargon), that is, the allowed sum of all distances from observations to margins. For $C=0$ there is no tolerance, so hyperplane and margins will be set according to MMC. At increasing values of C , the margins can be progressively widened, with progressively more observations being left at the wrong side of the margin or even at the wrong side of the hyperplane, and actually, the hyperplane itself changes as different observations are included. Parameter C thus controls the number of observations that are actually involved in computing the hyperplane and its margins. It is important to note that unlike other classification approaches such as LDA, SVM classification depends only on those observations lying on the margins or beyond them, and these observations are known as support vectors (as each observation is a vector of p descriptors). An optimal value for C is calculated by k -fold cross-validation in the training set for a predefined range of C values.

There are cases in which no hyperplane can separate the classes, no matter the value of C . In these cases, rather than fitting non-linear functions, the SVM approach is to increase the dimensionality by adding new dimensions that are non-linear transforms of the p original ones. As the parameters defining the hyperplane in SVC (or linear SVC) are found based on the inner product between all pairs of training observations, SVM classification (named SVC in scikit-learn) uses a kernel-based approach for this purpose. Popular kernels in SVM classification are the polynomial and, in particular, the radial basis function (RBF) because of its flexibility. For RBF, we have, for any two training observations \mathbf{x}, \mathbf{y} (vectors in space of p descriptors):

$$z = \exp(-\gamma || \mathbf{x} - \mathbf{y} ||^2)$$

where γ is a positive constant that can be thought of as the inverse of the variance: for low values of γ , the class will be very wide, and for small values, the class will be narrow.

As for the case of C , γ can be set by cross-validation within a grid-search strategy as is the case of EnMap Box. R package *caret*, instead, uses an analytical formula to get reasonable estimates of γ and fix it to that value [60].

3.3.3 Random Forest

Random forest classification [65] is a development of classification trees. Classification trees are produced by binary and hierarchically splitting the training set by thresholding the descriptors [63]. At each level, all descriptors are sequentially tried and the best threshold for the best predictor is selected to optimize a class purity metric such as the Gini index. Once the tree is grown in the top-down direction, it is then simplified ("pruned" in the classification trees jargon) in the bottom-up direction according to an error minimization rule. One problem with classification trees is that while the fit to the training set can be very good, the prediction of new cases is very dependent upon the specific training set that has been used.

Results are greatly improved by Random Forests, which introduce two modifications: a set of n trees is constructed (with no pruning) instead of just one, and only a subset of m descriptors is used. The training set is resampled with replacement in n bootstrapped sets and a classification tree is produced for each set, but only a random selection of m descriptors is considered at each level. Each tree is then applied to the test data, which yields as many predictions as bootstrapped training sets, and the final result is the set of most commonly selected ("voted") class. A value of $m = \sqrt{p}$ is a common recommendation, and the number of trees (n) should be large enough to ensure that all descriptors have been considered.

3.3.4 Validation.

Validation was conducted through the analysis of confusion matrices, which were built by cross-validation: Leave-one-out in the case of LDA, and 10-fold cross validation in the case of SVM and RF. We calculated standard metrics from the confusion matrices using R package *caret*: overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA), and F1 (which is the harmonic mean of user's and producer's accuracy [66–68]. Specifically, considering pixels in the digitized polygons,

$$OA = \frac{\text{Total number of Correctly Classified pixels}}{\text{Total number of pixels}}$$

$$PA(i) = \frac{\text{Number of Correctly Classified pixels in class } i}{\text{Number of pixels classified as class } i}$$

$$UA(i) = \frac{\text{Number of Correctly Classified pixels in class } i}{\text{Actual number of pixels in class } i \text{ ("ground truth")}}$$

$$F1(i) = \frac{2 \times UA(i) \times PA(i)}{UA(i) + PA(i)}$$

4. Results

4.1. Laboratory imaging spectrometry

Reflectance spectra of single-mineral polygons of cassiterite, malachite and muscovite measured in this work are similar to those found in the reference spectral libraries (Figure 3). Those of quartz targets are highly variable (as in the case of spectral libraries) because these spectra are very dependent upon the presence of other minerals or fluid inclusions as well as the effect on the optical properties of quartz of impurities and crystalline disorder. In the case of chalcopyrite, a set of spectra were much brighter than the

rest of studied minerals, and also than those from the spectral libraries. This is probably because of specular reflection from well-developed crystals. No reference spectra of wolframite could be found in the spectral libraries. In some cases, the spectra unveiled errors in the “de visu” identification of minerals in the hand samples. These spectra were discarded after the re-examination of the samples confirmed the erroneous original labelling.

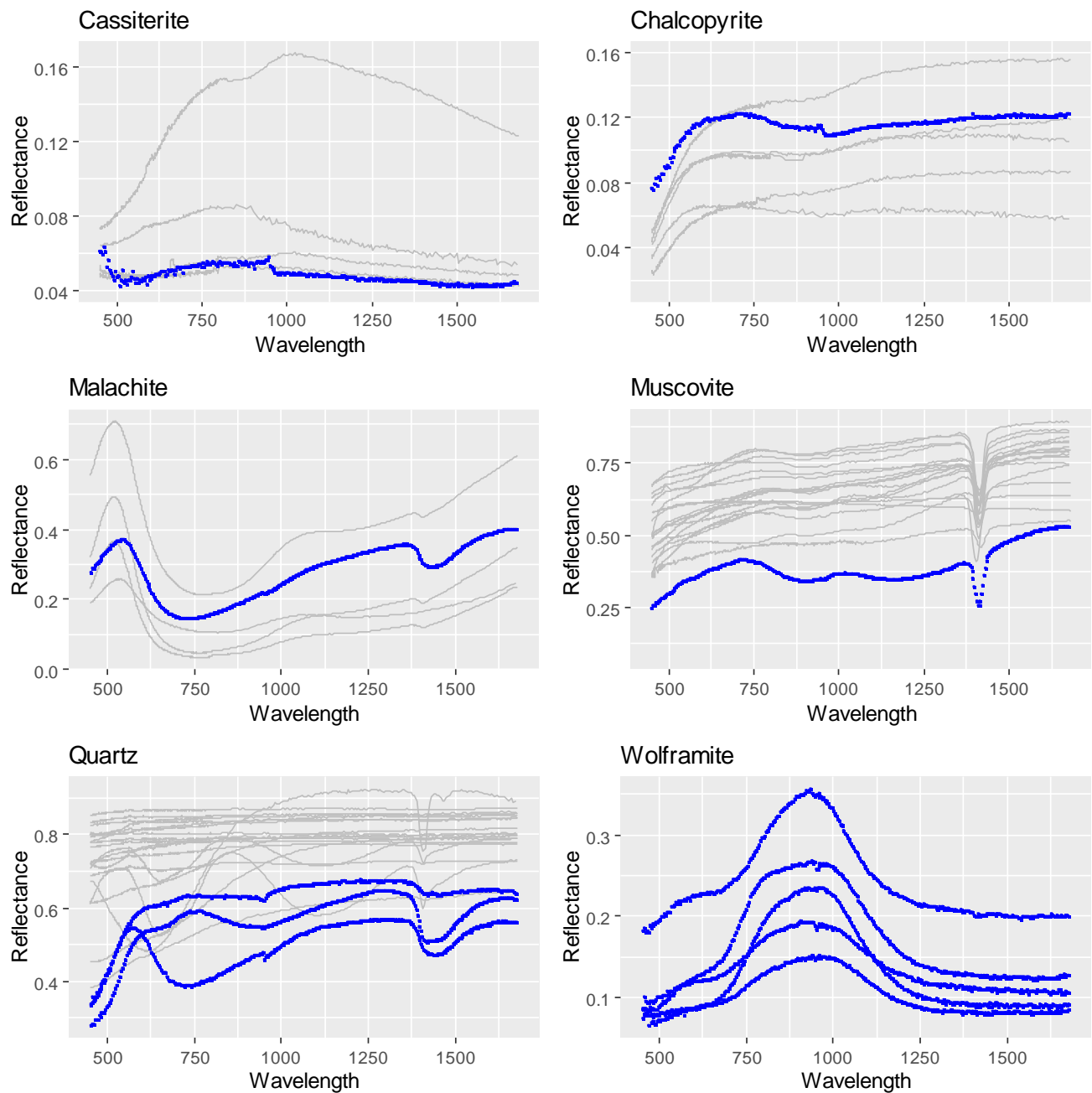


Figure 2. Spectra retrieved from hyperspectral images acquired by scanning samples with Specim cameras FX10 and FX17 in the laboratory (blue). Reference spectra from spectral libraries represented in light gray whenever available.

Spectral separability (Jeffries-Matutita index) among combined spectra of both cameras (450 – 1650 nm) is very high (>1.99) for all mineral pairs. In accordance, LDA classification had very high accuracy (98%, Table 2). Spectral differences are well mapped in the LDA space (Figure. 4), in which spectral samples are well ordinated in consistency with their respective mineral identity. Note that while spectra of wolframite and cassiterite overlap in the plane of the first two components, they are clearly separated by the 4th component. Only 2 spectral samples had conflicting labelling between the “de visu” inspection and the LDA classification.

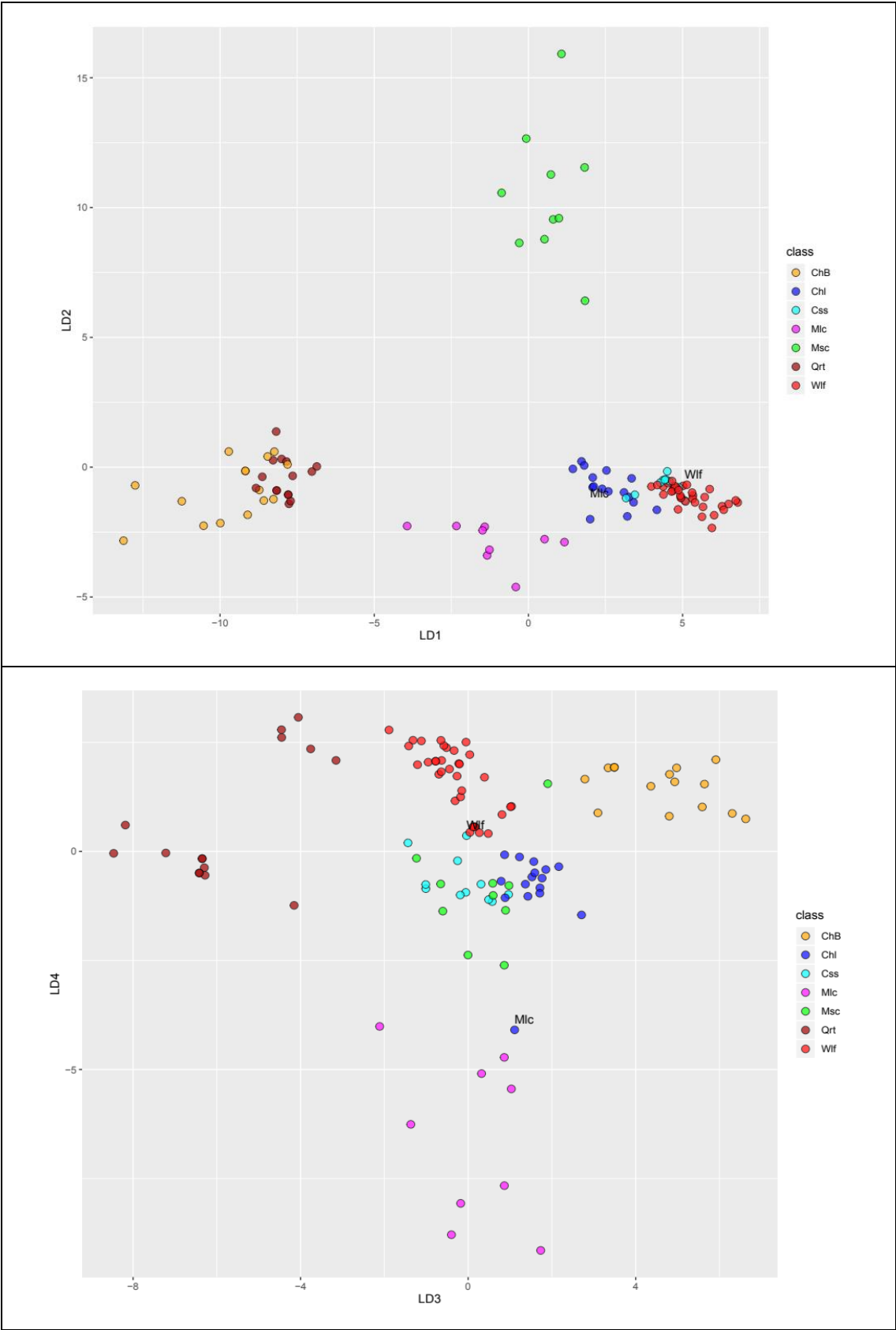


Figure 3. Ordination of laboratory spectra on planes defined by LD components 1 to 4. Colors correspond to classified spectra. Letters indicate the actual mineral of those spectra for which the classification was wrong.

Classification accuracy was very high with data from the hyperspectral cameras, slightly lower when spectra of each camera were taken separately (Table 3): 95% for the FX10 (450 – 950 nm) and 94% for the FX17 (950-1650 nm). Instead, classification accuracy was much lower (0.745) if only simulated RGB values were used. Confusion among minerals (Table 3) slightly increases for chalcopyrite, cassiterite and malachite if only the FX10 spectra are considered, and for cassiterite and muscovite if only the FX17 spectra are considered. Confusion significantly increases if only RGB values are considered, in particular for cassiterite, which is always confused with wolframite, but also for chalcopyrite and wolframite. Interestingly, more modern classification methods such as SVM and RF achieved similar or even slightly lower performance than LDA (Table 4).

Table 2. Accuracy values of classifications by Linear Discriminant Analysis using laboratory spectra with different spectral ranges.

Camera	Spectral range (nm)	Overall accuracy	Producer's accuracy	User's accuracy
FX10_FX17	397 – 1720	0.982	0.982	0.978
FX10	397 – 1004	0.954	0.939	0.964
FX17	936 – 1720	0.945	0.934	0.941
Canon 60D		0.745	0.662	0.828

Table 3. Confusion matrix of LDA classifications using laboratory spectra. Each value corresponds to the number of spectra observed as the mineral indicated by the row and predicted as indicated by the column. Spectral ranges: FX10_FX17: combined spectra (450 – 1650 nm); FX10 (450 – 950 nm); FX17 (950 – 1650 nm); RGB: Canon D60. Overall accuracies were 98.2%, 95.4%, 94.5% and 72.7% in the same order. Acronyms in the columns correspond to the full names in the rows.

	Predicted											
Observed	Instrument		ChalB	Chal	Css	Mal	Mus	Qtz	Wlf	Prod. acc.	Usr. acc.	F1
	Bright chalcopyrite	FX10_FX17	15	0	0	0	0	0	0	1.00	1.00	1.00
		FX10	15	0	0	0	0	0	0	1.00	1.00	1.00
		FX17	15	0	0	0	0	0	0	1.00	1.00	1.00
		RGB	13	0	0	0	0	2	0	0.87	1.00	0.93
	Chalcopyrite	FX10_FX17	0	15	0	0	0	0	0	1.00	0.94	0.97
		FX10	0	15	0	0	0	0	0	1.00	0.75	0.86
		FX17	0	13	2	0	0	0	0	0.87	0.87	0.87
		RGB	0	3	0	0	0	0	12	0.20	0.43	0.27
	Cassiterite	FX10_FX17	0	0	10	0	0	0	0	1.00	0.91	0.95
		FX10	0	2	8	0	0	0	0	0.80	1.00	0.89
		FX17	0	1	9	0	0	0	0	0.90	0.75	0.82
		RGB	0	0	0	0	0	0	10	0.00	0/0	0/0
	Malachite	FX10_FX17	0	1	0	9	0	0	0	0.90	1.00	0.95
		FX10	0	2	0	8	0	0	0	0.80	1.00	0.89
		FX17	0	0	0	10	0	0	0	1.00	1.00	1.00
		RGB	0	4	0	6	0	0	0	0.60	1.00	0.75
	Muscovite	FX10_FX17	0	0	0	0	10	0	0	1.00	1.00	1.00
		FX10	0	0	0	0	10	0	0	1.00	1.00	1.00
		FX17	0	1	0	0	8	0	1	0.80	1.00	0.89
		RGB	0	0	0	0	8	1	1	0.80	1.00	0.89
	Quartz	FX10_FX17	0	0	0	0	0	15	0	1.00	1.00	1.00
		FX10	0	0	0	0	0	15	0	1.00	1.00	1.00
		FX17	0	0	0	0	0	15	0	1.00	1.00	1.00
		RGB	0	0	0	0	0	15	0	1.00	0.83	0.91
	Wolframite	FX10_FX17	0	0	1	0	0	0	34	0.97	1.00	0.99
		FX10	0	1	0	0	0	0	34	0.97	1.00	0.99
FX17		0	0	1	0	0	0	34	0.97	0.97	0.97	
RGB		0	0	0	0	0	0	35	1.00	0.60	0.75	

Table 4. Accuracy values of different classification methods applied to laboratory spectra. LDA, Linear Discriminant Analysis; SVM, Singular Vector Machine; RF, Random Forest.

Classification method	Overall accuracy	Producer's accuracy	User's accuracy
LDA	0.982	0.982	0.978
SVM	0.982	0.987	0.971
RF	0.964	0.958	0.952

4.2. Ground-based panoramic hyperspectral imaging of simulated mine face

Each classification method applied to the FX17 hyperspectral image of the panel produces a map of the distribution of surface materials in the panel. Classification accuracy is lower than in the case of the laboratory spectra, but still high (Table 5): 0.906 – 0.914 if all classes are considered, and 0.813 – 0.849 if only relevant materials are included. Overall accuracy was highest with the RF classifier.

Table 5. Overall accuracy values of different classification methods applied to the FX17 hyperspectral image of the panel of hand samples. LDA, Linear Discriminant Analysis; SVM, Singular Vector Machines; RF, Random Forest.

Classification method	All classes	Relevant classes
LDA	0.906	0.845
SVM	0.908	0.813
RF	0.914	0.849

The confusion matrices (Table 6 and S2) reveal that while a high accuracy was achieved for wolframite, cassiterite is too often confused with the former (in the case of the RF classification, a 37.5% of the all actual Cassiterite pixels are labelled as wolframite), but more rarely confused with any other material (10.6%). In other words, while the identification of wolframite was reliable, pixels identified as cassiterite can actually be either cassiterite or wolframite. As cassiterite and wolframite are the main ores in this mine, an operationally interesting product can result by lumping together both minerals in an “ore” category that reaches 94.8% of user's accuracy and 93.9% of producer's accuracy.

Figure 4 displays the combined result: all three classification methods agree on 53.8% of the study area, while at least 2 methods agree in 87.5% of the study area. Within the training/validation polygons, the results of the three classifiers agree in 90% of the area and, in that case, the agreed class is always correct.

Table 6. Confusion matrix of main minerals in the Random Forest classification of the FX17 (950 – 1650 nm) hyperspectral image of the panel of hand samples. Each value corresponds to the number of spectra observed as the mineral indicated by the row and predicted as indicated by the column. Overall accuracy was 84.9% See table S1 for the complete table.

	Observed											
Predicted		Css	Chal	Mal	Mus	Wlf	Smg	Oxd	Qtz	User's ac- curacy	Pro- ducer's ac- curacy	F1
	Cassiterite	54	0	0	0	14	2	0	4	0.701	0.519	0.597
	Chalcopyrite	3	19	0	0	6	7	0	0	0.528	0.613	0.567
	Malachite	0	0	30	0	0	0	0	0	1.000	1.000	1.000
	Muscovite	0	0	0	19	0	10	0	0	0.655	0.594	0.623
	Wolframite	39	3	0	0	311	0	0	0	0.854	0.912	0.882
	Small grains	3	8	0	10	4	109	0	0	0.741	0.813	0.776
	Oxide	0	0	0	0	0	0	33	0	0.868	0.805	0.835
	Quartz	3	0	0	0	4	0	2	351	0.975	0.989	0.981

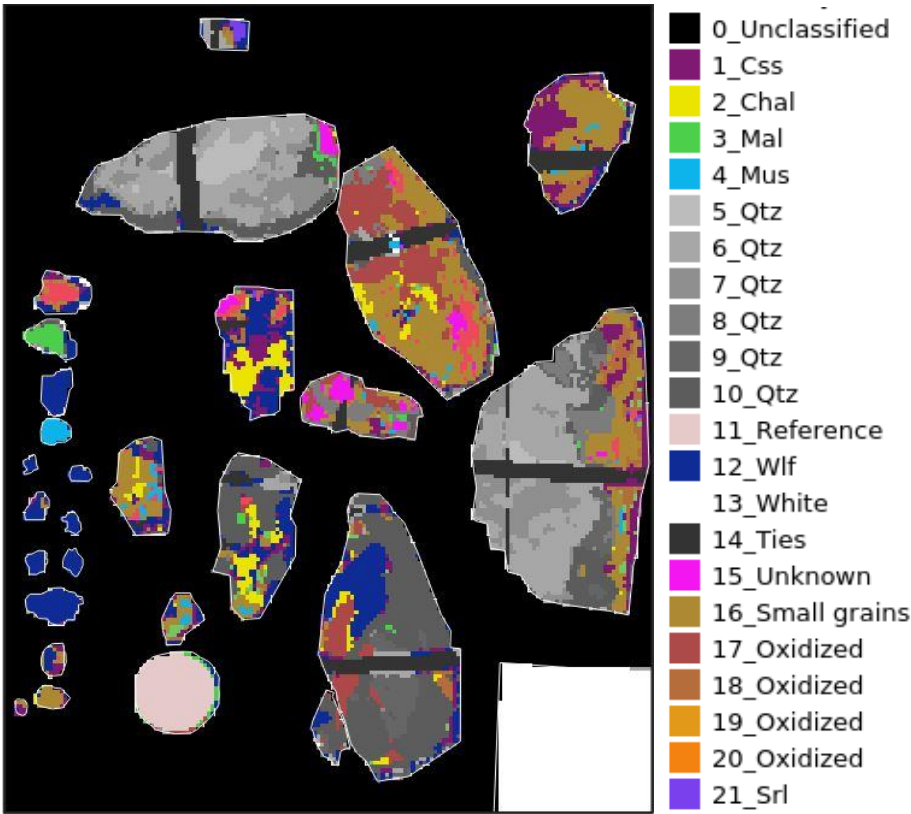


Figure 4. Distribution of materials in the panel of hand samples by combining the result of all 3 classification methods. The agreed class is selected wherever results of 2 or 3 classification methods are coincident (33.7% and 53.8% of the total area, respectively), while the result of RF is kept for those pixels with no agreement (12.5%).

5. Discussion and Conclusions

Three factors can be called to explain the decrease of accuracy between results obtained in the laboratory and in the simulation of gallery conditions: narrower spectral range (450 – 1780 nm in the laboratory while 950 – 1780 nm in the gallery simulation), poorer illumination (dedicated illumination system in the laboratory stand vs. standard photography lighting system in the gallery simulation), and coarser pixel size because increased camera – target distance (0.224 mm vs. 2.6 mm). The narrower spectral range is a consequence of the logistic requirement of using one single hyperspectral imaging system in the gallery. While scanning with a compact system covering the 450 nm - 1780 nm range (or more) would be better, our laboratory results indicate that, in this study, this has not been a substantial factor. Illumination could be more challenging in the actual gallery than in the simulated conditions, but our experience in this study indicates that no parts of the image were critically under-illuminated and that we were successful at correcting moderate illumination unevenness. Coarser pixels, instead, appear to be more responsible for the observed decreased accuracy. Large and uniform targets such as those of the standard reference have a crown of incorrect labels that put in evidence the effect of mixed pixels. While not so conspicuously, this effect is certainly present in other targets and must be more pervasive in those targets of smaller size, such as the case of cassiterite. Fragmentation probably makes the average size of ore minerals in the hand samples smaller than in the actual mine walls, and thus the effect of pixel coarseness at producing spurious mixtures might have been over-emphasized in this study. Notwithstanding, the fact that the minimum size of ore patches being worth extracting can be as small as 1 cm², indicates the interest of using a hyperspectral system with higher spatial resolution for forthcoming studies in the actual mine gallery.

Spatial resolution can also be improved by co-registration of hyperspectral images to conventional digital photography (“RGB”) and subsequent image fusion. Co-registered RGB images of high resolution are useful to add textural information from the high-resolution RGB image to the classification processing, which usually involves a previous segmentation step. Actually, a crude co-registration using 2D coordinates only was used in this study as an aid for photo-interpretation, but the correct co-registration requires 3D coordinates and thus generating a digital surface model [69], an involved task that will be worth addressing in the forthcoming study of the actual mine face.

Our results demonstrate the feasibility and interest of mapping materials of gallery mine faces in Sn-W deposits by analyzing hyperspectral images in the 950 nm – 1780 nm range through machine learning methods. A sequence of these maps at given time intervals as the excavation progress would improve orebody assessment and document the structure of the deposit as a tomography, opening the way to detailed studies of the spatial distribution of ore mineralization and the evaluation of geologic models of the deposit. Our study under simulated gallery conditions encourages forthcoming studies based on deploying ground-based hyperspectral systems in the actual mine gallery, which still will have to face challenges in illumination and surface modelling.

Supplementary Materials

Table S1. Description of target types in the panel of hand samples.

Class Acronym	Material	Description
Css	Cassiterite	Minerals
Chal	Chalcopyrite	
Mal	Malachite	
Mus	Muscovite	
Wlf	Wolframite	
Scl	Schorl	
Qtz_0	Quartz	Differently colored quartz fragments
Qtz_1		
Qtz_2		
Qtz_3		
Qtz_4		
Qtz_5		
Oxid	Oxidized	Surface with appearance of Iron oxide
Alter-1	Alteration	Hydrothermal alteration
Alter-2		
U-1	Unknown	Patches of consisting coloration in the PC composite of unknown material. Small grains of minerals can be observed
Smg	Small grains	Unidentified small grains with similar texture and color
Ref	Reference	WCS-MC-020 reference
White	White reference	Sphere Optics SG 3141 95%.
Tie	Tie	Ties to hold samples to the panel

Table S2. Accuracy metrics of all panel materials: LDA, SVM and RF classifications.

	LDA			SVM			RF		
	Producer's accuracy	User's accuracy	F1	Producer's accuracy	User's accuracy	F1	Producer's accuracy	User's accuracy	F1
Cassiterite	0.590	0.788	0.675	0.524	0.478	0.500	0.519	0.701	0.597
Chalcopyrite	0.450	0.871	0.593	0.833	0.588	0.690	0.613	0.528	0.567
Malachite	1.000	1.000	1.000	1.000	0.742	0.852	1.000	1.000	1.000
Muscovite	0.947	0.529	0.679	0.643	0.600	0.621	0.559	0.655	0.603
Quartz_0	1.000	0.863	0.926	0.894	0.868	0.881	0.849	0.816	0.832
Quartz_1	1.000	0.602	0.752	0.760	0.919	0.832	0.785	0.869	0.825
Quartz_2	0.627	1.000	0.771	0.966	0.838	0.898	0.942	0.903	0.922
Quartz_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Quartz_4	0.905	1.000	0.950	0.750	1.000	0.857	1.000	0.905	0.950
Quartz_5	0.878	0.966	0.920	0.986	0.922	0.953	1.000	0.937	0.967
Reference	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Wolframite	0.951	0.789	0.862	0.769	0.861	0.812	0.912	0.854	0.882
White	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.999
Ties	1.000	0.980	0.990	0.985	1.000	0.992	0.990	1.000	0.995
Unknown	0.596	0.939	0.729	0.808	0.840	0.824	0.818	0.730	0.771
Small grains	0.738	0.716	0.727	0.730	0.818	0.771	0.813	0.741	0.776
Oxidized_1	1.000	0.854	0.921	0.565	0.684	0.619	0.805	0.868	0.835
Oxidized_2	0.733	0.917	0.815	0.700	0.467	0.560	0.667	0.667	0.667
Oxidized muscovite	0.250	0.083	0.125	0.400	0.118	0.182	0.167	0.400	0.235
Oxidized_3	0.750	0.778	0.764	1.000	0.778	0.875	0.889	0.857	0.873
Schorl	0.917	1.000	0.957	0.778	0.583	0.667	0.000		

Author Contributions: Conceptualization, A.L., D.M., J.I. and J-L.F-T.; methodology, A.L.; software, A.L. and G.B.; validation, A.L., E.G. and G.B.; formal analysis, A.L. and G.B.; investigation, A.L., E.G. and G.B.; resources, D.M.; data curation, A.L. and E.G.; writing—original draft preparation, A.L.; writing—review and editing, J-L.F-T. and J.I.; visualization, A.L. and G.B.; supervision, A.L.; project administration, J.I.; funding acquisition, D.M. All authors have read and agreed to the published version of the manuscript.

Funding: EIT RawMaterials supported this research within the framework of the *iTARG3T* (Innovative targeting & processing of W-Sn-Ta-Li ores: towards EU’s self-supply) project nb. 18036.

Acknowledgments: We thank Valoriza Minería, in particular Ivan Losada, for providing the mine hand samples, as well as Marc Campeny-Crego and Fernando Tornos for their useful comments. Hyperspectral imaging was conducted by the Laboratory of Hyperspectral Imaging of GeoSciences Barcelona, CSIC.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Goetz, A.F.H.; Vane, G.; Solomon, J.E.; Rock, B.N. Imaging Spectrometry for Earth Remote Sensing. *Science* **1985**, 228, 1147–1153, doi:10.1126/science.228.4704.1147.

2. Amigo Rubio, J.M. *Hyperspectral Imaging; Data Handling in Science and Technology*; Elsevier, 2020; ISBN 978-0-444-63977-6.

3. Milton, E.J. Principles of Field Spectroscopy. *International Journal of Remote Sensing* **1987**, 8, 1807–1827, doi:10.1080/01431168708954818.

4. Clark, R.N. Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy. In *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*; Rencz, A.N., Ed.; John Wiley & Sons, Inc, 1999; Vol. 3, pp. 3–52.

5. Hunt, G.R.; Salisbury, J.W. Visible and Near-Infrared Spectra of Minerals and Rocks: I Silicate Minerals. *Modern Geology* **1970**, 1, 283–300.

6. Hunt, G.R. Spectral Signatures of Particulate Minerals in the Visible and near Infrared. *GEOPHYSICS* **1977**, 42, 501–513, doi:10.1190/1.1440721.

7. Clark, R.N.; King, T.V.; Klejwa, M.; Swayze, G.A.; Vergo, N. High Spectral Resolution Reflectance Spectroscopy of Minerals. *Journal of Geophysical Research: Solid Earth (1978–2012)* **1990**, 95, 12653–12680.

8. van der Meer, F.D.; van der Werff, H.M.A.; van Ruitenbeek, F.J.A.; Hecker, C.A.; Bakker, W.H.; Noomen, M.F.; van der Meijde, M.; Carranza, E.J.M.; Smeth, J.B. de; Woldai, T. Multi- and Hyperspectral Geologic Remote Sensing: A Review. *International Journal of Applied Earth Observation and Geoinformation* **2012**, 14, 112–128, doi:10.1016/j.jag.2011.08.002.

9. Gupta, R.P. *Remote Sensing Geology*; 3rd ed.; Springer-Verlag: Berlin Heidelberg, 2018; ISBN 978-3-662-55874-4.

10. Riaza, A.; Buzzi, J.; García-Meléndez, E.; Vázquez, I.; Bellido, E.; Carrère, V.; Müller, A. Pyrite Mine Waste and Water Mapping Using Hymap and Hyperion Hyperspectral Data. *Environmental Earth Sciences* **2012**, 66, 1957–1971, doi:10.1007/s12665-011-1422-0.

11. Buzzi, J.; Riaza, A.; García-Meléndez, E.; Weide, S.; Bachmann, M. Mapping Changes in a Recovering Mine Site with Hyperspectral Airborne HyMap Imagery (Sotiel, SW Spain). *Minerals* **2014**, 4, 313–329, doi:10.3390/min4020313.

12. Bolin, B.J.; Moon, T.S. Sulfide Detection in Drill Core from the Stillwater Complex Using Visible/Near-infrared Imaging Spectroscopy. *GEOPHYSICS* **2003**, 68, 1561–1568, doi:10.1190/1.1620630.

13. Kruse, F.A.; Bedell, R.L.; Taranik, J.V.; Peppin, W.A.; Weatherbee, O.; Calvin, W.M. Mapping Alteration Minerals at Prospect, Outcrop and Drill Core Scales Using Imaging Spectrometry. *International Journal of Remote Sensing* **2012**, *33*, 1780–1798, doi:10.1080/01431161.2011.600350. 49 50 51
14. Mathieu, M.; Roy, R.; Launeau, P.; Cathelineau, M.; Quirt, D. Alteration Mapping on Drill Cores Using a HySpex SWIR-320m Hyperspectral Camera: Application to the Exploration of an Unconformity-Related Uranium Deposit (Saskatchewan, Canada). *Journal of Geochemical Exploration* **2017**, *172*, 71–88, doi:10.1016/j.gexplo.2016.09.008. 52 53 54
15. Coulter, D.W.; Zhou, X.; Harris, P.D. Advances in Spectral Geology and Remote Sensing: 2008–2017. In *Proceedings of the Proceedings of Exploration 17: Sixth Decennial International Conference on Mineral Exploration*; Tschirhart, V., Thomas, M.D., Eds.; 2017; pp. 23–50. 55 56 57
16. Murphy, R.J.; Monteiro, S.T.; Schneider, S. Evaluating Classification Techniques for Mapping Vertical Geology Using Field-Based Hyperspectral Sensors. *IEEE Transactions on Geoscience and Remote Sensing* **2012**, *50*, 3066–3080, doi:10.1109/TGRS.2011.2178419. 58 59 60
17. Ramanaidou, E.R.; Wells, M.A. Hyperspectral Imaging of Iron Ores. In *Proceedings of the 10th International Congress for Applied Mineralogy (ICAM)*; Broekmans, M.A.T.M., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp. 575–580 ISBN 978-3-642-27681-1. 61 62 63
18. Okay, Ü.; Khan, S.; Lakshmikantha, M.; Sarmiento, S. Ground-Based Hyperspectral Image Analysis of the Lower Mississippian (Osagean) Reeds Spring Formation Rocks in Southwestern Missouri. *Remote Sensing* **2016**, *8*, 1018, doi:10.3390/rs8121018. 64 65 66
19. Okay, U.; Khan, S.D. Spatial Co-Registration and Spectral Concatenation of Panoramic Ground-Based Hyperspectral Images. *Photogrammetric Engineering & Remote Sensing* **2018**, *84*, 781–790, doi:10.14358/PERS.84.12.781. 67 68 69
20. Krupnik, D.; Khan, S. Close-Range, Ground-Based Hyperspectral Imaging for Mining Applications at Various Scales: Review and Case Studies. *Earth-Science Reviews* **2019**, *198*, 102952, doi:10.1016/j.earscirev.2019.102952. 70 71
21. Pirard, E. Multispectral Imaging of Ore Minerals in Optical Microscopy. *Mineralogical Magazine* **2004**, *68*, 323–333, doi:10.1180/0026461046820189. 72 73
22. Pirard, E.; Bernhardt, H.-J.; Catalina, J.-C.; Brea, C.; Segundo, F.; Castroviejo, R. From Spectrophotometry to Multispectral Imaging of Ore Minerals in Visible and Near Infrared (VNIR) Microscopy. **2008**, *6*. 74 75
23. Berrezueta, E.; Ordóñez-Casado, B.; Bonilla, W.; Banda, R.; Castroviejo, R.; Carrión, P.; Puglla, S. Ore Petrography Using Optical Image Analysis: Application to Zaruma-Portovelo Deposit (Ecuador). *Geosciences* **2016**, *6*, 30, doi:10.3390/geosciences6020030. 76 77 78
24. López-Benito, A.; Catalina, J.C.; Alarcón, D.; Grunwald, Ú.; Romero, P.; Castroviejo, R. Automated Ore Microscopy Based on Multispectral Measurements of Specular Reflectance. I – A Comparative Study of Some Supervised Classification Techniques. *Minerals Engineering* **2020**, *146*, 106136, doi:10.1016/j.mineng.2019.106136. 79 80 81
25. Murphy, R.J.; Taylor, Z.; Schneider, S.; Nieto, J. Mapping Clay Minerals in an Open-Pit Mine Using Hyperspectral and LiDAR Data. *European Journal of Remote Sensing* **2015**, *48*, 511–526, doi:10.5721/EuJRS20154829. 82 83
26. Jakob, S.; Zimmermann, R.; Gloaguen, R. The Need for Accurate Geometric and Radiometric Corrections of Drone-Borne Hyperspectral Data for Mineral Exploration: MEPHySTo—A Toolbox for Pre-Processing Drone-Borne Hyperspectral Data. *Remote Sensing* **2017**, *9*, 88, doi:10.3390/rs9010088. 84 85 86
27. Lorenz, S.; Salehi, S.; Kirsch, M.; Zimmermann, R.; Unger, G.; Vest Sørensen, E.; Gloaguen, R. Radiometric Correction and 3D Integration of Long-Range Ground-Based Hyperspectral Imagery for Mineral Exploration of Vertical Outcrops. *Remote Sensing* **2018**, *10*, 176, doi:10.3390/rs10020176. 87 88 89

28. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.; Heidebrecht, K.; Shapiro, A.; Barloon, P.J.; Goetz, A. The Spectral Image Processing System (SIPS) Interactive Visualization and Analysis of Imaging Spectrometer Data. **1993**, doi:10.1016/0034-4257(93)90013-N.
29. Varshney, P.K.; Arora, M. *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*; 2004; ISBN 978-3-642-06001-4.
30. van der Linden, S.; Rabe, A.; Held, M.; Jakimow, B.; Leitão, P.; Okujeni, A.; Schwieder, M.; Suess, S.; Hostert, P. The EnMAP-Box—A Toolbox and Application Programming Interface for EnMAP Data Processing. *Remote Sensing* **2015**, *7*, 11249–11266, doi:10.3390/rs70911249.
31. Kokaly, R.; Graham, G.E.; Hoefen, T.M.; Kelley, K.D.; Johnson, M.R.; Hubbard, B.E.; Buchhorn, M.; Prakash, A. Multiscale Hyperspectral Imaging of the Orange Hill Porphyry Copper Deposit, Alaska, USA, with Laboratory-, Field-, and Aircraft-Based Imaging Spectrometers. **2017**, 923–943.
32. Camps-Valls, G. Machine Learning in Remote Sensing Data Processing. In Proceedings of the 2009 IEEE International Workshop on Machine Learning for Signal Processing; IEEE: Grenoble, France, September 2009; pp. 1–6.
33. Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent Advances in Techniques for Hyperspectral Image Processing. *Remote Sensing of Environment* **2009**, *113*, S110–S122, doi:10.1016/j.rse.2007.07.028.
34. Cracknell, M.J.; Reading, A.M. Geological Mapping Using Remote Sensing Data: A Comparison of Five Machine Learning Algorithms, Their Response to Variations in the Spatial Distribution of Training Data and the Use of Explicit Spatial Information. *Computers & Geosciences* **2014**, *63*, 22–33, doi:10.1016/j.cageo.2013.10.008.
35. Krupnik, D.; Khan, S.D. High-Resolution Hyperspectral Mineral Mapping: Case Studies in the Edwards Limestone, Texas, USA and Sulfide-Rich Quartz Veins from the Ladakh Batholith, Northern Pakistan. *Minerals* **2020**, *10*, 967, doi:10.3390/min10110967.
36. Simancas, J.F. Variscan Cycle. In *The Geology of Iberia: A Geodynamic Approach: Volume 2: The Variscan Cycle*; Quesada, C., Oliveira, J.T., Eds.; Regional Geology Reviews; Springer International Publishing: Cham, 2019; pp. 1–25 ISBN 978-3-030-10519-8.
37. Ruiz Mora, J.E. Mineralizaciones estannovolfamíferas en Noia y Lousame: estudio previo. *Cadernos do Laboratorio Xeolóxico de Laxe: Revista de xeoloxía galega e do hercínico peninsular* **1982**, 595–622.
38. Mangas Viñuela, J.; Arribas Moreno, A. Estudio de las inclusiones fluidas atrapadas en cristales de casiterita y cuarzo del yacimiento de San Finx (La Coruña, España). *Boletín de la Sociedad Española de Mineralogía* **1989**, *12*, 241–259.
39. Rodríguez-Terente, L.-M.; Fernández-González, M.-Á.; Losada-García, I. Estudio cristalográfico de la Bertrandita de las minas de San Finx (A Coruña, España). *Macla* **2014**, *19*, 1–2.
40. R Core Team R: *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018;
41. Hijmans, R.J. *Raster: Geographic Data Analysis and Modeling*; 2018;
42. Bivand, R.; Keitt, T.; Rowlingson, B. *Rgdal: Bindings for the "Geospatial" Data Abstraction Library*; 2019;
43. Leutner, B.; Horning, N.; Schwalb-Willmann, J. *RStoolbox: Tools for Remote Sensing Data Analysis*; 2019;
44. O'Brien, J. *GdalUtilities: R Wrappers for the GDAL Utilities Executables Shipped with Sf*; 2018;
45. Reudenbach, C. *Link2GI: Linking Geographic Information Systems, Remote Sensing and Other Command Line Tools*; 2019;
46. GDAL/OGR contributors *GDAL/OGR Geospatial Data Abstraction Software Library*; Open Source Geospatial Foundation, 2019;

47. Grizonnet, M.; Michel, J.; Poughon, V.; Inglada, J.; Savinaud, M.; Cresson, R. Orfeo ToolBox: Open Source Processing of Remote Sensing Images. *Open Geospatial Data, Software and Standards* **2017**, *2*, 15.
48. McInerney, D.; Kempeneers, P. Pktools. In *Open Source Geospatial Tools: Applications in Earth Observation*; McInerney, D., Kempeneers, P., Eds.; Earth Systems Data and Models; Springer International Publishing: Cham, 2015; pp. 173–197 ISBN 978-3-319-01824-9.
49. Lowe, D.G. Object Recognition from Local Scale-Invariant Features. In Proceedings of the Proceedings of the Seventh IEEE International Conference on Computer Vision; IEEE: Kerkyra, Greece, 1999; pp. 1150–1157 vol.2.
50. EnMAP Core Science Team *EnMAP-Box 3 - A QGIS Plugin to Process and Visualize Hyperspectral Remote Sensing Data*;
51. QGIS Development Team *QGIS Geographic Information System*; Open Source Geospatial Foundation, 2009;
52. 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.; et al. *USGS Spectral Library Version 7; Data Series*; Reston, VA, 2017; p. 68;.
53. Baldridge, A.M.; Hook, S.J.; Grove, C.I.; Rivera, G. The ASTER Spectral Library Version 2.0. *Remote Sensing of Environment* **2009**, *113*, 711–715, doi:10.1016/j.rse.2008.11.007.
54. Meerdink, S.K.; Hook, S.J.; Roberts, D.A.; Abbott, E.A. The ECOSTRESS Spectral Library Version 1.0. *Remote Sensing of Environment* **2019**, *230*, 111196, doi:10.1016/j.rse.2019.05.015.
55. Richards, J.A.; Jia, X. *Remote Sensing Digital Image Analysis: An Introduction*; 4th ed.; Springer, 2005; ISBN 3-540-25128-6.
56. Jiang, J.; Liu, D.; Gu, J.; Susstrunk, S. What Is the Space of Spectral Sensitivity Functions for Digital Color Cameras? In Proceedings of the 2013 IEEE Workshop on Applications of Computer Vision (WACV); IEEE: Clearwater Beach, FL, USA, January 2013; pp. 168–179.
57. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*; Fourth.; Springer: New York, 2002;
58. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. *E1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*; 2020;
59. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *R News* **2002**, *2*, 18–22.
60. Kuhn, M. Building Predictive Models in R Using the Caret Package. *Journal of Statistical Software, Articles* **2008**, *28*, 1–26, doi:10.18637/jss.v028.i05.
61. Castle, M.; Keller, J. Rolling Ball Background Subtraction Available online: <https://imagej.net/plugins/rolling-ball.html> (accessed on 19 April 2021).
62. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat Meth* **2012**, *9*, 676–682, doi:10.1038/nmeth.2019.
63. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*; Springer Texts in Statistics; 1st ed.; Springer-Verlag New York: New York, 2013; ISBN 978-1-4614-7138-7.
64. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
65. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer-Verlag: New York, 2013; ISBN 978-1-4614-6848-6.
66. Congalton, R.G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote sensing of environment* **1991**, *37*, 35–46.
67. Maxwell, A.E.; Warner, T.A. Thematic Classification Accuracy Assessment with Inherently Uncertain Boundaries: An Argument for Center-Weighted Accuracy Assessment Metrics. *Remote Sensing* **2020**, *12*, 1905, doi:10.3390/rs12121905.

68. Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In *Advances in Information Retrieval*; Losada, D.E., Fernández-Luna, J.M., Eds.; Lecture Notes in Computer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; Vol. 3408, pp. 345–359 ISBN 978-3-540-25295-5.

175
176
177
178

69. Thiele, S.T.; Lorenz, S. Multi-Scale, Multi-Sensor Data Integration for Automated 3-D Geological Mapping Using Hylite. *Ore Geology Reviews* **2021**, *under review*.

179
180
181