International Journal of Mining Science and Technology 33 (2023) 659–674

Contents lists available at ScienceDirect

International Journal of Mining Science and Technology

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

Real-time ore sorting using color and texture analysis

David G. Shatwell [a,b], Victor Murray [c,d,*], Augusto Barton [a]

[a] Hochschild Mining PLC, La Colonia 180, Lima 15023, Peru
[b] Department of Electrical Engineering, Universidad de Ingenieria y Tecnologia - UTEC, Lima 15063, Peru
[c] John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge 02134, USA
[d] Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque 87131, USA

ARTICLE INFO

Article history:
Received 2 June 2022
Received in revised form 6 March 2023
Accepted 11 March 2023
Available online 27 April 2023

Keywords:
Ore sorting
Image color analysis
Image texture analysis
Machine learning

ABSTRACT

Sensor-based ore sorting is a technology used to classify high-grade mineralized rocks from low-grade waste rocks to reduce operation costs. Many ore-sorting algorithms using color images have been proposed in the past, but only some validate their results using mineral grades or optimize the algorithms to classify rocks in real-time. This paper presents an ore-sorting algorithm based on image processing and machine learning that is able to classify rocks from a gold and silver mine based on their grade. The algorithm is composed of four main stages: (1) image segmentation and partition, (2) color and texture feature extraction, (3) sub-image classification using neural networks, and (4) a voting system to determine the overall class of the rock. The algorithm was trained using images of rocks that a geologist manually classified according to their mineral content and then was validated using a different set of rocks analyzed in a laboratory to determine their gold and silver grades. The proposed method achieved a Matthews correlation coefficient of 0.961 points, higher than other classification algorithms based on support vector machines and convolutional neural networks, and a processing time under 44 ms, promising for real-time ore sorting applications.

1. Introduction

Gold grades have steadily declined over the years since most high-grade deposits are currently being mined or have already been depleted [1,2]. This trend has also been observed in metals such as copper [1] and manganese [3]. Mining low-grade deposits require more resources than high-grade deposits to produce the same amount of concentrate, which results in increased operating costs and lower profits [4]. New mining technologies are being developed to optimize mineral processing and make low-grade mining more accessible to keep up with the increasing demand for metals.

Typically, the gold extraction process consists of several stages that form a value chain of operations, including mining, crushing, grinding, and gold recovery [5]. Sensor-based ore sorting is a pre-concentration technology usually implemented between the crushing and grinding stages to classify mineralized rocks from waste rocks. Ore sorting aims to reduce the amount of material that goes into the processing plant without significantly reducing mineral recovery [6].

Ore sorting consists of three main stages. In the first stage, sensors are used to measure the physical properties of the rocks. For example, color cameras may be used to produce color images, X-ray transmission sensors to measure the atomic density, lasers and triangulation cameras to estimate the geometry of the rocks, or near-infrared hyperspectral cameras to produce spectral curves [7,8]. After measuring the particle properties, the data gathered by the sensors is sent to a processing unit, which uses classification algorithms to determine if the rocks are mineralized. Finally, the rocks are physically separated by ejecting them using a system of high-pressure jet nozzles depending on if the processing unit classifies them as ore or waste.

Gold and silver are usually scattered in minimal concentrations within a matrix of other rocks or minerals, making it difficult to detect them directly. However, gold and silver grades in many deposits are heavily correlated with the presence of proxy minerals or elements, which can be detected using the right set of sensors [9]. The problem with this approach is that each mine has unique mineralogy that may even vary in different areas of the same mine. Because commercial ore sorters use standard sorting algorithms, sometimes it is not possible to detect the minerals or elements correlated with gold or silver grade with high accuracy. In those cases, some of the mineralized rocks are classified as waste, and some of

* Corresponding author.
E-mail address: victormurray@seas.harvard.edu (V. Murray).

https://doi.org/10.1016/j.ijmst.2023.03.004

the waste rocks are classified as ore. When this happens, the higher grade of the mineral sent to the plant does not justify the value lost in the material rejected by the sorter, and the use of traditional ore sorting is not economically viable.

Visual ore sorting is a challenging problem because its efficiency depends on several different factors, such as particle size, feed rate, and the optical characteristics of the minerals [10]. In general, no single method works for every type of mineral. The diversity and complexity of the methods usually depend strongly on the mineral system [11]. However, in the past years, several papers have proven that it is possible to classify minerals with relatively high accuracy using only images, for example, from color cameras, which are inexpensive and more accessible to maintain than other sensors, such as X-ray transmission (XRT), which could enable low-cost sorting in the future.

Visual sorting methods can be grouped into two categories: feature engineering and feature learning. Feature engineering methods usually use different algorithms to extract color and texture features from an image. In [12,13], the authors proposed an ore sorting algorithm to classify rock particles from a ferromanganese metallurgical plant using color, Haralick texture, and neural networks. Similarly, in [11], the authors proposed a new ore sorting algorithm using loading vectors as color features, wavelets as texture features, and support vector machines. In [14], the authors proposed two improvements over the method presented by [11]. The first used the minimum redundancy maximum relevance (MRMR) algorithm to reduce the color and texture feature space, while the second used the watershed segmentation algorithm and a voting system to decide the final class of the rock. Then, in [15], the authors proposed another improvement over the method, which consisted in extracting texture features from the RGB (red, green, and blue) and HSV (hue, saturation, and value) channels of the images using Gabor filters without feature selection. Feature engineering methods have also been used in other mining areas, such as rock trace identification using features extracted from 3D point clouds [16].

More recently, different feature learning methods based on convolutional neural networks (CNN) have been proposed to solve problems in mining and other domains. In [17], the authors explored potential solutions using CNN models with different depths, structures, and dataset sizes for coal mineral classification. In [18], the authors used a CNN to estimate the particle density range of coal particles under different light sources. In other domains, in [19], the authors proposed a new type of CNN that uses the energy of the images produced by the last convolution layer in the network to generate texture features. Then, in [20], the authors proposed a CNN model that combines traditional CNN architectures with multiresolution analysis using Haar wavelets. Lastly, in [21], the authors presented a multi-scale rotation-invariant convolutional neural network (MRCNN) to classify different lung textures using Gabor-LBP (local binary patterns) images as inputs to a CNN.

Although the methods mentioned before achieved relatively high classification performances for different ore systems, none of them focused on classifying minerals containing gold and silver. Furthermore, these methods only focus on classifying rock fragments according to ore type but do not present classification results according to mineral grade or the processing time required to perform the classification.

This paper presents an ore sorting algorithm used to classify rocks from an underground mine in the Peruvian Andes, owned by Hochschild Mining PLC. Unlike some commercial ore sorters, the proposed method only uses color cameras, which are cheaper and easier to operate than other ore sorting sensors. The images taken by the color camera are used to extract color and texture features from the rock using image processing algorithms. Then, the features are used to train a classification model based on neural networks. The classification model was trained with two different datasets: (1) images of static dry rocks that were classified into four different classes by an expert geologist and (2) images of static dry rocks that were analyzed in a laboratory to determine their gold and silver grades. The classification performance was evaluated, and the algorithm was optimized to reduce its processing time for future real-time ore sorting. Finally, additional tests were performed with images of moving wet rocks to determine the algorithm's performance in a setting that more closely resembles an actual sorting plant.

The organization of this paper is as follows. Section 2 describes the minerals of the selected mine and the vision system. Section 3 presents the proposed system's digital image processing and machine learning methods. Finally, section 4 presents the performance of the proposed method and compares it to other methods using different classification algorithms.

2. Mineral and vision systems

The selected mine is a gold and silver underground mine located in the Peruvian Andes, over 4200 m above sea level. This mine is characterized by low- and high-sulphidation epithermal mineralized systems hosted by veins, breccias, and dissemination within Tertiary volcanic rocks. This section presents the characteristics of the four most abundant types of minerals found in one of the veins of the selected mine and the vision system used to implement the ore sorting algorithm.

2.1. Mineral characteristics

All the rock fragments used in this paper were extracted from one vein, composed chiefly of quartz veins surrounded by volcanic wall rock. The vein is approximately 2 m wide and oriented at a 45° angle. The transition between the main vein and the wall rock is not abrupt, as there is a zone in between where the wall rock contains smaller quartz veins and sulfides. The rocks found in this intermediate zone are called breccias. Fig. 1 shows a picture of the selected vein, where the breccia, quartz vein, and wall rock can be clearly seen.

The rocks extracted from the main vein (VE) are composed of quartz (SiO_2) and calcite ($CaCO_3$). Although quartz is not an economically valuable mineral, its presence is positively correlated with high gold and silver grades [22]. This is mainly due to the dissemination of sulfides [23], such as acanthite and argentite (Ag2), electrum (a gold-silver alloy), and free gold, which are deposited in the vein by hydrothermal fluids. In the selected mine, sulfides are usually scattered in small particles and are only visible as a light coloration on the rock.

1. Breccia; 2. Quartz vein; 3. Andesite.

Fig. 1. Vein in the selected mine showing breccia, quartz vein and andesite.

The breccia rock fragments (BX) consist primarily of quartz and calcite veins encapsulated in an andesite matrix. Rocks classified as breccia usually have a high grade for the same reason as vein rocks, which is that hydrothermal fluids deposit gold and silver particles in the mineral. Sulfides such as acanthite, galena, sphalerite, pyrrhotite, and pyrite may also be found on BX rocks.

The wall rock is mainly composed of andesite, a type of volcanic rock [24]. In the selected mine, two types of rocks have little to no economic value: pure andesite (AN) and altered andesite (AA). The difference between the two types of waste rock is the color and alteration: AN has a green and dark gray color and no significant alteration, while AA has a purplish-gray color due to the presence of oxides and clay. Other minerals found in the wall rock are pyrite and thin quartz veins, which may contain gold or silver particles in rare cases. Fig. 2 shows images of the four types of minerals described before.

Because gold and silver cannot be detected visually, correlations between the mineral classes and their respective equivalent silver grades are used to classify them as ore or waste. Table 1 shows grade statistics for the four different mineral classes acquired in two different assay campaigns performed from January to March 2020 and January 2022. All the rocks were extracted from the same vein at different times. The rocks from 2020 were manually selected by a geologist and are mostly ideal samples from each mineral class. In contrast, the rocks from 2022 were randomly selected and are more heterogeneous since a small portion of them contained multiple classes of minerals. We observed that the rocks from 2022 had lower average grades and a higher standard deviation, which could be explained by the fact that the section of the vein that was mined on January 2022 had lower grades in general, and also because the rocks selected in the first assay campaign were mostly ideal samples of the mineral classes. This difference can be further noticed by the percentage of rocks above the cut-off grade in both campaigns, which is also shown on Table 1 and graphically in Fig. 3. In 2020, all ore rocks were above the cut-off, but in 2022 only 66% of them followed this trend. Despite the difference, in both campaigns, breccia (BX) rocks had the highest average grade, followed by vein rocks (VE), andesite (AN), and altered andesite (AA).

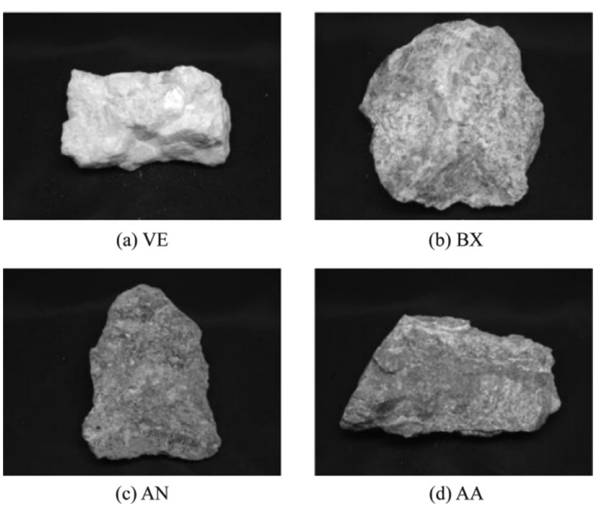


Fig. 2. Rock types found on the selected vein: vein rock (VE), breccia (BX), pure andesite (AN), and altered andesite (AA).

2.2. Vision system

A typical vision system consists of the hardware used to capture and process the images, and the software, which consists of the images and the algorithms used to analyze them. This section describes the hardware components and the images captured and processed by said hardware.

2.2.1. Hardware

The hardware used in the proof-of-concept stage to test the ore sorting algorithms consists of four main parts: a color camera, a structure to hold the camera, a professional light box to illuminate the rocks uniformly, and a processing unit that contains the algorithms. The camera that was used in the vision system is a Sony DSC-HX90V. This is a consumer photography camera that can capture high-resolution images using a 1/2.3″ type (7.82 mm) Exmor R™ CMOS sensor [25]. Later, we also performed additional tests with a semi-industrial camera under more realistic conditions. The settings used to capture the images are a focal length of 50 mm, an exposure time of 1/40 s, and an ISO (International Organization for Standardization) speed of 200. The structure used to hold the camera was fabricated using MDF (medium-density fiberboard) planks cut with a CNC (computer numerical control) router, aluminum tubes, and a tripod head. The structure was used to manually adjust the angle of the camera and the distance between the camera and the rocks. In this paper, all images were acquired at a distance of 25 cm and an angle of 70°. Also, a black cloth covered the structure in order to prevent external light from altering the measurements. The light box used to illuminate the rocks with uniform light was a gti PDV-2e/M3 [26], which includes four incandescent and three D65-type fluorescent light bulbs with a color temperature of 5000 K. For this vision system, only three fluorescent light bulbs were used. Finally, to deploy the analysis and classification algorithms, the vision system used a 2014 MacBook Pro with a 4-core Intel Core i7 processor running the macOS Catalina operating system. The camera, structure, and lightbox are shown in Fig. 4.

Table 1
Equivalent silver (AgEq) grade statistics of the four main classes of mineral found on the selected mine: average (avg.), standard deviation (SD), 0th, 25th, 50th, 75th, and 100th percentiles, and percentage of samples above the cut-off grade.

Mineral	Mass (kg)	Avg. (ppm)	SD (ppm)	Percentiles (ppm)					Above cut-off
				0	25	50	75	100	
First campaign: January-March 2020									
VE	3.7	214	116	97	118	197	230	449	100.0%
BX	0.8	333	229	172	172	333	495	495	100.0%
AN	9.5	26	22	3	13	18	28	86	0.0%
AA	2.1	61	53	14	24	48	98	134	25.0%
Second campaign: January 2022									
VE	16.6	201	158	15	74	167	271	828	66.3%
BX	30.1	225	242	10	63	139	278	1400	65.9%
AN	4.9	28	88	5	6	11	18	672	3.3%
AA	3.0	52	99	10	26	31	41	851	7.0%

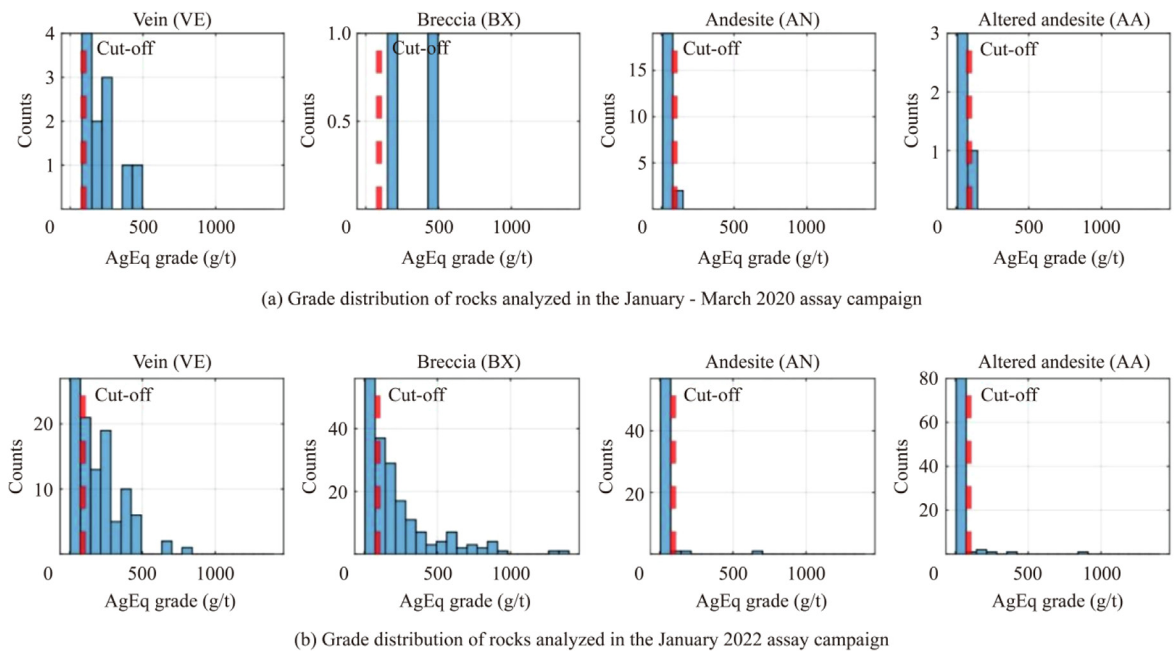


Fig. 3. Grade distribution of each class of mineral (Table 1).

Fig. 4. MDF structure used to position the camera and the light source at a fixed distance from the rocks, thus eliminating most of the external variations in brightness and pixel size.

2.2.2. Image dataset

Digital images are two- or three-dimensional numerical arrays that contain spatial information about a scene and are captured using an imaging device, such as a color camera. Color digital images are usually represented by a set of three matrices, where each one is a monochromatic image that captures light intensity in the red, green, or blue channels, centered in the 665, 550, and 470 nm wavelengths, respectively. Mathematically, color images are usually represented by $I(m, n, c)$, where m and n represent the vertical and horizontal spatial coordinates, and c represents the color or spectral channel.

The Sony camera used in this vision system captures images with a spatial resolution of 3672×4896 pixels and uses the standard RGB color space (sRGB). The pixel information is encoded in 3-byte strings, where each byte represents the intensity of one of the color channels. In addition, the images are saved in JPEG (Joint Photographic Experts Group) format with the camera's fine quality option, which uses less compression when storing images [25]. Although the images were taken with the consumer product Sony DSC-HX90V camera, in later tests, we used a semi-industrial Basler camera with a resolution of 5 megapixels (more information in the sub-section 4.5). Therefore, in order to be able to work with a resolution similar to that of the industrial camera, the images were decimated by a factor of two using a low-pass filter to avoid aliasing. This procedure generated new images with a spatial resolution of 1836×2448 pixels. Given that the images represented a surface of 151 mm $\times$ 201 mm, using the new resolution allowed the algorithms to analyze mineral structures of up to 82 μm, which is equivalent to a pixel density of 12.2 px/mm. All images contain only one rock, such as the examples shown in Fig. 2.

In order to design and test the mineral classification algorithms, three data sets of rock images were constructed. Initially, Hochschild Mining PLC provided 196 rock samples extracted from one of the main veins in the selected mine in 2020. This vein was selected because (1) in recent years, it has provided almost 17% of the extracted mineral, and (2) the mineral is highly representative of the whole mine. The 196 rocks were selected manually by a geologist and contained "ideal" features of each mineral class. The rocks in this data set are mostly homogeneous, which means that a randomly selected region in the rock is roughly equivalent to any other region. The 196 rocks were divided into two groups: the first with 156 samples (79.6%) and the second with 40 samples (20.4%), which follows the typical 80%/20% dataset split done to validate machine learning algorithms. The second group, with 40 samples, contains mostly quartz and andesite rocks, with only a few samples from the breccia and altered andesite classes. Initially, the plan was to include more rocks of the latter two classes; however, doing this was not possible because of budget constraints at the time. One to five images were captured for each rock since images can only capture one face of the rock, and rocks may have multiple faces with different visual properties. In total, 465 images were captured for the first and 155 for the second group, as shown in Table 2. Then, in 2022, Hochschild Mining PLC provided an additional 435 samples from the same vein to validate the algorithm further using images of moving rocks in the newly created test bench. The samples corresponding to the third group were photographed using a semi-industrial camera, producing 867 new images. More details about the imaging process of the rocks from the third group of samples are provided in the sub-section 4.5.

Table 2
Number of rocks and images used for training and testing the performance of the proposed algorithms.

Mineral	Group 1		Group 2		Group 3	
	Rocks	Images	Rocks	Images	Rocks	Images
VE	77	234	11	62	104	208
BX	33	44	2	5	185	370
AN	25	137	23	77	60	117
AA	21	50	4	11	86	172
Total	156	465	40	155	435	867

After forming the image data set, the samples from groups two and three were sent to a geochemical laboratory to perform assays and estimate their gold and silver grades. As mentioned in Section 2.1, the assays were performed on two different campaigns in 2020 and 2022 and are shown on Table 1. The methods used by the laboratory to estimate the grades were fire assays and aqua regia, which are the industry standard analysis processes. The detection range of the assays was 0.005–10 ppm for gold and 10–1000 ppm for silver. After estimating the individual gold and silver grade of each rock in the second and third groups, the equivalent silver grade (G_{AgEq}) was calculated as follows

$$G_{\mathrm{AgEq}} = G_{\mathrm{Ag}} + fG_{\mathrm{Au}} \tag{1}$$

where G_{Ag} is the silver grade; G_{Au} the gold grade; and f a conversion factor that is based on the ratio between the gold and silver price. This paper uses $f=81$ to calculate the equivalent silver grade.

It is worth mentioning that although the assays are performed on the whole rock, the images captured with the color cameras only contain information about the surface. This means that there could be rocks that appear to be barren on the surface but are actually mineralized and vice versa. However, because of the small size of the rocks (mesh between 3/4″–5″), we assume that the surface of the rock contains roughly the same physical properties as the interior, which most sensors and algorithms of commercial sorters also do.

3. Image processing and machine learning algorithms

This section describes the main methods proposed for the ore sorting algorithm. First, the machine vision problem is presented to give context to the algorithms from an image processing perspective, and then, the analysis and classification algorithms are described. The method consists of four main stages: (1) segmentation and partition of input images, (2) feature extraction, (3) classification of sub-images, and (4) the voting algorithm to decide the overall class of the rocks.

3.1. Machine vision challenge

In most cases, ore sorting is not a trivial problem because the algorithms that are needed strongly depend on the types of minerals found in the specific mine [11]. Several convolutional neural network (CNN) architectures have been proposed in the last decade to solve various image classification problems. This new type of algorithms can automatically learn the feature representations from input images [27], thus eliminating the need to use a manual feature extraction stage. However, CNNs usually require large datasets, which are costly and challenging to create in mining. This is especially true for ore sorting algorithms, in which the classification results need to be validated by performing chemical assays to the rocks on the test set and often also on the training set.

In the selected mine, developing ore sorting algorithms has additional challenges. Although vein (VE) and andesite (AN) rocks can be easily classified using color features, breccia (BX) and altered andesite (AA) rocks have very similar colors. The only difference is the light shades of purple found on AA compared to BX. The other challenge related to BX and AA is their high-class variability. The amount of quartz present in BX and the amount of white clay present in AA may change drastically in each rock. Because quartz has a similar color to clay, this creates further problems in their classification. Classifying these two types of minerals correctly is crucial because BX rocks usually have a high grade, while AA rocks usually have a low grade.

Finally, the last challenge related to the development of sorting algorithms is the processing time. Commercial sorters may use belt speeds of up to 3 m/s to process the material extracted from the mines. Considering an image surface area of 151 mm × 201 mm, which is the area of the scene captured by the camera in the proof-of-concept stage, the entire processing time for a single image should be less than 70 ms, equivalent to a frame rate of 14.3 fps (frames per second). This requirement severely limits the amount of color and texture features that can be used for classification, as each additional feature that is computed increases the system's processing time. For this reason, it is convenient to use features that require a low number of operations to compute or are highly parallelizable.

Solving the machine vision problem requires a processing pipeline composed of several stages (Fig. 5). Each image is segmented and split into several sub-images in the first stage. The use of sub-images allows the algorithm to identify all the possible minerals present in the rock and use different strategies to decide whether they should be sent to the processing plant or waste dump. The second stage involves extracting each sub-image color and texture features using statistical and image processing algorithms. In the third stage, one or more neural networks in parallel are used to assign one mineral class to each sub-image, creating mineral distribution maps. Finally, the fourth stage uses a voting algorithm to assign a single class to the whole image by counting the number of sub-images that belong to each mineral class.

3.2. Segmentation and sub-images

This section presents the methods used in the first stage of the algorithm, which contains the image segmentation and partition blocks, shown in Fig. 5. The input to this stage is the color image $I(m.n.c)$, and its output is a set of N_{v} sub-images of the rock. In this case, N_{v} represents the number of valid sub-images that do not contain part of the background, as will be explained later in this section. Fig. 6 shows a detailed block diagram representing all the operations performed in this stage.

Many of the rocks in a real mining setting contain more than one type of mineral. The most common cases are andesite rocks with small quartz veins, which may have a low or medium grade depending on the number of veins in the rock. Identifying the distribution of the different minerals in the rocks is essential in order to classify them as ore or waste correctly. For this reason, the algorithm splits the input image into 1064 sub-images with dimensions of 64×64 pixels (5.25 mm × 5.25 mm), denoted by $S_p(m,n,c)$, where p=1, … , 1064. Because each sub-image

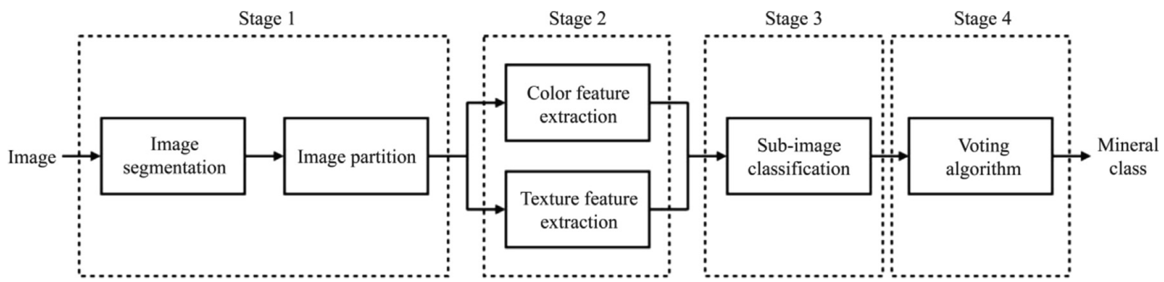


Fig. 5. Block diagram of the proposed mineral classification algorithms.

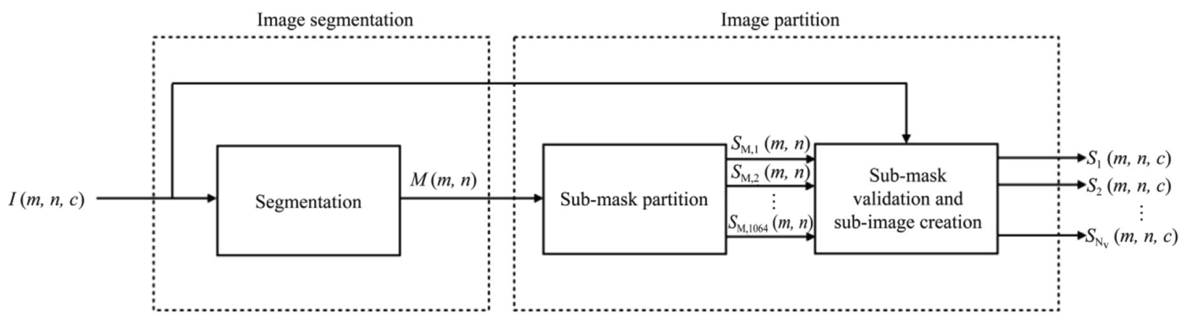


Fig. 6. Block diagram of the preprocessing stage of the algorithm.

$S_p(m,n,c)$ contains information about the local color and texture of the image, classifying sub-images instead of whole images has two main advantages: (1) each sub-image has a higher probability of containing a single type of mineral due to its small size, and (2) classification at the sub-image level generates a much larger number of training samples, which makes the classification algorithm more robust and accurate.

In order to avoid classifying sub-images that are part of the background and do not contain any part of the rock, the algorithm uses a basic but effective segmentation method. This method consists of two steps: first, it produces a grayscale image by adding the three color channels of the color image $I(m,n,c)$, and then, it thresholds the image in order to obtain a binary mask, denoted by $M(m,n)$. Mathematically, this method is defined by

$$M(m,n) = \begin{cases} 1, \text{ if } \sum_{c=1}^{C} I(m,n,c) \geq T_M \\ 0, \text{ if } \sum_{c=1}^{C} I(m,n,c) < T_M \end{cases} \tag{2}$$

where $\sum_{c=1}^{C} I(m,n,c)$ represents the sum of the image's red, green, and blue channels; C=3 the total number of color channels; and T_M the decision threshold, determined experimentally as $T_M = 100$ by using image intensity histograms. Because the image and mask have the same number of pixels, each sub-image has a corresponding sub-mask $S_M(m,n)$ of 64×64 pixels. Therefore, to decide whether a sub-image should be processed, the algorithm counts the number of pixels in the sub-mask equal to 1, representing those containing a section of the rock. If at least 90% of the pixels in the sub-mask is equal to 1, the sub-image is processed in the subsequent stages of the algorithm (Fig. 7 as an example of the process). As mentioned earlier in this section, the result of this stage is a set of N_v valid sub-images $S_p(m,n,c)$ that contain sections of the rock but not of the background.

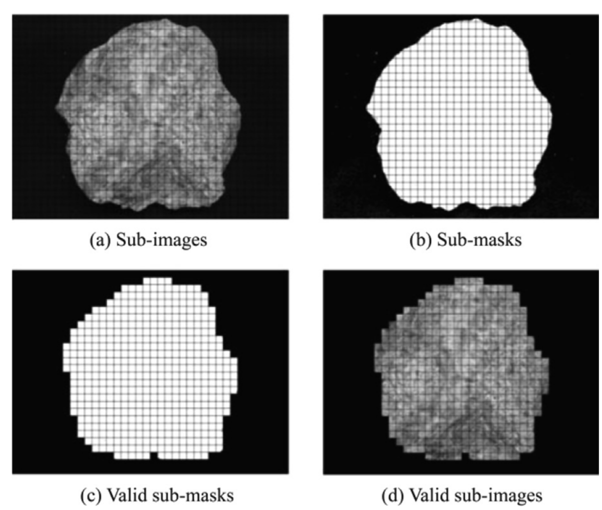


Fig. 7. Visual example of the image segmentation and partitioning process.

3.3. Feature extraction

This section presents the methods used in the second stage of the algorithm, which contains the color and texture feature extraction blocks, shown in Fig. 5. The input to this stage is the set of N_v valid color sub-images, and the output is a set of N_v feature vectors, each one with 24 color features and 12 texture features (36 features in total). Fig. 8 shows a detailed block diagram representing the algorithms used in this stage.

Once all the valid sub-images have been identified, the next stage in the algorithm pipeline is used to extract a feature vector from each one of them. As mentioned in Section 1, several papers, such [11,14,15], have demonstrated that color and texture features can be used to classify different types of minerals. The color features used successfully in the past include color statistics and multi-way principal component analysis (MPCA), while the texture features include Haralick features, wavelets, Gabor filters, or amplitude modulation frequency-modulation (AM-FM) [28]. This paper presents a feature extraction stage that uses color statistics, principal component analysis, and wavelets. Other methods, such as Haralick features and local binary patterns (LBP), were also eval-

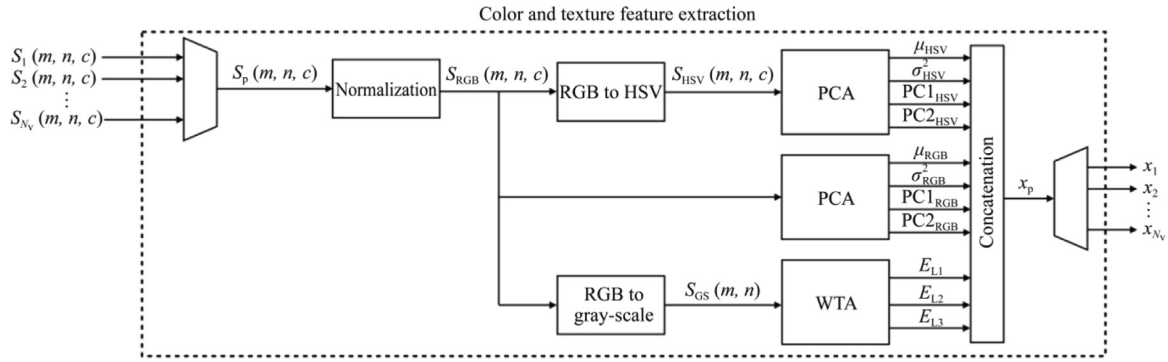


Fig. 8. Block diagram of the color and texture feature extraction stage.

uated, although the results are not presented in this paper because of their high computational time.

3.3.1. Color feature extraction

The color feature extraction stage begins by normalizing each sub-image $S_p(m,n,c)$ to [0,1] and storing the resulting values in the new sub-image $S_{RGB}(m,n,c)$. Then, the sub-image is transformed from the RGB to the HSV color space, producing the new sub-image $S_{HSV}(m,n,c)$. The HSV color space describes colors similarly to how humans interpret them. Hue is a characteristic that describes pure colors, saturation refers to how diluted the color is with white, and value is related to the brightness of the color. The equations used to convert a color from the RGB to HSV space are widely known and can be found in [29]. In the early stages of the algorithm development, we found that using both sets of channels produced better classification results in the training set than using only the RGB color space. The main advantage of using the HSV color space, in addition to the RGB color space, is that it separates color from intensity, providing a new way to perform feature extraction over the channels. Another advantage is that it provides more color features for the classification algorithms in later stages, and performing the transformation is computationally inexpensive.

The RGB and HSV sub-images are each composed of three color channels each, which are represented by the grayscale sub-images $S_R(m,n)$, $S_G(m,n)$, $S_B(m,n)$, $S_H(m,n)$, $S_S(m,n)$, and $S_V(m,n)$. The first 12 color features computed for the sub-images are the mean (μ_{RGB}, μ_{HSV}) and variance ($\sigma^2{}_{RGB}$, $\sigma^2{}_{HSV}$) vectors of the pixels in each color space. These two statistical parameters are useful color features in this ore sorting problem because their computation is already needed in the principal component analysis (PCA) algorithm, so they do not require additional computational costs.

The next set of color features is the principal components of the RGB and HSV sub-images, which contain information about the directions of the most significant variance of both color spaces. Unlike statistical parameters such as the mean and variance, principal components consider the correlation between color channels, which makes them good descriptors of the overall color and contrast of the image [30]. Using the principal components and the mean and variance of each color space allows the algorithm to fully represent the sub-image color information. PCA can be computed using several methods, such as singular value decomposition (SVD) [31], eigenvalue decomposition of the covariance matrix [32], and the alternating least squares algorithm [33]. The method proposed in this paper uses the eigenvalue decomposition of the covariance matrix to compute the principal components because it is the fastest algorithm of the three mentioned before.

The first step to compute the principal components of a sub-image $S_{CS}(m,n,c)$ with dimensions 64×64×3 is to reshape it into a matrix $\mathbf{X_{PCA}}$ with dimensions 1064×3, where the subscript *CS* can be used to represent either the RGB or HSV color spaces. The main objective of PCA is to find a new set of uncorrelated variables that maximize the variance of the data while minimizing the loss of information [34]. These new variables, called principal components (PC), are represented by orthonormal vectors, whose directions are estimated by calculating the eigenvectors of the covariance matrix of the dataset.

Mathematically, a data point that belongs to matrix $\mathbf{X_{PCA}}$ can be represented by a column vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$ of size $n \times 1$, where each value $x_1, x_2, \cdots, x_n$ represents a random variable and T is the transpose operator. To find the principal components of a data set, the first step is to compute the covariance matrix given by

$$\mathbf{C_x} = \mathrm{E}\left\{(\mathbf{x} - \mathbf{m_x})(\mathbf{x} - \mathbf{m_x})^T\right\} \tag{3}$$

where $\mathbf{m_x}$ is a vector composed of the mean value of each random variable in $\mathbf{x}$. The next step is to find the eigenvalues λ_1, λ_2, ..., λ_n and eigenvectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, ..., $\boldsymbol{e}_n$ of the covariance matrix $\mathbf{C_x}$. In PCA, the eigenvalues are proportional to the variance contribution of their respective eigenvector. Thus, the eigenvector with the highest eigenvalue is considered the first principal component, followed by the second eigenvector with the highest eigenvalue, and so on. The eigenvectors can then be concatenated horizontally in order to form a new matrix-$\mathbf{A} \in \mathbb{R}^{n \times n}$, which is called the Hotelling transform [35]. This transform assigns each data point, represented by $\mathbf{x}$, a new vector $\mathbf{y}$, whose components are uncorrelated.

Because the eigenvectors with lower eigenvalues account for a small portion of the total variance of the data, it is possible to discard them to reconstruct $\mathbf{X_{PCA}}$ with minimal loss of information. In this case, the matrix $\mathbf{A}_k$, which is composed of the k eigenvectors with higher eigenvalues, is used to reconstruct the dataset so that the approximation of each data point $\mathbf{x}$ is given by

$$\hat{\mathbf{x}} = \mathbf{A}_k^T\mathbf{y} + \mathbf{m_x} \tag{4}$$

Applying the PCA algorithm to the RGB and HSV sub-images produces, in total, six principal components $PC_{i,RGB}$ and $PC_{i,HSV}$, where $i = 1, 2, 3$, and each component is a vector of dimensions 3×1. The color feature extraction stage proposed in this paper only uses the first and second principal components because they represent, on average, 99.93% of the variance in RGB sub-images and 99.54% of the variance in HSV sub-images. These two values were found experimentally by analyzing the cumulative variance of the principal components of each sub-image in the dataset. Because the four selected principal components are three-dimensional vectors, the algorithm produces 12 additional color features in total. The 12 PCA features are then concatenated with the 12 features derived from channel statistics in order to form a color feature vector for each sub-image with dimensions 1×24.

3.3.2. Texture feature extraction

The visual texture of an image is defined by the local variations in intensity generated by the roughness or unevenness on the surface of an object [36]. Wavelet decomposition is a method commonly used to analyze image textures that decomposes a grayscale image into a set of lower resolution images containing frequency information at different scales [35]. Wavelet analysis decomposes an image using high-pass and low-pass filter banks, which must meet specific properties so that the image can be perfectly reconstructed [37]. The filtered images are then decimated by a factor of 2, reducing their resolution by half and eliminating redundant information. The filters used to decompose the image are closely related to a specific family of wavelets, chosen according to the specific application.

Analyzing a grayscale image using wavelets produces four new matrices. The first matrix, called the approximation image, contains the lower frequencies of the original image. The second, third, and fourth matrices, called the horizontal, vertical, and diagonal detailed images, contain high-frequency information about the original image in their respective directions. Because the approximation image is just a lower resolution version of the original image, wavelet analysis can be applied recursively to the approximation image an arbitrary number of times to produce a multiresolution analysis [38].

The filters used by the proposed texture feature extraction stage are associated with the two-coefficient Haar wavelet. Although many wavelets can be used to decompose an image, the Haar wavelet is the simplest to implement and the least computationally demanding [39], which is desirable in a real-time classification system. The filters associated with this type of wavelet are $h_{\phi}(n)=\frac{1}{2}\left\{\sqrt{2},\sqrt{2}\right\}$ and $h_{\psi}(n)=\frac{1}{2}\left\{\sqrt{2},-\sqrt{2}\right\}$, where $h_{\phi}(n)$ is the low-pass filter and $h_{\psi}(n)$ is the high-pass filter. The proposed feature extraction stage uses these filters to generate the first three wavelet analysis levels, including 4 approximations and 8 detailed images. Wavelet texture analysis (WTA) is a method for extracting texture information from approximation and detailed images. The most common method for extracting texture information is by computing the energy (E) of each matrix [40], defined by

$$E=\sum_{m=1}^{M}\sum_{n=1}^{N}\left|S_{\mathrm{GS}}(m,n)\right|^{2} \tag{5}$$

where $S_{\mathrm{GS}}(m,n)$ is the grayscale sub-image, m and n are the vertical and horizontal coordinates of each pixel; M the numbers of rows of the sub-image; and N the numbers of columns of the sub-image. WTA is based on the assumption that the feature vectors of similar textures form clusters in the feature space, which are different from other clusters that belong to different textures. For this reason, the proposed algorithm uses the energy of the 12 resulting matrices as texture feature vectors to characterize the texture of each sub-image.

After the color and texture feature extraction stages, the N_{v} valid sub-images are transformed into feature vectors x_p with 36 elements, where the first 12 come from color statistics, the next 12 come from the vector coefficients obtained through PCA, and the last 12 come from wavelet texture analysis. These vectors are a low-dimensional representation of the sub-images, which reduces the dimension space from 12228-pixel intensity values (4096 for each color channel) to just 36 real decimal numbers, which enables the use of simple classification algorithms in later stages.

3.4. Classification mode

This section presents the methods used in the third stage of the algorithm, which contains the machine learning models used to classify the sub-images as shown on Fig. 5. The input to this stage is the set of N_{v} feature vectors, each representing the color and texture of a sub-image, and the output is a set of probability maps. Depending on the data set used to train the algorithm, the probability maps represent the likelihood that the sub-images belong to each of the four classes of minerals or the likelihood that the sub-images are mineralized. Fig. 9 shows a detailed block diagram representing the algorithms used in the classification and voting stages.

After extracting the 36 color and texture features from each sub-image, the next stage uses classification models to find their respective classes. The ore sorting algorithms proposed in this paper use only artificial neural networks (ANN) in the classification stage, although other classification methods were also tested, as will be explained in section 4.

ANN are a class of regression and classification algorithms that can be used to approximate unknown functions from a collection of input and output data points. One of the main advantages of neural networks over other regression and classification algorithms is that they do not require prior knowledge of the data distribution, which is especially useful when working with data with a large number of variables [41].

Neural networks are composed of basic units, called artificial neurons or neural units, which are nonlinear functions that take as input a feature vector $\mathbf{x}\in\mathbb{R}^{d}$, a weight vector $\mathbf{w}\in\mathbb{R}^{d}$, and a bias $b\in\mathbb{R}^{d}$. The output of the neural unit is given by

$$\widehat{y}=g\left(\mathbf{w}^{T}\mathbf{x}+b\right) \tag{6}$$

where $g(\cdot)$ is a non-linear function known as the activation function. In order to train a neural network from a set of input and output data points, one must find the weights that minimize the cost function of the neural network, which is usually done with the gradient

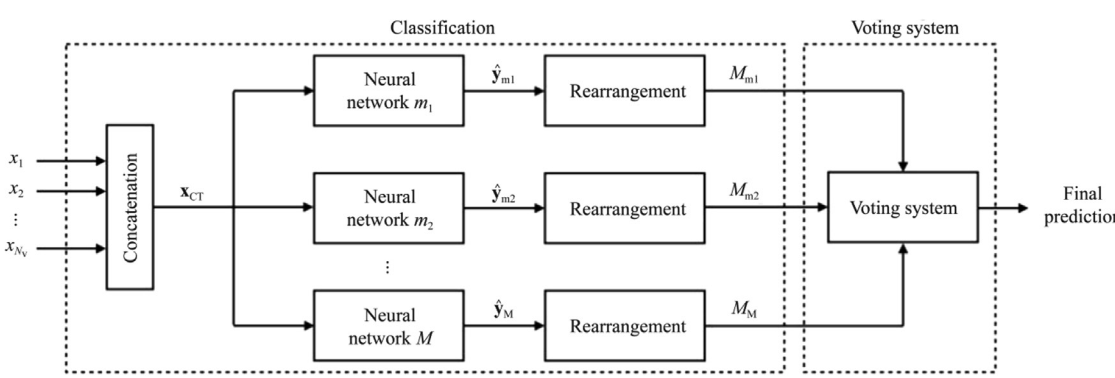


Fig. 9. Block diagram of the classification and voting stages of the algorithm.

descent algorithm [42]. This algorithm is a numerical optimization method that calculates the gradient of the cost function concerning the neural network's weights and takes small steps in the steepest direction until the optimal point is found. The backpropagation method uses the chain rule to decompose the gradient into several simpler derivatives based on known functions. This process is widely known and explained thoroughly in other sources [35].

The classification algorithm presented in this paper consists of a set of neural networks operating in parallel, where each is specialized in classifying one specific mineral. We found that using parallel one-vs-all neural networks achieved better classification results than using a single multi-class neural network. When the algorithm is trained with the feature vectors from rock group 1 and the labels created by the geologist, it uses four neural networks because there are four main classes of minerals. In contrast, when the algorithm is trained with the feature vectors from rock group 2 and the labels created by binarizing the rock's grade with the cut-off, it only uses one neural network that predicts whether each sub-image is ore or waste since the problem reduces to binary classification. All the neural networks used in this work have one input layer, two hidden layers, and one output layer. Using two hidden layers or more allows the algorithm to find arbitrarily complex decision boundaries, unlike using one hidden layer, limiting the algorithm to convex boundaries [35]. Each hidden layer comprises 200 neural units with the rectified linear unit (ReLU) activation function. The output layer of each neural network has only one neuron and uses the logistic activation function, which maps the networks' output to a probability between [0,1].

Each neural network of the classification stage has the same input, which is a matrix $\mathbf{X}_{\mathbf{CT}}$ with dimensions $N_v \times 36$ that contains the 24 color and 12 texture features of all N_v valid sub-images of the rock. Before being classified by the neural networks, the algorithm normalizes all columns of matrix $\mathbf{X}_{\mathbf{CT}}$ by computing their z-score, which has zero mean and a similar scale for all columns. Normalizing the feature matrix can reduce the number of iterations needed for the weight of the network to converge [43].

Once the input matrix is normalized, it is processed by the neural networks in parallel, and each one of them produces a vector of probabilities $\widehat{\mathbf{y}}_m$ with dimensions $N_v \times 1$. Each element of this vector is a number between [0,1], representing the probability that each valid sub-image belongs to the mineral class m. When the model is trained with the feature vectors from rock group 1 and the mineral labels identified by the geologist, the neural networks generate four probability vectors: $\widehat{\mathbf{y}}_{\mathrm{VE}}$, $\widehat{\mathbf{y}}_{\mathrm{BX}}$, $\widehat{\mathbf{y}}_{\mathrm{AN}}$, and $\widehat{\mathbf{y}}_{\mathrm{AA}}$. In contrast, when the model is trained with images from rock group 2 and the labels obtained by binarizing the rock's grade with the cut-off, it only produces one probability vector $\widehat{\mathbf{y}}_{\mathrm{ore}}$. By tracking the spatial location of the valid sub-images, the algorithm rearranges the probability vectors into probability maps, denoted by M_{VE}, M_{BX}, M_{AN}, M_{AA}, and M_{ore}. Fig. 10 shows the four probability maps of a breccia (BX) rock generated by the model when it is trained with the geologist's labels (Fig. 10 b-e) and the probability map generated by training the model with the labels obtained through chemical assays (Fig. 10f).

3.5. Voting system

After generating the probability maps, the classification stage uses a voting system that analyzes the probability of each sub-image in order to decide whether a rock is mineralized or not and, consequently, whether it should be sent to the processing plant or waste dump. The inputs to this stage are the probability maps, and the output is a single classification value, which takes a value of 1 if the algorithm decides that the rock is mineralized, or 0 if it decides that the rock is barren. The voting system corresponds to the fourth stage of the proposed algorithm, represented graphically in Fig. 5.

The voting system is composed of two main stages. In the first stage, the algorithm analyzes each sub-image and compares the score given by each probability map. The map with the highest score for a given sub-image determines its mineral class when the algorithm is trained using the geologist's labels. In contrast, when the algorithm is trained with the labels obtained by performing chemical assays, it determines whether the sub-image is mineralized or not by comparing the predicted probability with a binary threshold, whose value is determined experimentally in the training stage. After assigning a class to each sub-image, the algorithm produces a mineral distribution map, represented by M_{MD}. This process is presented graphically in Fig. 11.

The second stage of the voting system counts the number of sub-images that belong to each mineral class. The mineral class with the highest amount of sub-images determines the overall class of the rock. When trained with the geologist labels, if the rock as a whole is classified as either vein (VE) or breccia (BX), the algorithm assumes that the rock is mineralized and should be sent to the processing plant, and if the rock is either pure andesite (AN) or altered andesite (AA), it should be sent to the waste dump. This decision process is based on the fact that, on average, VE and BX have a grade above the cut-off, as explained in the sub-section 2.1.

4. Results and discussion

The proposed algorithm was first trained and tested using the images of static rocks, which correspond to rock groups one and

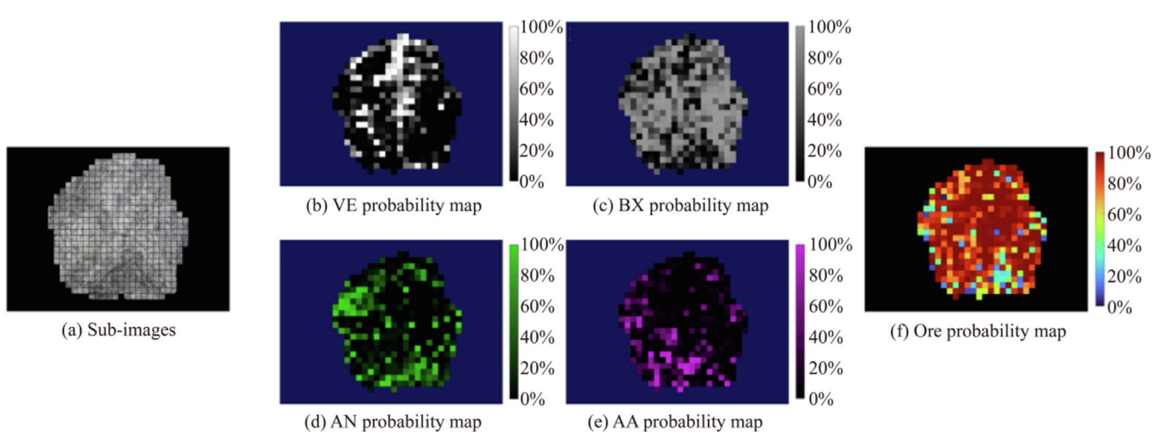


Fig. 10. Probability maps of a sample image.

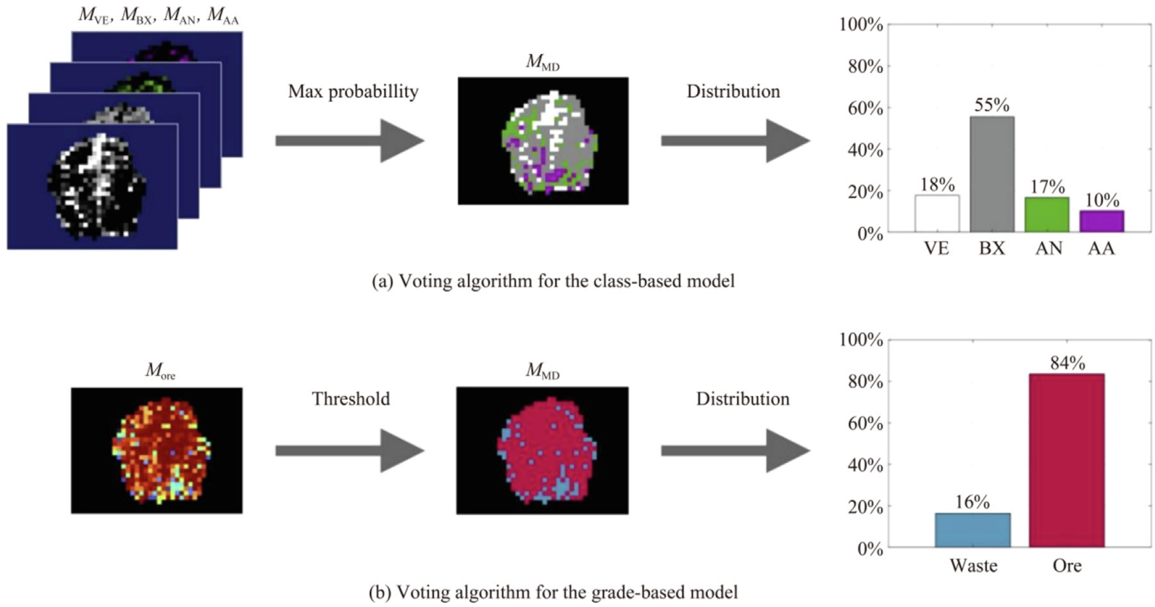


Fig. 11. Graphical representations of the voting system. Note: (a) shows the voting system used when the classification algorithm is trained with the geologist's labels, assigning the class with the highest probability to each sub-image, and (b) shows the voting system used when the classification algorithm is trained with the equivalent silver grade.

two. This first test was used as a proof of concept to validate the algorithm in an ideal setting, with dry rock samples highly representative of each mineral class and photographed using a high-resolution image. Then, after validating the algorithm in this ideal setting, more tests were performed using images of moving rocks corresponding to group three, which were photographed in a new test bench using an industrial camera. Although performed in a lab, these tests are more representative of the conditions that could be expected in an actual sorter since the rocks are photographed moving at a constant speed while still being wet. Sections 4.1–4.4 describe the training and testing process of the proof-of-concept stage, as well as the results that were obtained, while Section 4.5 shows the results with the moving rock images.

4.1. Training the classification models

In order to train and test the classification model in the proof-of-concept stage, the image dataset was divided into a training set and a test set. As explained in the Section 2.2.2, the geologist classified all 196 rocks from rock groups 1 and 2 into four mineral classes, while the 40 rocks from rock group 2 were also analyzed in a geochemical laboratory to obtain their gold and silver grades. Because of the small size of the rocks (mesh between 3/4″–5″), the algorithm assumes that the whole rock is homogeneous and assigns the same label to all sub-images from the same rock. This applies to both the geologist's labels as well as the grades obtained from the assays. All images obtained from rock group 1 were used to train the classification algorithm using the geologist labels,

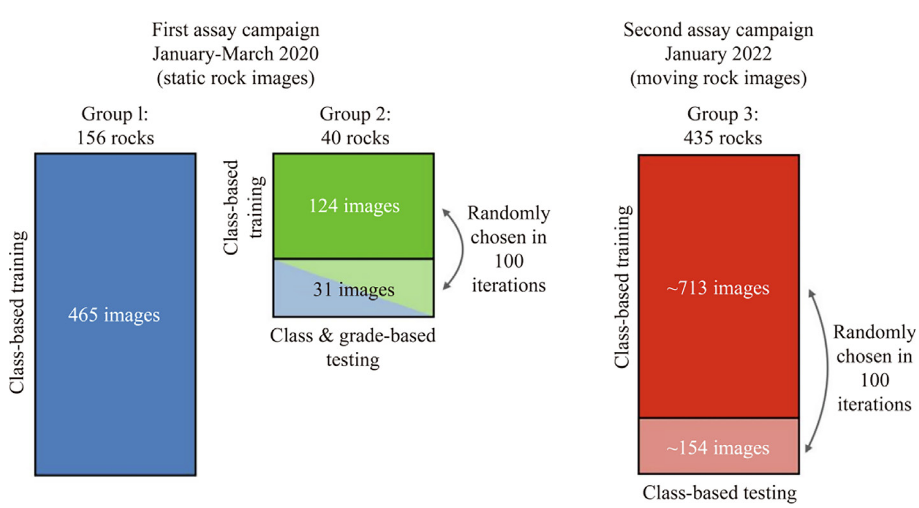
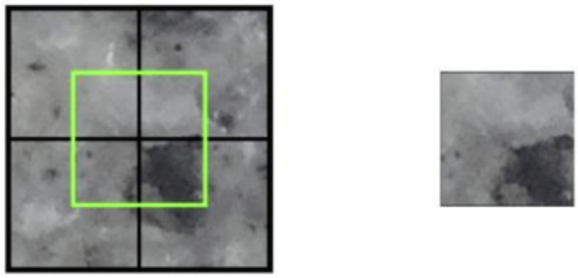


Fig. 12. Graphical representation of the separation between the training and test set for each of the three rock groups used in this paper. Note: 465 images from rock group 1 were exclusively used to train the algorithm using class-based labels in the proof-of-concept stage. 124 images (80%) from rock group 2 were used to train the algorithm using grade-based labels also in the proof-of-concept stage, and the remaining 31 images were used to test the algorithm using both types of labels. Because of the low amount of images in rock group 2, 100 train-test iterations were performed randomly, choosing the train and test sets. Finally, in rock group 3, on average, 713 images were used to train the algorithm with class-based labels using moving rocks, while the remaining 154 images were used for testing. Images from this group were also randomly assigned to the train and test sets in 100 iterations. The exact amount of images in the train and test sets is not defined since the train set is constructed using mass proportions found in the actual mine.



(a) Adjacent sub-images (b) New sub-image

Fig. 13. Graphical representation of the data augmentation process.

while only 80% of images from rock group 2 were used to train the algorithm using the labels obtained through chemical assays. The remaining 20% of images from rock group 2 were used to test the algorithm trained using both sets of labels. It is worth mentioning that because of the lower quantity of rocks in group 2 compared to group 1, the training and testing stages were performed on 100 iterations, randomly choosing the train and test rocks from group 2. Fig. 12 explains this process clearly by showing a graphical representation of how the rock groups were divided into train and test sets.

After choosing the training set rocks, their images were processed using the methods described in subsections 3.2 and 3.3 to create feature matrices and label vectors. Data augmentation using an offset when partitioning the images into sub-images was also used to double the number of sub-images in the training set. This method effectively creates one new sub-image from every-four adjacent sub-images, as illustrated in Fig. 13.

Because there was a different number of images for each mineral class, the feature matrices and label vectors were unbalanced. In classification problems, having an unbalanced dataset can result in poor performance for standard algorithms [44], such as neural networks. Thus, to avoid this problem, the feature matrices and label vectors were resampled to match the target class size so that the algorithms could be trained with the same number of positive and negative data points. This means that half of each feature matrix and label vector rows contained data from the target mineral, while the other half contained data from the rest of the minerals, as shown in Table 3.

After balancing the feature matrices and label vectors, they were further divided into a training set used to train the models and a validation set used to tune the model's parameters and evaluate the performance at the sub-image level. The neural networks were trained using the Adam (Adaptive Moment Estimation) algorithm. This algorithm is a stochastic optimization method based on the gradient descent algorithm that finds the optimal weights of the neural network by combining the main advantages of the AdaGrad and RMSProp algorithms [45]. Also, the proposed algorithm uses the dropout regularization method, in which neurons in the hidden layers have a fixed probability of being temporarily eliminated in one iteration of the training process. This method makes the neural network less dependent on any particular neuron, which means that the network's weights are kept small [46]. All the neural networks of the proposed models were trained for 200 epochs, recording the value of the cost function at each iteration to verify convergence.

4.2. Testing the classification models

The model's classification performance was first evaluated using test sub-images from the feature matrices and then using test images from rock group 2. The first test, which evaluates the algorithm's accuracy in classifying sub-images, is a good indicator of which minerals are harder to classify and provides insights into the performance of the individual neural networks. In contrast, the second test evaluates the performance of the algorithm in classifying complete images as ore or waste, which is the most important metric when comparing ore sorting algorithms.

4.2.1. Testing with sub-images

The sub-image classification performance of the model was evaluated using the validation set of the feature matrices. The metric used to evaluate the performance of the neural networks is the Matthews correlation coefficient (MCC), also known as the phi-coefficient, which is given by

$$\mathrm{MCC} = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{\mathrm{PP} \times \mathrm{P} \times \mathrm{N} \times \mathrm{PN}}} \tag{7}$$

where PP=TP+FP, P=TP+FN, N=TN+FP, PN=TN+FN, TP is the number of true positives; FP the false positives; FN the false negatives; and TN the true negatives. The MCC is a classification metric used to evaluate binary classifiers that produces a high score only if the prediction obtained good results in all of the four categories of the confusion matrix (true positives, false negatives, true negatives, and false positives), proportionally both to the size of positive elements and the size of negative elements in the dataset [47]. The MCC is a real decimal number between $+1$ and -1, where a value of $+1$ means that the model classified all samples correctly, 0 means that it did not find any relationship between inputs and outputs and is working as a random classifier, and a value of -1 means that the model classified all samples incorrectly. The classification performance of the neural networks was also evaluated using more common metrics, such as the true positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), and negative predictive value (NPV).

The results from the sub-image tests are presented in Table 4. The VE, BX, AN, and AA results correspond to the single test performed to evaluate the model's performance when trained with the geologist's labels. In contrast, the classification results associated with the ore versus waste classifier were obtained when the model was trained with the equivalent silver grades and is presented as the average obtained in 100 iterations. The proposed model achieved an excellent classification performance, quantified by the MCC, for the VE (0.938) and AN (0.902) classifiers and good performance for the BX (0.723) and AA (0.779) classifiers. VE and AN rocks are easier to classify because they have very different

Table 3
Number of sub-images per mineral class used to train the neural networks.

Classifier	Feature matrix	Data-points per class						Total
		VE	BX	AN	AA	Ore	Waste	
VE vs. other	X_{VE}	11251	3750	3750	3750			22501
BX vs. other	X_{BX}	1039	3118	1039	1039			6235
AN vs. other	X_{AN}	1613	1613	4839	1613			9678
AA vs. other	X_{AA}	896	896	896	2689			5377
Ore vs. waste	X_{ore}					11063	8304	19367

Table 4
Sub-image classification performance of the individual neural networks for the vein (VE), breccia (BX), pure andesite (AN), and altered andesite (AA) versus all the rest, and for ore (O) versus waste (W).

Model	TPR	TNR	PPV	NPV	MCC
VE	97.5%	96.0%	96.0%	97.5%	0.938
BX	86.5%	81.0%	79.0%	87.9%	0.723
AN	96.1%	93.6%	93.7%	96.0%	0.902
AA	85.6%	89.3%	89.5%	85.3%	0.779
O vs. W	90.5%	87.9%	89.8%	87.4%	0.778

Note: The neural networks are compared using TPR, TNR, PPV, NPV, and MCC.

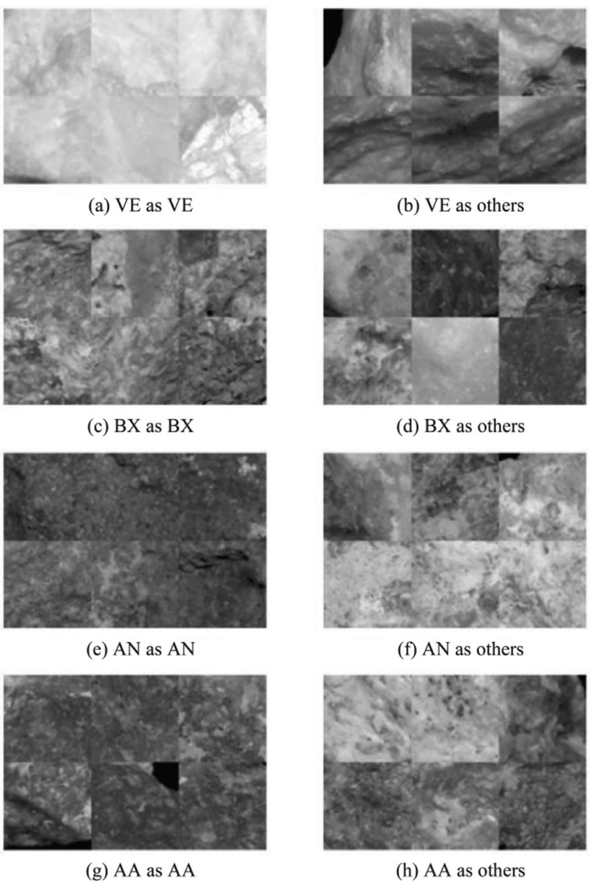


Fig. 14. Comparison between sub-images that were correctly classified (a, c, e, and g) and those that were incorrectly classified (b, d, f, and h) for each mineral type.

color characteristics from the rest of the minerals. VE is almost entirely white due to its high quartz content, while AN has a dark green to black color due to its high andesite content. The BX and AA minerals classifiers have lower accuracy than the others because they are visually very similar. Both minerals are gray with white sections. In the case of BX, the white sections are composed chiefly of quartz, while those of the AA mineral contain clay. By looking at the sub-images presented in Fig. 14, it can be clearly seen that most of the BX class sub-images that were misclassified are very similar to the AA class sub-images, with a few exceptions that resemble the VE and AN minerals. The same is true for the misclassified sub-images of the AA class, which resemble those of the BX class. The ore versus waste classifier also achieved a good classification performance (0.778), which was 0.055 points higher than the BX classifier but 0.001 points lower than the AA classifier.

4.2.2. Testing with full images

After training the neural networks and testing the classification performance using sub-images, the model trained with both sets of labels was evaluated using the full images. The tests were performed on 100 iterations, using 31 randomly chosen rock images from group 2 not used for training the models. Then, the test images were processed using the segmentation, partitioning, and feature extraction methods described in Section 3, producing a set of feature vectors for each image. The neural networks then classified the feature vectors to make a single prediction value for each image. An image was considered to be correctly classified if the prediction matched the grade of the rock. For example, if the prediction of one of the models was "ore" and the rock associated with the image had an equivalent silver grade above the cut-off, then the classification was considered to be correct. The same applied if the model's prediction was "waste" and the grade was below the cut-off. It is important to note that the same test images were used to evaluate both models in each iteration to compare the classification results. Similar to the tests using sub-images, the classification performance using complete images is quantified using the MCC.

Table 5 shows the classification results for the complete image tests. The model trained with the geologist's labels (NN-G) achieved a higher MCC than the model trained with the labels obtained through chemical assays (NN-A) in every test iteration, which on average, was 0.032 points higher. The models are compared using the true positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), negative predictive value (NPV), and Matthews correlation coefficient (MCC). Such metrics as TPR, TNR, PPV and NPV were also higher for the model trained with the geologist's labels, which proves that the model is the superior method for mineral classification for this particular mineral system.

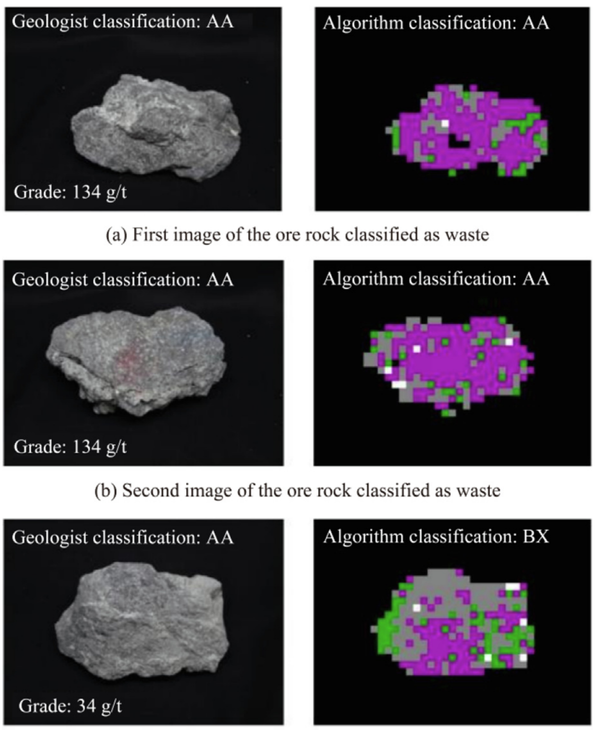
After performing the complete image tests, the images misclassified by the model trained with the geologist's labels were analyzed to determine the cause of the errors. Only 3 out of 155 images from group 2 were misclassified. The first two images shown in Fig. 15 come from the same rock, which the geologist and the algorithm probably classified as altered andesite (AA) because of its purple tones. However, after analyzing the rock using chemical assays, it was found that its equivalent silver grade was higher than the cut-off, which means that it should have been classified as ore by the geologist and the algorithm. This particular example is an outlier in the altered andesite category, and it is inferred that most of the gold and silver content is hosted on small quartz veins on the rock. The third misclassified image is a rock classified correctly by the geologist as an altered andesite but misclassified by the algorithm as breccia (BX). In this case, the algorithm classified the other image from the same rock correctly as waste.

4.3. Comparison with other classification algorithms

Previous papers have used other methods besides neural networks in their mineral classification algorithms. One of the most common methods is support vector machines (SVM), which are classification and regression algorithms that aim to find the deci-

Table 5
Image classification performance of the proposed method compared to the support vector machines (SVM) and the VGG-19 convolutional neural network models using all images from the second rock group.

Model	TPR	TNR	PPV	NPV	MCC
NN-G	97.1%	98.8%	98.5%	97.7%	0.961
NN-A	96.1%	96.9%	95.9%	96.9%	0.929
SVM	92.8%	97.7%	97.0%	94.4%	0.909
VGG-19	100.0%	95.4%	94.2%	100.0%	0.948



(a) First image of the ore rock classified as waste

(b) Second image of the ore rock classified as waste

(c) Image of the waste rock classified as ore

Fig. 15. Images from the group 2 rocks that the model misclassified.

sion boundary that maximizes the separation between two classes, in the case of classification, or to find the curve that best fits the trend of a data set, in the case of regression. For example, [11,14,15] used SVM to classify rocks from nickel mineral systems successfully.

Another method that has been used extensively for solving many challenges related to image classification is convolutional neural networks (CNN). CNNs are a class of neural networks specialized in processing data arranged in a grid, such as digital images. These classifiers comprise at least one convolutional layer containing filter banks whose coefficients are learned in a supervised learning stage [48]. CNN models learn to identify arbitrarily complex structures without designing a prior feature extraction system [48,49]. Due to their good classification potential and ease of use, CNNs have become one of the dominant algorithms in several areas of computer vision. For example, in ore sorting applications, CNNs were successfully used by [17,50] to classify different coal ore classes.

The classification algorithm proposed in this paper was also compared with SVM and CNN-based algorithms to confirm that our approach is the best choice for the ore sorting problem. The SVM-based algorithm used the same feature extraction method as the proposed model, but instead of neural networks, it used a regression SVM with a Gaussian kernel to predict the class of the sub-images. The CNN-based algorithm, unlike the SVM, did not use the color and texture features of the rock but instead used the raw sub-images as inputs. We used a modified version of the VGG-19 model [51], which is a commonly used CNN for many classification problems. This network consists of 16 convolution layers, 5 max pool layers, 3 fully connected layers, and 1 SoftMax layer. The modifications are the following: (1) the input size of the image was reduced from $224\times224\times3$ to $64\times64\times3$ since the sub-images have the latter dimensions, (2) the fully connected layer size was reduced from 4096 to 200 neurons, because that was the number of neurons used by the proposed model and we only want to compare the feature extraction backbone, and (3) the output layer was reduced from 1000 to 4 neurons because we only want to classify 4 classes of minerals. The results of the two new algorithms, presented in Table 5, show that the model using neural networks and feature engineering achieved a better classification performance, evidence that the proposed algorithm is the best alternative for this specific mineral classification problem.

4.4. Processing time

The proposed method was optimized to classify rocks in real time. The processing time of the algorithm was measured from the moment after the image was loaded from memory until the moment the algorithm finished counting the number of sub-images from each class. The time the computer takes to load the image from memory was not considered because, in a real-time implementation, the camera transfers data directly to the memory buffer [52]. Fig. 16 shows the distribution of processing times for the test set images, with a mean of 19.2 ms and a standard deviation of 7.9 ms. The maximum processing time was 44 ms and belonged to a AA class rock, the largest rock in the test set and the entire dataset. This indicates that the processing time is within the limit proposed by the mining company of 70 ms. There is a surplus of 25 ms, which another neural network could use to classify other classes of minerals if necessary.

On the other hand, Fig. 17 shows the correlation between the number of valid sub-images in an image and the time it takes to classify them. This correlation is linear and has a coefficient of

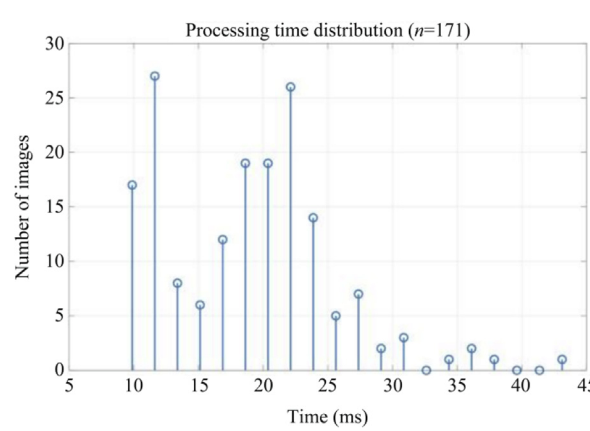


Fig. 16. Distribution of the processing times of the images from rock group 2.

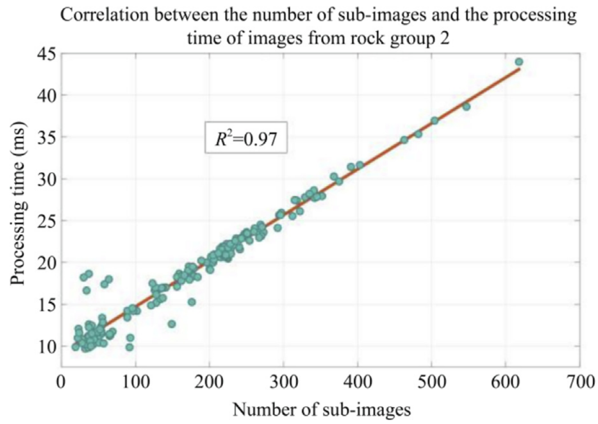


Fig. 17. Correlation between the number of valid sub-images in each rock and their processing time, represented by a linear regression with a coefficient of determination (R^2).

determination R^2=0.97, which indicates a strong correlation between the number of sub-images and the processing time. To calculate the maximum number of sub-images that the algorithm can classify, we used the line equation found by the linear regression method, which is given by

$$t = 0.0543n + 9.7002 \tag{8}$$

where t is the processing time in milliseconds; and n the number of sub-images. Solving Eq. (8) using the maximum possible number of sub-images in a given image (n=1064) results in a processing time of just 67.5 ms, which is less than the limit of 70 ms needed for real-time ore sorting. Thus, the algorithm can classify an entire image of 1836×2448 pixels using four neural networks. The maximum processing time was calculated using a theoretical conveyor belt speed of 3 m/s. However, this speed could change depending on the company's mining plan.

It is worth mentioning that the voting strategies presented in this paper could be further optimized to improve classification accuracy. For example, the voting system could automatically accept material when it detects a small amount of a specific mineral instead of the current strategy, where the rock is accepted only if the predominant mineral is ore. As another example, the voting system could be further improved by using a second threshold that determines the minimum amount of sub-images classified as ore to classify the rock as a whole as ore. Also, both voting systems can be tuned to accept or reject more material, to maximize profits according to a financial model.

4.5. Performance test with moving rocks

After doing the tests with static rocks to choose the best algorithm, additional tests were performed with moving rocks on a test bench. The test bench consists of a conveyor belt and an MDF structure with the camera and lighting system, as shown on Fig. 18. The camera is a Basler daA2500-14uc [53], located perpendicularly at a distance of 500 mm from the conveyor belt. It has a resolution of 1342×1960 pixels and sends raw image data to the PC through a USB cable. By using this resolution and distance to the conveyor belt, the images had a pixel density of 6.25 px/mm, representing a scene with dimensions of 215 mm × 317 mm. The images were illuminated using 2 m of 16.2 W/m LED strips, with a color temperature of 4000 K. The conveyor belt operated at a speed of 1 m/s, which is only a third of commercial sorters. However, commercial sorters use line-scan cameras and powerful LED bars, which are able to reduce the sensor exposure time to obtain sharper images.

Using the test bench, we acquired 867 images from 435 rocks (54.4 kg) since, in most cases, they were photographed from two different sides. All rocks were extracted from the same vein as the tests with static rocks but approximately-two years later. Additionally, the rocks were washed before performing the tests in order to remove dust and mud from their surface and then photographed while they were still wet. After creating the image database, the 460 rocks were dried and analyzed in a geochemical laboratory using aqua regia and fire assays to estimate their gold and silver grades. The grade statistics of this group of rocks are shown on Table 1 and correspond to the assay campaign performed on January 2022.

Classification tests were performed using the selected algorithm on the static tests, which consists of extracting color and texture features from sub-images and then classifying them with neural networks trained with geologist's labels. The only modification that was made to the algorithm was that sub-images had a side length of 32 pixels instead of 64 because of the lower camera resolution and pixel density. The algorithm was trained and tested on 100 iterations, choosing the images for each set randomly. However, unlike the tests with static rocks, the test set with moving

(a) Outside view (b) Inside view

Fig. 18. Test bench used to capture images of rocks moving at speeds of 1 m/s.Note: (a) Outside view of the test bench: a treadmill is used as a conveyor belt, and the camera, located inside the MDF box, is connected to a laptop with the image acquisition program. (b) Inside view of the test bench: 2 m of 16.2 W/m LED strips are used to illuminate the rocks moving through the conveyor belt, while a Basler daA2500-14uc camera is used to acquire color images.

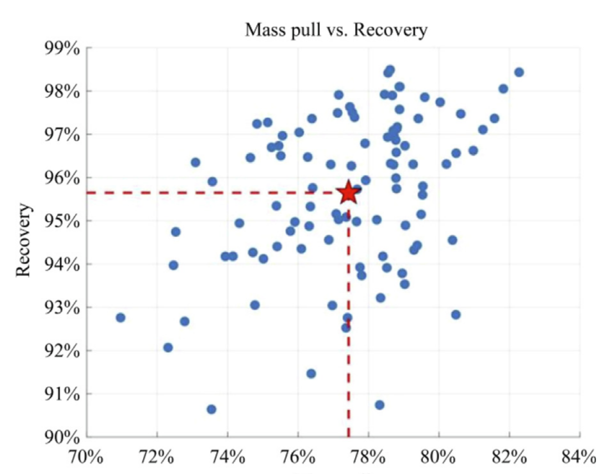


Fig. 19. Recovery and mass pull calculated on each of the 100 test iterations, represented by blue dots. The red star represents the average value of the 100 iterations, which is a recovery of 95.6% and a mass pull of 77.4%. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rocks was formed by choosing mineral proportions that closely resemble those found in the selected mine. In particular, the test set proportions are 39% VE, 39% BX, 19% AN, and 3% AA. Although these proportions can vary in different veins and periods of time, the values represent averages and were validated by geologists of the mining company.

After performing the 100 test iterations, we found that the average MCC of the moving rock tests was 0.901, which is lower than the average MCC of 0.961 obtained using static rocks. The decrease in performance might be explained by one or several of the following factors: lower grades, non-ideal rock samples, wet surfaces, a lower camera resolution, and the blur produced by the motion of the rocks. Motion blur is detrimental to the algorithm because it distorts the texture of the rock. Although the first three factors cannot be controlled since they depend exclusively on the characteristics of the mineral, the last two factors might be controlled in future work by using better sensors. Particularly, using a line-scan camera and a brighter light source is key to reducing sensor exposure time and, therefore, motion blur.

The algorithm was also evaluated using two additional metrics commonly used in ore sorting: mineral recovery and mass pull (also called yield). The mass pull is the percentage of mass that is classified as "ore" by the algorithm, while the recovery is the percentage of the mass of gold and silver that can be recovered from the rocks classified as "ore". These two metrics were not calculated for the static rock tests because they need actual mineral proportions to be interpreted correctly. In the moving rock tests, the recovery and mass pull were calculated on each of the 100 iterations, as shown on Fig. 19. We found that the average recovery was 95.6%, while the average mass pull was 77.4%. This means that if the sorting algorithm was implemented in a real mine, by processing only 77.4% of the mineral, 95.6% of the gold and silver could potentially be recovered. Finally, we calculated the weighted-average grade of the rocks in the input and output streams on each iteration and found that the proposed algorithm increases the equivalent silver grade from 151 to 186 g/t, which is equivalent to a grade upgrade of 23.6%. All of these metrics could be used in the future to estimate the potential economic benefit of implementing the proposed algorithm in a real mining setting.

5. Conclusions

This paper presented a novel ore sorting algorithm capable of classifying rock particles in real-time using color and texture analysis. The algorithm was trained with two different data sets. The first dataset consisted of rocks that a geologist manually labeled according to their mineral content, while the second one included rocks analyzed in a geochemical laboratory to determine their grade. The algorithm was tested with images of gold and silver-bearing rocks extracted from an underground mine in the Peruvian Andes. The ore sorting problem was particularly challenging because of the color and texture similarities between high and low-grade rocks. The main findings are:

(1) The highest performance was obtained when training the algorithm with mineral class labels identified by a geologist, with an average MCC of 0.961 points. In contrast, when the algorithm is trained directly using mineral grades as labels, its performance is significantly worse, with an average MCC of 0.929 points.
(2) The algorithm using color and texture analysis in the feature extraction stage and neural networks in the classification stage outperforms other algorithms trained with the same images. In particular, the MCC of the proposed method is 0.052 points higher than the one using SVMs instead of neural networks and 0.013 points higher than the one using a VGG-19 backbone for feature extraction.
(3) The algorithm is capable of classifying rock particles screened with a 3/4"−5″ mesh with an average and maximum processing time of 19.2 ms (52.1 fps) and 44 ms (22.7 fps), respectively.
(4) Testing the algorithm with an additional 54.4 kg of non-ideal wet rock particles moving at a speed of 1 m/s on a conveyor belt, with a lower-resolution industrial USB camera, and with real mineral proportions produced an MCC of 0.901 points, which are still highly desirable results for ore sorting. This performance translates to gold and silver recovery of 95.6% (or grade upgrade of 23.6%) and a mass pull of 77.4%.

The results of our algorithm indicate that it could be implemented in a pilot plant to perform real-time ore sorting. Also, in the future, the proposed method will be combined with hyperspectral analysis using multi-modal learning to improve its classification performance.

References

[1] Priester M, Ericsson M, Dolega P, Löf O. Mineral grades: An important indicator for environmental impact of mineral exploitation. Miner Econ 2019;32(1):49–73.
[2] Cho K, Kim H, Myung E, Purev O, Choi N, Park C. Recovery of gold from the refractory gold concentrate using microwave assisted leaching. Metals 2020;10(5):571.
[3] Elliott R, Barati M. A review of the beneficiation of low-grade manganese ores by magnetic separation. Can Metall Q 2020;59(1):1–16.
[4] Ulrich S, Trench A, Hagemann S. Grade-cost relationships within Australian underground gold mines—A 2014–2017 empirical study and potential value implications. Resour Policy 2019;61:29–48.
[5] La Brooy SR, Linge HG, Walker GS. Review of gold extraction from ores. Miner Eng 1994;7(10):1213–41.
[6] Lessard J, de Bakker J, McHugh L. Development of ore sorting and its impact on mineral processing economics. Miner Eng 2014;65:88–97.
[7] Lessard J, Sweetser W, Bartram K, Figueroa J, McHugh L. Bridging the gap: Understanding the economic impact of ore sorting on a mineral processing circuit. Miner Eng 2016;91:92–9.
[8] Robben C, Wotruba H. Sensor-based ore sorting technology in mining—Past, present and future. Minerals 2019;9(9):523.
[9] Simmons SF, White NC, John DA. Geological characteristics of epithermal precious and base metal deposits. Society of Economic Geologists (100th anniversary vol.); 2005.
[10] Gülcan E, Gülsoy ÖY. Performance evaluation of optical sorting in mineral processing—A case study with quartz, magnesite, hematite, lignite, copper and gold ores. Int J Miner Process 2017;169:129–41.
[11] Tessier J, Duchesne C, Bartolacci G. A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts. Miner Eng 2007;20(12):1129–44.
[12] Singh V, Rao SM. Application of image processing and radial basis neural network techniques for ore sorting and ore classification. Miner Eng 2005;18(15):1412–20.
[13] McCoy JT, Auret L. Machine learning applications in minerals processing: A review. Miner Eng 2019;132:95–109.
[14] Perez CA, Estévez PA, Vera PA, Castillo LE, Aravena CM, Schulz DA, Medina LE. Ore grade estimation by feature selection and voting using boundary detection in digital image analysis. Int J Miner Process 2011;101(1–4):28–36.
[15] Perez CA, Saravia JA, Navarro CF, Schulz DA, Aravena CM, Galdames FJ. Rock lithological classification using multi-scale Gabor features from sub-images, and voting with rock contour information. Int J Miner Process 2015;144:56–64.
[16] Chen JY, Huang HW, Cohn AG, Zhang DM, Zhou ML. Machine learning-based classification of rock discontinuity trace: SMOTE oversampling integrated with GBT ensemble learning. Int J Min Sci Technol 2022;32(2):309–22.
[17] Liu Y, Zhang ZL. Liu X, Wang L, Xia XH. Ore image classification based on small deep learning model: Evaluation and optimization of model depth, model structure and data size. Miner Eng 2021;172:107020.
[18] Bai FY, Fan MQ, Yang HL, Dong LP. Fast recognition using convolutional neural network for the coal particle density range based on images captured under multiple light sources. Int J Min Sci Technol 2021;31(6):1053–61.
[19] Andrearczyk V, Whelan PF. Using filter banks in convolutional neural networks for texture classification. Pattern Recognit Lett 2016;84:63–9.
[20] Fujieda S, Takayama K, Hachisuka T. Wavelet convolutional neural networks for texture classification. 2017:arXiv: 1707.07394.

[21] Wang QC, Zheng YJ, Yang GP, Jin WD, Chen XJ, Yin YL. Multiscale rotation-invariant convolutional neural networks for lung texture classification. IEEE J Biomed Health Inform 2018;22(1):184–95.
[22] Corbett G. Epithermal gold for explorationists. AIG Journal 2002;67:1–8.
[23] Sillitoe RH. Styles of high-sulphidation gold, silver and copper mineralisation in porphyry and epithermal environments. In: Proceedings of the Australasian Institute of Mining and Metallurgy. Parkville, Vic.: Australasian Institute of Mining and Metallurgy; 2000. p. 19–34.
[24] Gibson HL, Watkinson DH, Comba CDA. Silicification; hydrothermal alteration in an Archean geothermal system within the Amulet Rhyolite Formation, Noranda. Quebec Econ Geol 1983;78(5):954–71.
[25] Sony, Digital Still Camera-Instruction Manual-DSC-HX90V/DSC-HX90/DSC-HX80/DSC-WX500 (2015). URL.
[26] GTI Graphic Technology, Inc., PDV-2e/M Multi-Source Portable Desktop Viewers (Dec. 2017). URL.
[27] Rawat W, Wang ZH. Deep convolutional neural networks for image classification: A comprehensive review. Neural Comput 2017;29(9):2352–449.
[28] Loizou CP, Murray V, Pattichis MS, Pantziaris M, Nicolaides AN, Pattichis CS. Despeckle filtering for multiscale amplitude-modulation frequency-modulation (AM-FM) texture analysis of ultrasound images of the intima-media complex. Int J Biomed Imaging 2014;2014:1–13.
[29] Ibraheem NA, Hasan M, Khan R, Mishra PK. Understanding color models: A review. ARPN J Sci Technol 2012;2(3):265–75.
[30] Yu HL, MacGregor JF. Multivariate image analysis and regression for prediction of coating content and distribution in the production of snack foods. Chemom Intell Lab Syst 2003;67(2):125–44.
[31] Edward JJ. A User's Guide to Principal Components, Vol. 587. John Wiley & Sons; 2005.
[32] Jolliffe I. Principal Component Analysis. Encyclopedia of Statistics in Behavioral Science. 2005.
[33] Ilin A, Raiko T. Practical approaches to principal component analysis in the presence of missing values. J Mach Learn Res 2010;11:1957–2000.
[34] Jolliffe IT, Cadima J. Principal component analysis: A review and recent developments. Philos Trans Ser A Math Phys Eng Sci 2016;374 (2065):20150202.
[35] Gonzalez RC, Woods RE. Digital Image Processing. (2nd Edition). Pearson; 2017.
[36] Davies ER. Introduction to texture Analysis. In: Handbook of Texture Analysis. Imperial College Press; 2008. p.1-32.
[37] Strang G. Wavelets and dilation equations: A brief introduction. SIAM Rev 1989;31(4):614–27.
[38] Moulin P. Multiscale image decomposition and wavelets. In: The Essential Guide to Image Processing. Academic Press; 2009.
[39] Sun ZH, Miller R, Bebis G, Dimeo D. A real-time pre-crash vehicle detection system. In: Proceddings of the sixth IEEE workshop on applications of computer vision. Orlando: IEEE; 2003.
[40] Laine A, Fan J. Texture classification by wavelet packet signatures. IEEE Trans Pattern Anal Mach Intell 1993;15(11):1186–91.
[41] Bishop CM. Neural networks. In: Pattern recognition and machine learning. Springer; 2006. p. 225–90.
[42] Murphy KP. Machine Learning: A Probabilistic Perspective. Cambridge, MA: The MIT Press; 2012.
[43] LeCun Y, Bottou L, Orr GB, Müller KR. Efficient BackProp. Neural Networks: Tricks of the Trade. Berlin, Heidelberg: Springer, 2012:9–48.
[44] Zhang J, Mani I. KNN approach to unbalanced data distributions: A case study involving information extraction. Proc ICML'2003 Work Learn From Imbalanced Datasets 2003: p.1–7.
[45] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014:arXiv: 1412.6980.
[46] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15(1):1929–58.
[47] Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 2020;21(1):6.
[48] Rosebrock A. Deep Learning for Computer Vision with Python (1st Ed., Vol.1). Starter Bundle, PyImageSearch, 2017.
[49] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: An overview and application in radiology. Insights Imaging 2018;9(4):611–29.
[50] Liu Y, Zhang ZL, Liu X, Wang L, Xia XH. Deep learning-based image classification for online multi-coal and multi-class sorting. Comput Geosci 2021;157:104922.
[51] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014:arXiv: 1409.1556.
[52] Patterson DA, Hennessy JL. Computer Organization and Design (ARM Edition): The Hardware-Software Interface. Morgan Kaufmann; 2016.
[53] Basler AG. Basler product documentation: daa2500-14uc; 2022.