Araştırma Makalesi

Applıcatıon Of Support Vector Machıne To Separate Hydrothermal Alteratıon Zones Usıng The Lıner Sensor

Shezad SAPHIRI

Kabul State University

*Sorumlu yazar: ssaphırı@gmail.com

Abstract

In this research, an attempt is made to investigate the application of support vector machine algorithm in separating hydrothermal alterations of porphyry copper deposits using a primary sensor. A total of 2204 pixels were selected from mineralized areas to teach this algorithm. The 4, 6, 7, and 8 bands of the primer were selected to detect filter and argillic changes, and the 9 band visible and near infrared range was selected to detect proplitic changes as input to this algorithm. The fusion matrix was examined to evaluate the classification error. The results of the fusion matrix show that the classification error for filite and argillic zones is relatively high and it is not possible to easily separate the two regions while classifying the proplitic change well. The study also identified a new error function called blind error, which uses a fusion matrix to calculate the ratio of unclassified algorithm pixels. Based on the amount of this error, the backup vector machine does not classify 6.73% of the algorithm pixels. Also, the support vector machine algorithm has an overall classification accuracy of 66.06% and a kappa coefficient 0.6522.

Keywords: Argylic Alteration, Phyllic Alteration, Propylitic Alteration, Porphyry Copper

INTRODUCTION

In recent years, with the rapid growth of remote sensing technology, the use of satellite data is one of the most common methods of mine exploration. Remote sensing science uses the unique spectral behavior of minerals to identify and explore areas of mineral potential. In this method, the reflection and absorption index curves are very important as they provide information about the spectral behavior of objects of different wavelengths. In fact, each rock type has its own spectral interaction that can be identified by the same spectral reaction and distinguished from other rocks. This sensor has 92 bands that allow it to detect certain minerals in a specific and suitable range of electromagnetic waves. The primary captures data in three spectral ranges, including visible and near infrared (VNIR), short wave (SWIR) and thermal infrared (TIR). The spatial resolution of the near infrared bands is 15 meters, the spatial resolution of the short-wave infrared bands is 30 meters, and the spatial resolution of the thermal infrared images is 09 meters (1), (2). Bandwidth analysis, principal component analysis, shell method, and composite images are the most important traditional satellite image processing methods for determining the hydrothermal change regions in the region that were widely used in previous research. Past research has shown that these classical methods fail to separate regions of change from vegetation or their ability to discriminate hydrothermal change regions is poor. The reason for this can be thought of as the spectral fingerprints of the alteration minerals overlap each other and with the vegetation. Alteration minerals often have a hydroxyl bond in their structure, such as clays consisting mostly of minerals containing Mg-OH and Al-OH. These chemical bonds are strong absorption from electromagnetic radiation in the wavelength range of 2 / 1-2 / 4 μm and 1 / 7-1 / 6. It shows high reflection in the μm range. Therefore, it is difficult to distinguish between alteration minerals due to their participation in spectral behavior. In addition, the presence of water in plant tissues causes spectral overlap with hydroxyl minerals, which makes their separation difficult. Past research confirms that conventional image processing methods such as band ratio, principal component analysis, shell and composite images may not be significantly successful in distinguishing different shifting regions from each other. However, new methods from various sciences can be used to solve this problem. The standby vector machine method is a classification algorithm widely used in satellite image processing due to its high accuracy. Guderzi Mihr et al. (1391) compared the effectiveness of the support vector machine method to classify 6 lithological units in the region with neural network and maximum probability classification methods, and examined the three functions used in this method, namely the radial base function, e and the line, in terms of classification power. They found that it achieved the highest accuracy compared to the other two methods. Kasımı et al. (1390) evaluated the data on two porphyry copper deposits in Aliabad and Berberi Valliy by supporting vector classification method, linear differentiation analysis and nonlinear differentiation analysis. They concluded that the support vector machine method performed better than the other two methods with an accuracy of 80% on the Aliabad bed. In the Barberry Valley deposit, the support vector machine method was more successful than the other two methods with 90% accuracy. Huang et al. (2008) used the support

vector machine method to classify the forest and evaluated its accuracy according to the effect of the slope and direction of the ground, the height and azimuth of the sun, and the relative position of the trees. After topographic corrections, they found that the classification accuracy increased by 3/6% across all grades. Oman et al. (2008) used the support vector method and the maximum probability method to classify lithological units and their results show that the support vector machine method is more accurate than the maximum probability method. This research attempts to use the support vector machine algorithm to identify and isolate propilitic, argillic, and filter change regions and compare the capability of this algorithm with traditional satellite image analysis methods. Since the porphyry copper deposits known in the study area include turbidity, Godcollovari, Ijo, Serno and Turquoise wells, the best classification method is expected to detect the change in the specified mineral regions.

Geology of the Area and Data Used

The study area is located in the southern part of the volcanic-sedimentary belt of Central Iran. The study area is the origin of important porphyry deposits such as turbidity, Godkolvari, Ijo, Serno and Çah-ı Firuze and is located southwest of Kerman. The Cretaceous color mix is the oldest and the Quaternary and alluvial gravel deposits are among the youngest outcrops in the study area. Cretaceous sediments are mainly composed of flysch rocks. Sedimentary rocks The sedimentary volcano complex is usually sandstone and light limestone. Volcanic rocks are occupied by Miocene Oligocene intrusive rocks, granodiorite, quartz diorite, monzonite, tonalite and granite. Volcanic rocks have been modified in many cases due to the

unprincipled proximity of these intrusive masses. Most intrusive and volcanic rocks have undergone in situ hydrothermal alteration and mineralization. Argillic, sericite and propolytic alterations are among the most common hydrothermal alterations in the region. In the southeast part of the area, the Miocene Oligocene sediments are metamorphically covered by Eocene formations. The thickness of these series is 299 meters. Neogene sediments are separated by a metamorphism between the lowest and highest members. The lower limb is mainly composed of medium to fine grained sandstones with gypsum veins and halite interference. The upper limb contains deposits that show weak and incorrect classification and do not have sufficient strength. These sediments are conglomerate and sandstone in the lower parts and rock fragments and dacite abrasive fragments of round veins in the upper parts, [3], [4], [4], [5], [6], [7], [9], [10] .

In this study, three primary sensor images at 1B level were used. 98 Two images were taken on April 2000, and the other was taken on June 15, 2007. Images were first radiographically corrected by Crossstock and then IARR and placed side by side in a mosaic. Figure 2 shows a satellite image of an area showing known mineral fields. 3 band visible range and 6 band near infrared range were used as inputs of this study.

Backup Vector Machine Algorithm

The standby vector machine is one of the most well-known methods of data classification. Backup machines are often used for problems where there are two sets of data. Under this method, the two pages are spaced enough to be on the boundary of the two data classes. The main purpose of this algorithm is to find the boundary between these two pages, which creates the greatest separation

between the two data sets (Figure 3). In cases where the data cannot be separated linearly, the RBF can be generalized to a wider area using functions such as linear, polynomial, and sigmoid, and then separated linearly.

Consider the data set $\delta =$} X1,...., Xn} and (Xi ε X = Rp) consisting of N data. Each of these data has a corresponding (yi) tag indicating which class the data belongs to. (Y1, ..., yn ε Y). Support vector machines are a group of kernel-based methods that learn the f: X → Y function during the training phase using a tagged training dataset (δ). Later in the walktest phase, they use this function to predict the new data tag that enters the system. In this section, we only consider the case where each model belongs to one of classes 9 and 2, we call these two classes +9 and -9 ((1 + 1-}) ε y). This simple problem is called binary classification. In this case, the backup vector machine tries to separate the two data classes (w ε Rp, b ε r) using a linear function of the form f (x) = wT x + b. One such function labels data x ε X, f (x)> 0 +1 and data f (x) <0 -1 [11]. So the problem with the backup vector machine

is to learn such a function (f) on the observation set δ

A linear decision boundary can generally be written as:

x is a point on the decision boundary and w is an n-dimensional vector perpendicular to the decision boundary. b / || w || The distance from the beginning to the decision boundary and the w.x represents the internal product of the vectors x and w. Since a constant is always equal to the product of a constant on both sides, the following conditions apply to unique definitions of x and w:

If Xi is a substitute vector → yi (w.xi) = 1

If Xi is not a replacement vector → yi (w.xi)> 1

The first step in calculating the optimal decision limit is to find the closest teaching examples of the two classes. In the next step, the distance between these points is calculated in the direction perpendicular to the boundaries that completely separate the two classes. The limit for optimal decision making is the limit with the maximum optimal margin. The optimal decision boundary is calculated by solving the relation (3) optimization problem:

$$\max_{w,\ b}\ \min_{i=1,...,L}[yi\frac{w.x++b}{|w|}]$$

(3)

Solving the optimization problem is a difficult task. To simplify, Lagrange's indefinite coefficient method

(4) $f(x, w, b) = sgn(w.x + b)$

has been used. The last classifier is obtained from Equation (4).

The algorithm above gives the linear boundary of two completely separate classes, but it is always an error to separate the classes by the linear decision boundary in case the classes overlap. To solve this problem, the data can first be transferred from the original

space with Rn using a nonlinear φ-transform to the space Rm with larger dimensions that have less overlap with each other in the new class space. Then, in the new space, using the previous equations and replacing xi with a kernel function xi (φ), the optimal decision

boundary is calculated. A core function is actually a function in the first space that is equal to the internal product of two vectors in the feature space. In this study, radial base function was used to classify porphyry copper.

Mapping Swap Regions

This work uses the support vector machine method to map hydrothermal changes. This method is one of the most practical methods of heuristic data modeling based on learning principles. The main features of this method are higher power and greater accuracy in using less instructive samples. In this algorithm, deposition is divided into two educational and experimental areas. To specify training data, you must first select areas of the image with uniform pixels, then classify the data by crossing all bands and a linear boundary using an algorithm. In this way, a set of training points with a minimum distance from the decision boundary are considered support vectors. The support vector machine method for detecting shifting zones requires training grounds for learning. These areas are marked as red, green and blue polygons in Figure 7. A total of 2204 pixels were selected to promote the training areas; of these, 516 pixels belong to argylic alteration zone, 1278 pixel phyllite change zone and 500 pixel propolytic alteration zone. The pixels of the training grounds were selected from all mineral fields of Kadar, Godkolvari, Ijo, Serno and Çah-i Firuze, showing Figure 7 of the training pixels of the different mineral fields.

Mapping of Filit and Argillic Alteration Zone

According to the field and laboratory studies conducted by Majdifar et al. (2013) and Hunermend et al. (2011), 1278 pixels from the mineralized areas of the phyllic zone and 516 pixels from the argillaceous zone were selected as the training area. The locations of these pixels in green and red for filigree and agile regions, respectively, are shown in Figure 7. The substitute vector machine algorithm was used to define the filamentous and argillic alteration region using the radial basis function. In order to learn the support vector machine better, 8, 4, 6, 7 bands were chosen as the input of this algorithm known as effective bands in spectral fingerprints of sericite and kaolinite minerals. Figure 8 shows the main structure of the input data for the filtrate and clay regions.

(Red: argillism, green: filigree, blue: propolitic)
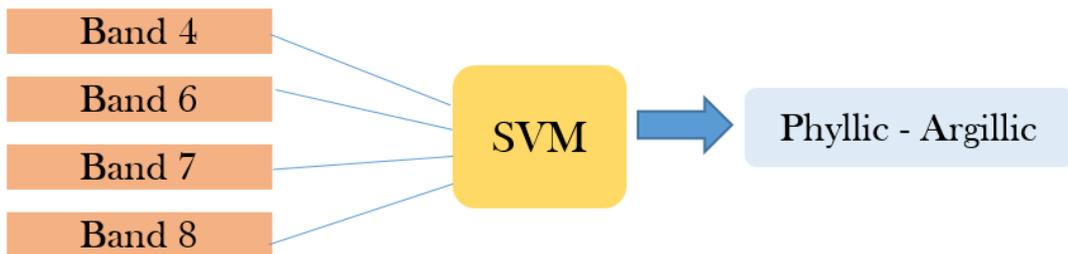
Figure 3. Educational areas of different displacement zones

Figures 1 and 10 show the output of the support vector machine algorithm for filtrate and clayey regions. According to the research of Hunermend et al. (2011), Felicity change is dominant in Ijo, Serno, Çah-i Firuze and Kadar reserves. The backup vector machine was able to detect the distribution of filite variation in fuzzy, eju, serno, and turquoise well beds. In addition, argillic alteration was observed in turbid, serno and well ovary reserves, which were well detected by the support vector machine in fuzzy and serno areas in terrestrial studies. From the results it is quite clear that the backup vector machine cannot be zoned. Distinguish between phyllic and argillic alterations, as they do not show argillic alteration in Goodcolvari and show an almost similar classification to the phyllic region. The reason for this can be considered in the spectral similarity of kaolinite and sericite minerals.

Mapping the Propolytic Alteration Zone

In order to detect the propylitic region with the support vector machine, the radial base function was chosen as the basic function of the algorithm and 500 pixels was chosen as the training area. The location of the propolytic zone training pixels is shown in blue in Figure 7. In this study, 9 visible and infrared bands near the liner sensor were used as input to the backup vector algorithm. Figure 99 shows the structure of the propolytic alteration zone inputs.

Hünermend et al. (2011) state that phyllic, argillic and propolytic changes are observed as uniform in the opaque region. Their research also shows that proplitic alteration is dispersed around most of the porphyry mineral deposits in the region. Figure 12 shows the output of the propolytic change using a boost vector machine. Figure 12 fully confirms the results of the soil and laboratory investigations and propolytic alteration has been detected around the porphyry deposits in the area.

Verifying the Backup Vector Machine Method

Phillite, argillic and propilitic alteration zones are known as sericite, kaolinite, chlorite and epidote minerals. Of course, other common secondary minerals such as goethite, jarosite and hematite have also been found in most porphyry deposits in the region. The most important minerals, showing a wide overlap range, are serist and kaolinite minerals that are well seen in support vector machine results (Figures 9 and 10). This algorithm was unable to make a significant distinction between these two replacement regions. In addition, the current study attempts to measure the classification quality of the backup vector algorithm in the form of a fusion matrix. A matrix is a tabular interaction with the number of rows and columns equal to the number of classes that can measure the capability of the classification algorithm. For this purpose, the fusion matrix for three classes of phyllite, argillic and propolytic alteration is presented in Table 3. In the Blending Matrix in Table 3, each column gives information about the actual pixels used in a region. This means that the support vector machine detects 93 pixels as an argylic zone, 127 pixels as a filigree zone, and 295 pixels as an unmodified or unclassified zone, the number of pixels removed from the agile zone is 516. Each row of the fusion matrix also provides information about the pixels lined up by the support vector machine in each region. Therefore, the number of pixels selected by the backup vector machine as the clayey region is 335 pixels, of which only 94 pixels have been correctly selected. And 237 pixels of phyllic zone and 4 pixels of propolytic zone are incorrectly classified as clayey

zones. To measure the classification quality of the support vector machine algorithm using the fusion matrix, the deletion and realization error functions are calculated [14]. The total accuracy parameter and the kappa coefficient are also used [14]. A deletion error for the slanted zone means that there are 46.57% of the pixels belonging to the agile zone. Other zones are incorrectly classified, and an agile zone rendering error means that 91.71% of the pixels from other zones are misplaced into the curved zone. The total accuracy parameter indicating the ratio of correctly classified pixels to total pixels should ideally be 100%. The kappa coefficient, which expresses the degree of conformity of the classification to reality suggested by Kangalton and Greene (2009), will be equal to one at best. According to Table 3, it can be seen that despite the small number of training pixels in the propolytic region, the backup vector machine performs best relative to this region. And the errors in removing and executing this zone are equal to 6.96% and 61.4%, respectively. However, the performance of the algorithm used in the filtrate and agile regions is weaker because the spectral overlap between the mineral characteristic of these regions is important and causes the results of the support vector machine to be affected. The results of the algorithm in the argil region reach their worst state, which can be attributed to a smaller number of training pixels than the filigree region. However, another point that is ignored in the error functions mentioned above in Table 3 is that the real change pixels are the use of a support vector machine to separate the real shift regions that the algorithm does not classify in any shift region. And they are called unclassified or inverted pixels. To measure this behavior of the backup vector algorithm, this study defines the error parameter as the blind error, which introduces the ratio of the actual pixels of the unclassified change in each region. For the argylic region, the blind error means that 57.17% of the pixels of the argylic alteration region are not recognized as replacement regions at all and are introduced as unchanged regions. The computed blind error in Table 3 shows that the backup vector machine has the highest blind error despite the lowest error in separating pixels in the propolytic region, that is, there are definitely change pixels in the region not detected by this algorithm. In addition to the fusion matrix, this study uses traditional band and shell ratio analysis methods to compare the degree of separation of the displacement regions compared to the support vector machine algorithm. Ahmadi Rouhani et al. (2014) thought that 4/6 and 4/5 Bandi ratios were effective in determining phyllic and argillic alteration regions, respectively. They also introduced a band ratio of 7.4 to identify minerals in the propolytic alteration zone. As shown in Figure 13, the above band ratios are used and the mineral regions are shown in a circle. Comparison of the three output images shows the overlap of the three displacement zones in the zone and the mineral zones have almost no inseparability.

Aliani et al. (9313) used the shell selector method to detect phyllite, argillic and propolytic regions. As input to the principal component analysis method, they used bands 4, 5 and 7 for argillic alteration, bands 4, 6 and 7 for filteric change, and bands 7, 8, and 9 for proplitic change. This research tried to compare the resolution of the changes with the support vector algorithm using the specified main components (Figure 92). As Figure 92 shows, detection of the proplitic change by components 7, 8 and 9 is poorly performed, and the change is

poorly performed, and the filamentous and argillic changes have considerable overlap. According to the results of Bandi and Kirista ratio analysis, it is not possible to separate the changes. And a different method must be used to show the results of the backup vector machine that makes this resolution more accurate.

RESULT

In this research, an attempt has been made to investigate the separation of phyllite, argillic and propolytic alterations in porphyry copper deposits of the Kerman Cenozoic magmatic belt using the support vector machine algorithm. For this purpose, 9 bands visible and infrared range near the primary are chosen as inputs of this algorithm. The results show that it is not possible to separate filite and argillic alterations using a support vector machine, because kaolinite as a feature of the argylic region has a significant spectral overlap with the sericite of the characteristic mineral of the filigree region. This fact is also observed in the computed entanglement matrix. This matrix indicates that while the propolytic region is classified well, the classification errors for the phyllite and

.

argylic regions are high. On the other hand, the purpose of this research is to introduce an error function called blind error, which is calculated using the fusion matrix. This error is a measure of pixels that are not detected by the algorithm as replacement regions. According to this criterion, the backup vector machine classifies most of the pixels of the propolytic region as unchanged regions, but can distinguish them well from filtrate and argillic regions if they are classified as replacement regions. In addition, traditional image processing results such as band and shell ratio analysis show that these methods cannot distinguish changes from each other. If the backup vector splits the machine more accurately. Finally, it can be said that the support vector machine algorithm is a relatively good tool for detecting filteric, argillic, and propolytic changes. However, it cannot make an acceptable distinction between phyllic and argillic zones, and the propolytic zone is carefully selected. However, this region has a high blind error and needs to be corrected by other methods