Article

Identification of Hydrothermal Alteration Minerals for Exploring Gold Deposit Based on SVM and PCA Using ASTER Data: A Case Study of Gulong

Kai Xu[1,2,3], Xiaofeng Wang[1], Chunfang Kong[1,2,4]*, Ruyi Feng[1,2]*,Gang Liu[1,2], and Chonglong Wu[1],

[1] School of Computer, China University of Geosciences, Wuhan, 430074, China; xukai@cug.edu.cn, wxf@cug.edu.cn, kongcf@cug.edu.cn, fengry@cug.edu.cn, Liugang@cug.edu.cn, Wucl@cug.edu.cn
[2] Hubei Key Laboratory of Intelligent Geo-Information Processing (China University of Geosciences (Wuhan)), 430074, P. R. China; Xukai@cug.edu.cn, kongcf@cug.edu.cn, Liugang@cug.edu.cn
[3] Key Laboratory of Tectonics and Petroleum Resources (China University of Geosciences), Ministry of Education, Wuhan, 430074, China; Xukai@cug.edu.cn,
[4] National-Local Joint Engineering Laboratory on Digital Preservation and Innovative Technologies for the Culture of Traditional Villages and Towns, Hengyang, 421000, China; kongcf@cug.edu.cn
* Correspondence: kongcf@cug.edu.cn, fengry@cug.edu.cn; Tel.: +86-027-6788-3286

Abstract: Dayaoshan, as an important metal ore producing area in China, is faced with the dilemma of resource depletion due to long-term exploitation. In this paper, remote sensing method is used to circle the favorable metallogenic areas and find new ore points for Gulong. Firstly, vegetation interference bas been removed by using mixed pixel decomposition method with hyperplane and genetic algorithm (GA) optimization; then, altered mineral distribution information has been extracted based on principal component analysis (PCA) and support vector machine (SVM) method; Thirdly, the favorable areas of gold mining in Gulong has been delineated by using ant colony algorithm (ACA) optimization SVM model to remove false altered minerals; Lastly, field survey verified that the extracted alteration mineralization information is correct and effective. The results show that the mineral alteration extraction method proposed in this paper has certain guiding significance for metallogenic prediction by remote sensing.

Keywords: gold deposit; Alteration information; ASTER image; support vector machine (SVM); principal component analysis (PCA)

1. Introduction

Surrounding rock alteration is one of the important interpretation markers of mineral exploration using remote sensing technology [1, 2]. The metasomatism of hydrothermal mineralization can cause the minerals to alter and generate groups or ions such as Fe3+, magnesium hydroxyl group and aluminum hydroxyl group, which show different hue and spectral characteristics compared to the non-altered rock in remote sensing images [3–5]. In the process of mineral exploration, according to the different hue and reflectance spectra of altered minerals, the composition and spatial distribution of altered minerals can be analyzed by remote sensing technology, and the favorable metallogenic areas can be found out.

With the development of remote sensing technology, most scholars have focused on mechanism of the mineral spectral [6–9].Others have studied extraction method of altered mineral [10–12]. Still some have researched on identification of hydrothermal alteration minerals using multi- and hyper-spectral images [13–16]. In recent years, altered mineral identification have been investigated by using different methods in different locations for mineral exploration, including alteration mineral mapping in Northwestern Junggar basin using Landsat TM data and principal component analysis (PCA) [17], predictive mineral prospectivity modeling for Cu deposits in Varzaghan district, NW Iran base on support vector machine (SVM) [18], as well as predictive

models for Rodalquilar mining district mineral prospectivity in the southeast of Spain with machine learning [19]. All these researches contribute to the development of metallogenic prognosis in arid and semi-arid area. However, study on the altered mineral extraction in high vegetation covered areas is still limited, particularly in a specific region as an important metal mineral area [20, 21], where many old mines are facing resource depletion after long-term exploitation, such as Gulong, China.

In response to the above problems, the objective of this paper is to: (1) select the image identification marks of altered minerals according to the distribution of gold deposits and the ore-bearing strata in Dayaoshan; (2) remove vegetation interference by using mixed pixel decomposition method with hyperplane and genetic algorithm (GA) optimization; (3) extract altered mineral information based on PCA and SVM method by the ASTER data; (4) verify the altered mineral extraction result using field survey method. Lastly, the research results not only delineate the favorable metallogenic area of gold deposits, but also provide reasonable suggestions for mineral exploration with remote sensing data.

2. Study area and Materials

Gulong Town, located at the junction of Yangtze paleo-plate and Cathaysia paleo-plate, is part of the Guangxi Dayaoshan Au ore belt, dotted with gold deposits [22–25] (Figure 1). It belongs to the subtropical monsoon climate region, where the annual average temperature is 21°C, the frost-free days are about 320 days, and the average annual rainfall is 1600 mm. There are a lot of pristine forests and all kinds of herbaceous plants, and forest cover can reach more than 90%.

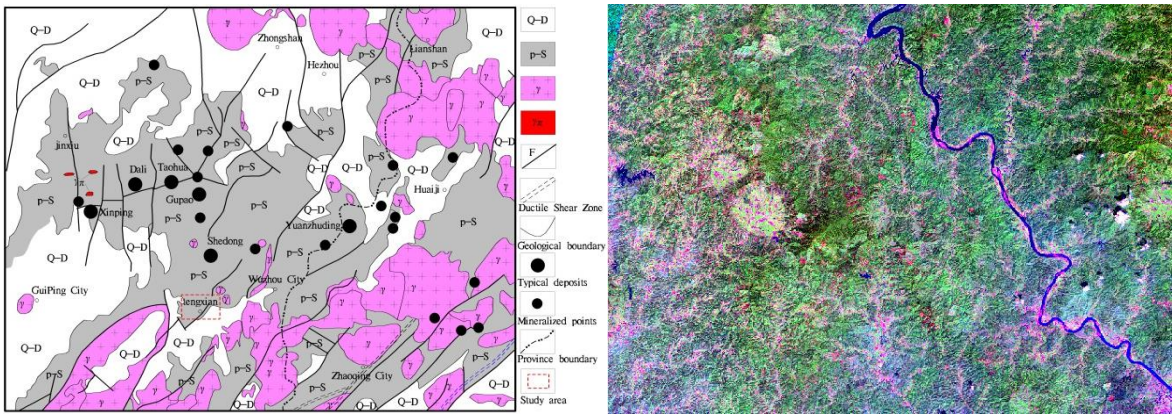


Figure 1. Reginal geological sketch map and ASTER image of Dayaoshan area in Guangxi

The study area is strongly influenced by multistage structural evolution, such as sedimentation, uplift, and folds and faults in different directions in a long process of geological evolution, which makes the structures of the research area very complex. The exposure strata are mainly Sinian and Cambrian clastic rock series, and scattered of small rock mass, rock strain and vein of the Yanshanian period, which plays a vital role in formation of porphyry deposits of the whole study area. At present, the endophytic metal deposits found in the study area can be divided into porphyry deposits, broken zone altered rock deposits, and quartz vein deposits. Those three kinds of deposits are often associated with each other, that is to say, there are altered rock deposits and quartz vein deposits of different scales in the outer part of the porphyry deposits.

The ASTER images adopted in this paper were from the Guangxi Remote Sensing Center, and consist of six level 1B scenes for the study area. Before identifying the altered mineral, the images were pretreated, such as mosaic and subset, atmospheric correction by the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) [26, 27], crosstalk correction [28], water disturbance removed by using a water index. The flow chart of ASTER image preprocessing is shown in Figure 2.

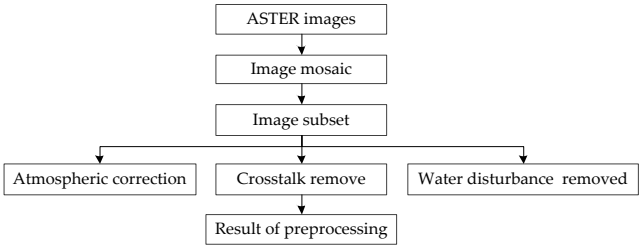


Figure 2. Flow chart of ASTER image preprocessing

3. Methods

The extraction of altered mineral information is usually based on the spectral characteristics of minerals[9, 29–31]. Although minerals have their own spectral characteristics, rocks are usually formed by various mineral assemblages, so the spectral information of rocks on remote sensing images is very complex. In addition, due to the influence of other surface features, such as soil and vegetation, the altered mineral information is very weak in the remote sensing images [32–34]. Therefore, this paper firstly focuses on eliminating vegetation interference, extracting mineralized information samples with remote sensing images, and then establishing PCA and SVM method to improve the accuracy of altered mineral information extraction.

3.1. Mixed Pixel Decomposition with Hyperplanar Optimized by GA

In the area with high vegetation coverage, the spectral characteristics of pixels in remote sensing images are not a single ground object, but a mixed reflection of the several objects. The altered mineral information is seriously hided by high vegetation coverage, so it is very difficult to extract the information which is needed. For mixed pixels, an unmixed pixel decomposition method with hyperplane and genetic algorithm optimized [35–39], which is established in this paper, will eliminate the influence of vegetation, and obtain altered mineral information from the sub-pixel level.

In the feature space of remote sensing image, the same kind of mineral is assembled in one region because of its similar spectral characteristics, while other similar minerals are gathered in another region. By calculating the hyperplane, these kinds of features can be separated, and the proportion of the objects contained in each pixel can be obtained, that is, the decomposition result of the mixed pixels, which can be expressed as formula (1):

$$f(x) = M * X > 0,\ f(x) = M * X \le 0 \tag{1}$$

Where M=m_1,m_2...m_n，X=(x_1,x_2,…,x_n)，n is the dimension of feature space.

The position of hyperplane in multidimensional remote sensing data space can be well represented by iterative recursive mode, and the equation is expressed as (2):

$$\begin{aligned} d &= \alpha_N \cos\alpha_{N-1} + \beta_{N-1}\sin\alpha_{N-1} \\ \beta_{N-1} &= \alpha_{N-1}\cos\alpha_{N-2} + \beta_{N-2}\sin\alpha_{N-2} \\ \beta_{N-2} &= \alpha_{N-2}\cos\alpha_{N-3} + \beta_{N-3}\sin\alpha_{N-3} \\ &\cdots \\ \beta_1 &= \alpha_{N-2}\cos\alpha_0 + \beta_0\sin\alpha_0 \end{aligned} \tag{2}$$

Where α_{N-1} is the angle between the hyperplane unit normal and the X_N, α_{N-2} is the angle between the normal projection in the $X_1, X_2, \cdots, X_{N-1}$ space and the X_{N-1} axis, X_{N-3} is the angle between the normal projection in the $X, X_2, \cdots, X_{N-2}$ space and the X_{N-2} axis, α_1 and α_0 is the angle between the normal projection of 2D space and the 2nd feature axis, and 1D space and the 1st feature axis, respectively. if α_0=0, d is the vertical distance between the hyperplane and the origin, In this way, a hyperplane of N-dimensional space can be determined by N $\geqslant$ 1 angle $\alpha_1, \alpha_2, \cdots \alpha_{N-1}$ and a vertical distance d [40].

It can be seen from formula (2), if hyperplane is used to classify the feature space composed of multi-band remote sensing images, the first problem needs to be solved is how to determine the N parameters of hyperplane equation, which can best achieve the classification effect. In this paper, to establish hyperplane classification model optimized by genetic algorithm for classification, firstly, the pattern description and pattern matching of the samples in the training sample set are carried out; secondly, compare and select different pattern schemes with evolution, and get the best pattern classification scheme. Finally, extend to the whole image to achieve the purpose of pattern classification. The process is as follow Figure 3.

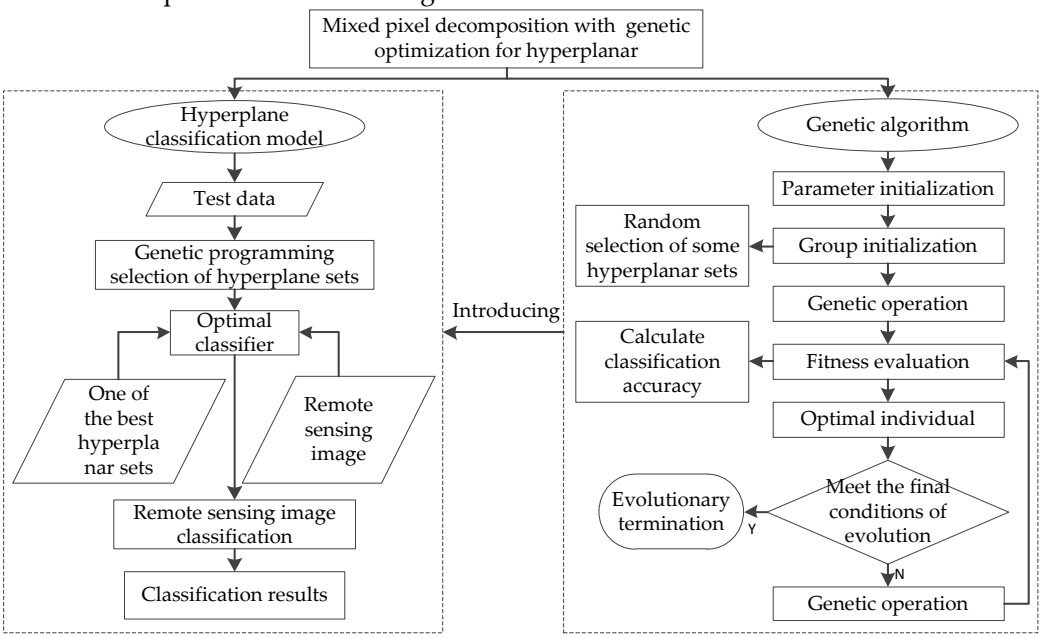


Figure 3. Mixed pixel decomposition with genetic-hyperplane flow chart

3.2. Mineral Altered Samples Collection with Ratio Analysis and PCA

By using the ratio of reflection and absorption band, the spectral difference among various geological information can be enhanced, and the influence of topography can be reduced [41–44]. Therefore, according to the characteristic of spectrum, the selection of appropriate ASTER band for ratio operation can enhance the weak altered mineral.

Because of high correlation among the bands of images, some of the data are redundant and repetitive [13, 45–47]. In this paper, PCA of images is carried out, and the independent quantities are obtained in order to reduce the correlation influence in the process of alteration information extraction.

PCA combines the original n features linearly and establishes m principal components, which are not correlated with each other, and the variance decreases gradually. The expression of the transformation is as follow (3):

$$Y = AX = \sum_{i=1}^{n} A_i X_i \quad (3)$$

Transformation is to reduce the dimension of the original data, using m-dimensional vector can cover the original n-dimensional information, and make its standard deviation within the allowable range, then there is (4):

$$Y = AX = \sum_{i=1}^{m} A_i X_i + \sum_{i=m+1}^{n} A_i X_i \quad (4)$$

Where X_i is denoted as b_i, then the formula (4) becomes (5):

$$Y(m) = \sum_{i=1}^{m} A_i X + \sum_{i=m+1}^{n} A_i b_i \quad (5)$$

Let the error is ε between Y and Y(m),

$$\varepsilon = \sum_{i=m+1}^{n} A(X_i - b_i) \quad (6)$$

3.3. Altered Mineral Extraction Using SVM and ACA

To improve the extraction accuracy of mineral alteration information, SVM model, an excellent remote sensing image classification technology [48–52], was trained using alteration sample data by PCA. SVM firstly maps samples to high-dimensional spaces by nonlinear transformation; and then finds out the optimal classification hyperplane in high-dimensional space; at last, classifies the sample data.

The samples are set (X_i,Y_i), i=1,2,3,…n, $Y_i \in \{-1,1\}$, $\omega X+b=0$ is the classified hyperplane of the sample, so construct discriminant function (7):

$$g(x)=\omega X+b \tag{7}$$

where ω represents the normal vector, b represents the displacement, which determines the distance between the hyperplane and the origin. For the optimal classification hyperplane, the SVM is:

$$\min_{\omega,b} \frac{1}{2}\left|\left|\omega\right|\right|^2$$

$$\text{s.t.}\, y_i \cdot (\omega^T \cdot x_i + b) \geq 1, i = 1,2,\dots,l \tag{8}$$

For nonlinear data samples, relaxation variables (ξ_i) and penalty factors (C) are introduced to deal with the error problem, Formula 8 becomes 9.

$$\min_{\omega,b} \frac{1}{2}\left|\left|\omega\right|\right|^2 + C\sum_{i=1}^{l}\xi_i$$

$$\text{s.t.}\, y_i \cdot (\omega^T \cdot x_i + b) \geq 1 - \xi_i,\ i = 1,2,\dots,l \tag{9}$$

In order to map the sample to the higher dimensional feature space, and to avoid the difficulty of calculation, the Gaussian kernel function is introduced.

$$k(x_i, x_j) = \exp(-\gamma||x_i, x_j||^2) \tag{10}$$

where $\gamma = \frac{1}{2\sigma^2}$

Classification decision function of optimal classification hyperplane is (11).

$$h(x) = sgn(\sum_{j=1}^{l} a_i y_i k(x_i, x_j) + b) \tag{11}$$

In summary, determining the parameters C and γ of SVM is the key to improve the classification accuracy. So the ACA is introduced to solve the problem of SVM parameter optimization [53–57]. ACA, as a heuristic algorithm based on global optimization, can solve the parameter optimization problem very well. The optimization operation steps are as follows: (1) determine parameter search interval, $C \in [0,10]$,$\gamma \in [0,1]$; (2) initialize, each ant carries on the random allocation operation, named {C, γ}, the grid point on each space corresponds to one state, and each state corresponds to a solution on the space. The movement trajectory of ants in the grid is recorded, and the pheromone is arranged according to the objective function value between each grid point, so that the ant can search according to the size of the pheromone on the grid point; (3) the root mean square error is set as the objective function, and then the best combination is selected among the many parameter combinations. The optimization algorithm flow of ACA is shown in Figure 4.

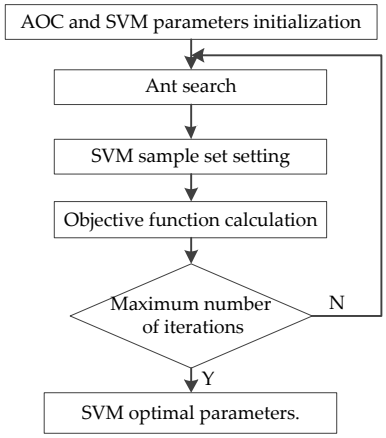


Figure 4. Flow chart of SVM parameters with ACA

To sum up, the steps of remote sensing altered mineral information extraction with PCA and SVM methods are as follows:

(1) The remote sensing images in the study area are preprocessed, including radiometric calibration, atmospheric correction and descrambling, and eliminate the influence of vegetation with mixed pixel decomposition method.

(2) The threshold density segmentation of pyrite, sericite and chlorite principal components is carried out with ratio analysis and PCA, and the training samples are collected according to the known deposits.

(3) Optimize the kernel parameters and penalty factors of SVM with ACA.

(4) SVM classifier trained by training samples is used to process the remote sensing image to obtain the abnormal information of altered mineral.

4. Results and Discussion

4.1. Vegetation Disturbance Elimination

In the study area, there are three main land cover types of vegetation, bare land and water, and 8460, 6320, 4210 training samples are selected respectively, about 2.32% of the total data of the study area (Table 1). The training samples and reference samples are carried out in ENVI4.8, and the number of samples meets the requirements of training and testing.

Table 1. Training samples

Band	Type	Number of samples	Total number
C1	vegetation	8460	
C2	bare land	6320	18990
C3	water	4210	

Three bands of remote sensing image are selected, and three kinds of training samples are classified by using two-dimensional feature space. According to formula (2), the hyperplane equation can be simplified to:

$$d = x_2 \times \cos\alpha + x_1 \times \sin\alpha \tag{12}$$

Because a chromosome represents a set of hyperplane sets, and there are three kinds of objects to be classified, a group of hyperplanes has three chromosomes. A hyperplane consists of an angle α and a distance d, and there are 30 chromosomes in the genetic initialization population in this paper.

The parameter setting of GA is very important to the result accuracy and time efficiency of hybrid pixel decomposition. In this paper, appropriate parameters are determined through the following three tables (Table 2–4).

Table 2 The relationship between variation rate and training accuracy & time
(Number of samples: 500; Type: 3; Crossing rate: 0.99)

Variation rate	0.01	0.02	0.05	0.07	0.09	0.1	0.2	0.3
Generations	1060	1120	2000	2000	1560	2000	2000	2000
Correct number of classifications	432	456	391	416	381	372	365	371
Training accuracy	0.864	0.912	0.782	0.832	0.762	0.744	0.73	0.742
Training time (s)	261	259	302	431	462	294	309	192

When the variation probability is less than 0.1, the training accuracy is high. When the variation probability is set to 0.02, the relationship between the cross rate and the training accuracy and training time is shown in the following table when the other parameters remain unchanged (Table 3).

Table 3 The relationship between Crossing rate and training accuracy & time
(Number of samples: 500; Type: 3; Variation rate: 0.02)

Crossing rate	0.02	0.05	0.1	0.3	0.4	0.6	0.8	0.99

Generations	711	623	658	110	1120	812	421	1128
Correct number of classifications	457	451	451	445	453	472	465	471
Training accuracy	0.914	0.902	0.902	0.89	0.906	0.944	0.93	0.942
Training time (s)	149	138	141	92	245	193	121	207

When the cross rate is 0.3, the training time is the least and the training accuracy is high. When the cross rate is equal to 0.3 and the variation rate is equal to 0.02, the relationship between the number of training points and the training accuracy & time is shown in the following table 4:

Table 4 The relationship between training points and training accuracy & time (Type: 3; Variation rate: 0.02; Crossing rate:0.3)

Number of samples	100	200	300	500	900	1200	2000	4000
Generations	631	292	273	446	68	2000	2000	2000
Correct number of classifications	94	192	273	359	465	864	1187	2739
Training accuracy	0.94	0.96	0.91	0.718	0.517	0.72	0.594	0.685
Training time (s)	0.0548	0.089	0.132	0.184	0.231	0.734	0.891	1.632

When the cross rate is equal to 0.3 and the variation rate is equal to 0.02, the increase of training points will not lead to the increase of training time in a certain range.

In this paper, the genetic algorithm uses binary coding. The initial population is randomly selected, and then according to the above research, cross rate is equal to 0.3, and variation rate is equal to 0.02.

After 212 generations of genetic evolution, 18250 training samples were successfully classified, and the training success rate was 96.1%. The three hyperplane equations and angles acquired by genetic training were as Table 5:

Table 5 Hyperplane set

Number	Distance	Angle	Hyperplane equations
H1	1.546207	44.136729	$1.546207 = x_2 \times \cos 44.136729 + x_1 \times \sin 44.136729$
H2	0.760832	161.892638	$0.760832 = x_2 \times \cos 161.892638 + x_1 \times \sin 161.892638$
H3	1.546207	14.834792	$1.546207 = x_2 \times \cos 14.834792 + x_1 \times \sin 44.14.834792$

The mixed pixel decomposition with hyperplane and GA is used to process the remote sensing images, and the categories are determined according to the final assignment, and the vegetation interference are removed, and results are as Figure 5.

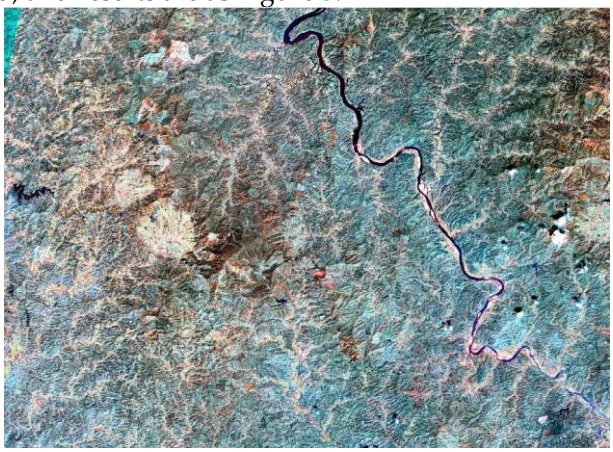


Figure 5. Results of vegetation removal by mixed pixel decomposition

4.1. Collection of Altered Mineral Samples

Through the analysis of the distribution of gold deposits and the ore-bearing strata of the study area, the pyrite, sericite and chlorite are the most closely related to altered mineral of surrounding rock, which is a good prospecting indicator of gold ore.

Pyrite mineralization, in the light of the minerals reflectance spectra, B1, B2, B5/B3, and B4 of ASTER data are adopted for PCA. The eigenvector of principal component is shown in Table 6.

Table 6 Pyrite eigenvector matrix

Eigenvector	B1	B2	B5/B3	B4
1	0.166731	0.120621	0.822539	0.530171
2	0.106507	0.233433	-0.562449	0.786013
3	0.599849	0.734171	0.024860	-0.317107
4	-0.775267	0.626062	0.080392	0.023353

In Table 6, the eigenvector of B1 and B2, B4 are opposite, the eigenvector of B1 is negative, and the eigenvector of B2 is positive. Therefore, PC4 reflects the alteration information of iron mineralization.

In order to get pyrite mineralization, PC4 needs a series of treatments. First, the PC4 is filtered by median filter to eliminate the noise. Then, mean and standard deviation are counted by linear stretching. Finally, the pyrite mineralization is obtained by using mean and standard deviation as the threshold for abnormal segmentation, and the segmentation results are as Figure 6:

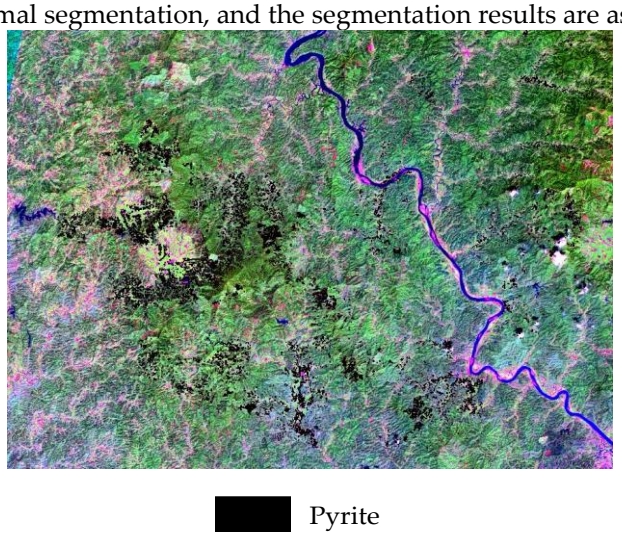


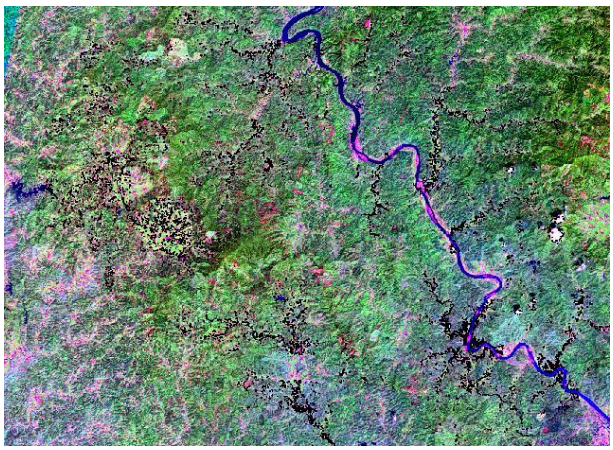
Figure 6. Alteration distribution of pyrite mineralization

Figure 6 shows that the alteration zone is mainly located in the South Central of the study area, and the alteration reaction is strong in some areas of South Central, which is the possible point of mineralization.

Table 7 Sericite eigenvector matrix

Eigenvector	B1	B4/B6	B7	B9
1	0.164443	0.811045	0.523520	-0.202710
2	0.070343	-0.569964	0.690800	0.439304
3	0.859394	0.038299	-0.353052	0.367872
4	-0.479012	0.126017	0.352244	-0.794098

Table 7 shows that the PC4 reflects sericite alteration information. And PC4 needs to be processed in the same way as pyrite, and the segmentation results are as Figure 7.



Sericite

Figure 7. Alteration distribution of sericite mineralization

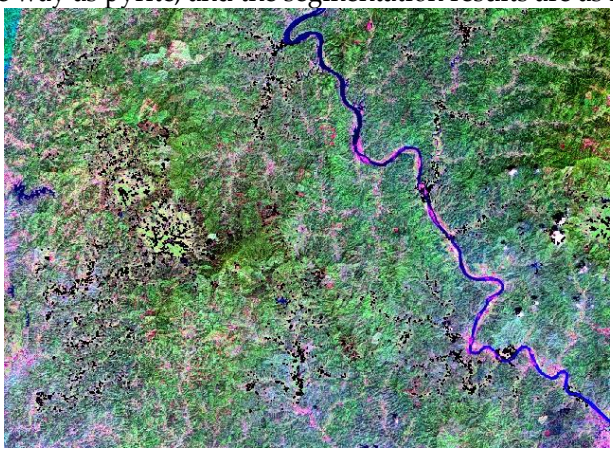
Figure 7 indicates that the alteration zone is mainly located in the south-central part of the study area, and the alteration reaction is strong in some areas of central and southern, which is the possible place of prospecting.

For chlorite mineralization, B1, B2, B5/B8 and B9 from ASTER data are adopted for PCA. The eigenvector of principal component is shown in Table 8.

Table 8 Chlorite eigenvector matrix

Eigenvector	B1	B2	B5/B8	B9
1	-0.162074	0.799186	0.516097	0.262064
2	0.063939	-0.592936	0.688779	0.412216
3	0.909014	0.073300	-0.309002	-0.269885
4	0.378598	-0.065995	0.404654	0.829797

Table 8 illustrates that PC4 can reflect the alteration information of chlorite. And PC4 needs to be processed in the same way as pyrite, and the segmentation results are as Figure 8.



Chlorite

Figure 8. Alteration distribution of chlorite mineralization

Figure 8 illustrates that the alteration zone is mainly located in the southwest of the study area, and the alteration reaction is strong in some areas of southwestern, which is the possible areas of prospecting.

4.3. Altered Mineral Information Extraction

In order to obtain mineralization alteration sample, the mineralization alteration information of pyrite, sericite, and chlorite were loaded into the previous survey results map (Figure 9).

Three kinds of mineralized alteration samples, 300 in each category, and 900 samples of non-mineralized alteration were selected from the overlay map to train the SVM model. Among them, 80% of the samples (half of them were mineralization alteration, 1/3 of per mineralization

alteration) were randomly selected as training samples, and the remaining 20% were used as the test samples.

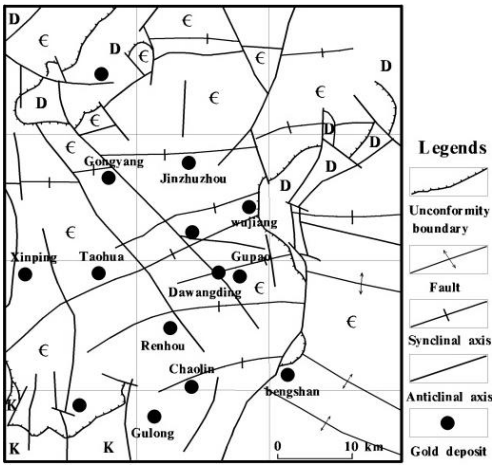


Figure 9. Distribution map of gold deposits in Dayaoshan area

On the basis of selecting SVM kernel function, the SVM is trained by the above sample data, at the same time, SVM is optimized by ACA, in which the maximum number of iterations K=10, ant colony size N=20, pheromone evaporation coefficient Rho=0.8, pheromone increasing intensity Q=0.9, ant crawling speed Lambda=0.3. The mineralization alteration information obtained by PCA method in the study area are input into the trained SVM, the false information will be removed, and the mineralization alteration results will be obtained as shown in Figure 10.

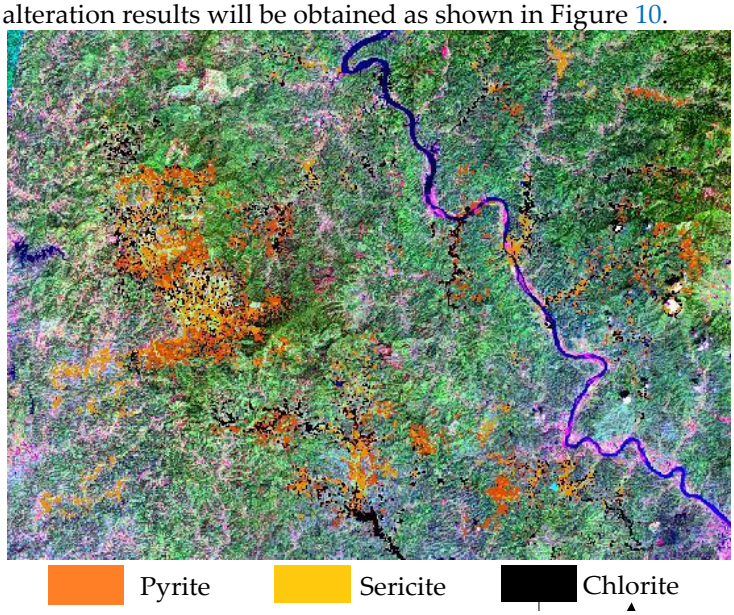


Figure 10. Extraction results of mineralization alteration information

Figure 10 indicates that the orange, black and yellow areas represent the extraction results of pyrite, sericite and chlorite in the study area, respectively. Compared with Figure 6-8, it can be seen that the range of the alteration area is smaller, which is due to the removal of some false anomalies. Figure 10 also indicates that mineral alteration does not exist alone, but is associated with each other, which is a good prospecting indicator for gold deposits.

4.4. Verification of Mineralization Alteration Information Extraction Results

In order to verify the accuracy of the extracted mineral alteration information, a special field investigation was carried out. First of all, the working scope of field verification is determined according to the spatial distribution of remote sensing alteration information. Then, the field survey route is planned according to the regional geological characteristics. Finally, the rock specimens with alteration anomalies are collected in different locations. After 15 days of field sampling and

verification in Gulong of Dayao Mountain, 116 rock samples were collected at 52 sampling points, and the rock sample verification analysis is as follows.

(1) The exposed rocks in Sanxianding have obvious pyritization with a light brass color and bright metallic luster, which showing that the altered minerals contained in the outcrop rocks in the field are consistent with the results extracted by ASTER images (Figure 11). In hydrothermal deposits, pyrite is symbiotic with other sulfides, oxides, quartz and so on. This phenomenon indicates that such rock composition often contains gold, silver and other elements.

Figure 11. Samples at Sanxianding (110.945°E, 23.641°N; 2015-2-7)

(2) The rocks in Hecun sampling area are mainly granite porphyry with 15-20% phenocryst, the main rock components are quartz and feldspar, and sometimes biotite and hornblende. Quartz phenocrysts usually to be hexagonal bipyramid, and biotite and hornblende sometimes were existed darkening edges (Figure 12). This phenomenon indicates that there are intrusive rocks in the sampling area, which directly proves the existence of mineral alteration in surrounding rock (Figure 12). Metallic minerals related to granitic porphyry are gold, silver, copper, and so on.

Figure 12. Samples at Hecun (111.027°E, 23.540°N; 2015-2-8)

(3) The sampling point is mainly meso-basic igneous rock and metasomatic rock, which is the surrounding rock related to chloritization (Figure 13). Chloritization rarely occurs alone, often accompanied by pyritized, sericitization, propylitization, and carbonate and so on.

Figure 13. Samples at Sunjia mountain (111.013°E, 23.652°N; 2015-2-10)

The results of field verification certified that 40 of the 52 sampling sites can collect rocks containing pyrite, sericite and chlorite, which is consistent with the results of mineralization and alteration extraction using remote sensing image. The results of this paper indicate that the extraction results of alteration information in this paper have very high accuracy and could play a vital role in the exploration of metal minerals in the study area. According to the extracted

mineralization alteration and the results of field verification, the favorable areas of gold deposit can be delineated (Figure 14).

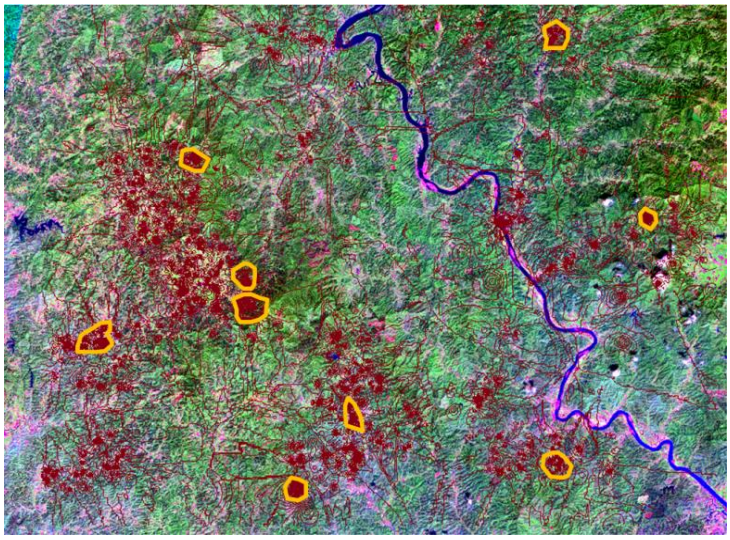


Figure 14. The favorable areas of gold deposit

5. Conclusions

(1) The mixed pixel decomposition method with hyperplane optimized by genetic algorithm is used to remove the influence of vegetation interference information from the interior of the pixel. Thus provides a methodological reference for areas with high vegetation coverage and strong disturbance information, and also provides the possibility for extracting other weak information in this paper.

(2) Extraction model of altered mineral information based on PCA and SVM is constructed, and successfully extracted the abnormal information of mineralization alteration in the study area, by comprehensive processing the multi-source heterogeneous data, such as remote sensing ground object spectral data, laboratory mineral spectrum data, and surface rock mineral distribution data and so on.

(3) Field survey verified that the extracted alteration mineralization information is correct and effective. At the same time, the rock samples collected in the field proved that the method of extracting alteration mineralization information in this paper is feasible. The results of the study can delineate favorable areas for gold mining in Gulong area. Simultaneously, the results can provide technical support for the multi-source heterogeneous data fusion and geological big data metallogenic prediction.

Author Contributions: All the authors made significant contributions to the work. Kai Xu, Xiaofeng Wang, Ruyi Feng, and Chunfang Kong designed the research, analyzed the results, and accomplished the validation work. Kai Xu, and Xiaofeng Wang completed field investigation, sampling, analysis, and validation; Gang Liu, and Chonglong Wu provided advice for the revision of the paper.

Acknowledgments: The authors would like to thank the Remote Sensing Center of Guangxi for providing the various data sets used in this paper. This work has been supported by the National Natural Science Foundation of China (No: 41201193); the Research on evaluation method of Oil and Gas Resources investigation (prediction) in "Belt and Road Initiative" Spatial Information Corridor; the Open Research Project of Key Laboratory of Intelligent Geo-Information Processing of Hubei (No.KLIGIP-2017B07); the Open research project of key laboratory of Tectonics and Petroleum Resources (China University of Geosciences), Ministry of Education (No: TPR-2019-11); and the Open fund project of National-Local Joint Engineering Laboratory on Digital Preservation and Innovative Technologies for the Culture of Traditional Villages and Towns (CTCZ19K01). The authors would like to thank the anonymous reviewers for providing valuable comments on the manuscript.

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

References

1. Sabins, F.F. Remote sensing for mineral exploration. Ore Geo. Rev. **1999**, 14, 157–183.
2. Tommaso, I.D.; Rubinstein, N. Hydrothermal alteration mapping using ASTER data in the Infiernillo porphyry deposit, Argentina. Ore. Geo.Rev. 2007, 32, 275–290.
3. Dilles, J.H.; Einaudi, M.T. Wall-rock alteration and hydrothermal flow paths about the Ann-Mason porphyry copper deposit, Nevada–a 6-km vertical reconstruction. Econ. Geol. **1992**, 87, 1963–2001.
4. Hubbard, B.E.; Crowley, J.K. Mineral mapping on the Chilean–Bolivian Altiplano using co-orbital ALI, ASTER and Hyperion imagery: Data dimensionality issues e solutions. Remote Sens.Environ. **2005**, 99, 173–186.
5. Ducart, D.F.; Crosta, A.P.; Filio, C.R.S. Alteration mineralogy at the Cerro La Mina epithermal prospect, Patagonia, Argentina: field mapping, short-wave infrared spectroscopy, and ASTER images. Econ. Geo. **2006**, 101, 981–996.
6. Mars, J.C.; Rowan, L.C. Spectral assessment of new ASTER SWIR surface reflectance data products for spectroscopic mapping of rocks and minerals. Remote Sens. Environ. **2010**, 114, 2011–2025.
7. Molan, Y.E.; Refahi, D.; Tarashti, A. Mineral mapping in the Maherabad area, eastern Iran, using the HyMap remote sensing data. Int. J. Appl. Earth Obs. Geoinf. **2014**, 27, 117–127.
8. Zadeh, M.H.; Tangestani, M.H.; Roldan, F.V.; Yusta, I. Spectral characteristics of minerals in alteration zones associated with porphyry copper deposits in the middle part of Kerman copper belt, SE Iran. Ore Geol. Rev. **2014**, 66, 191–198.
9. Rajendran, R.; Nasir, S. Characterization of ASTER spectral bands for mapping of alteration zones of volcanogenic massive sulphide deposits. Ore Geol. Rev. **2017**, 88, 317–335.
10. Inzana, J.; Kusky, T.; Higgs, G.; Tucker, R. Supervised classifications of Landsat TM band ratio images and Landsat TM band ratio image with radar for geological interpretations of central Madagascar. J. Afr. Earth Sci. **2003**, 37, 59–72.
11. Abedi, M.; Norouzi, G.H.; Bahroudi, A. Support vector machine for multi-classification of mineral prospectivity areas. Comput Geosci-UK. **2012**, 46(3), 272–283.
12. Zadeh, M.H.; Tangestani, M.H.; Roldan, F.V.; Yusta, I. Mineral exploration and alteration zone mapping using mixture tuned matched filtering approach on ASTER data at the central part of Dehaj-Sarduiyeh Copper Belt, SE Iran. IEEE J. Sel. Top.Appl. Earth Obstet. Remote Sens. **2014**, 7, 284–289.
13. Zhang, X.F.; Pamer, M.; Duke, N. Lithologic and mineral information extraction for gold exploration using ASTER data in the south Chocolate Mountains (California). ISPRS J. Photogramm. Remote Sens. **2007**, 62, 271–282.
14. Bedini, E. Mineral mapping in the Kap Simpson complex, central East Greenland, using HyMap and ASTER remote sensing data. Adv. Space Res. **2011**, 47, 60–73.
15. Zheng, S.; Fu, B.H. Lithological mapping of granitoids in the western Junggar from ASTER SWIR-TIR multispectral data: case study in Karamay pluton, Xinjiang. Acta Petrol. Sin. **2013,** 29, 2936–2948.
16. Alimohammadi, M.; Alirezaei, S.; Kontak, D.J. Application of ASTER data for exploration of porphyry copper deposits: a case study of Daraloo-Sarmeshk area, southern part of the Kerman copper belt, Iran. Ore Geol. Rev. **2015**, 70, 290–304.
17. Dai, J.J.; Qu, X.M.; Xin H.B. Extraction of mineral information using ASTER remote sensing data in Duolong area, Tibet, China. Geol. Bull. China **2010**, 29(5), 752-759.
18. Ghezelbash, R.; Maghsoudi, A.; Carranza, E.J.M. Performance evaluation of RBF- and SVM-based machine learning algorithms for predictive mineral prospectivity modeling: integration of S-A multifractal model and mineralization controls. Earth Sci. Inform. **2019**, 1–17.
19. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. **2015**, 71, 804–818.
20. Shao, C.K.; Xue, Y.G.; Zhang P.; Tian H.; Wan J.Y.; Zhou Y.; Zhao Z.N. Comprehensibe anomaly characteristics and metallogenic law of Tongmu sub-Cu-polymetallic ore belt in the west side of Dayaoshan mountain. Contrib. Geol. Miner. Resour. Res. **2019**, 34(3), 445–452.
21. Li, Z.Y.; Dang Y.; Wei, Z.R.; Ye Y.L.; Chen, M.H. Molybdenite Re-Os dating for the Dali porphyry Mo-Cu deposit of Dayaoshan area and its geological implications. J. Guilin Univ. Techn. **2019**, 39(2), 249–257.

22. Xu, Z.X.; Liu, L.M.; Sun, T. Specific features of the porphyry-related gold deposits in the middle of the Dayaoshan Caledonian orogenic belt and constraints of tectonic setting: an example from the Dawangding gold deposit. Geol. Explor. **2012**, 48(2), 305–312.
23. Zhou, G.F.; Liang, B.Z.; Wu, X.K; Zhou, W.J.; Wang, X.Y.; Wei, A.W.; Liang, G.K.; Kang, Z.Q. Metallogenic mechanism of porphyry copper-tungsten-molybdenum-gold deposits in Dayaoshan area. J. Guilin Univ. Techn. **2015**, 35(4), 649–659.
24. Wang, Q.; Yang, Y.Q. Geological characteristics, ore genesis and prospecting direction of Renhou gold deposit in Guangxi province. Gold Sci.Techn. **2016**, 24(1), 41–45.
25. Nong, J.N.; Huang, X.Q.; Guo, S.Y.; Sun, M.H.; Deng, B.; Zou, Y. Discovery of Caledonoan basic rocks in Dayaoshan region Eastern Guangxi and its geological significance. Geol. Sci. Techn. Inform. **2017**, 36(6), 113–121.
26. Thome, K.; Palluconi, F.; Takashima, T.; Masuda, K. Atmospheric correction of ASTER. IEEE Trans. Geosci. Remote Sens. **1998**, 36(4), 1119–1211.
27. Ayoobi, I.; Tangestani, M.H. Evaluation of relative atmospheric correction methods on ASTER VNIR–SWIR data in playa environment. Carbonate. Evaporite. **2017**, 32(4), 539–546.
28. Iwasaki, A.; Tonooka, H. Validation of a crosstalk correction algorithm for ASTER/SWIR. IEEE Trans. Geosci. Remote Sens. **2005**, 43(12), 2747–2751.
29. Hunt, G.R. Spectral signatures of particulate minerals in the visible and near infrared. Geophysics **1977**, 42, 501–513.
30. Cooper, B.L.; Salisbury, J.W.; Killen, R.M.; Potter, A.E. Mid-infrared spectral features of rocks and their powders. J. Geophys. Res. Planets **2002**, 107.
31. Son, Y.S.; Kang, M.K.; Yoon, W.J. Lithological and mineralogical survey of the Oyu Tolgoi region, Southeastern Gobi, Mongolia using ASTER reflectance and emissivity data. Int. J. Appl. Earth Obs. Geoinf. **2014**, 26, 205–216.
32. Bedini, E.; Van Der Meer, F.; Van Ruitenbeek, F. Use of HyMap imaging spectrometer data to map mineralogy in the Rodalquilar caldera, southeast Spain. Int. J. Remote Sens. **2009**, 30 (2), 327–348.
33. Ines, A.V.M.; Mohanty, B.P.; Shin, Y. An unmixing algorithm for remotely sensed soil moisture. Water Resour. Res. **2013**, 49(1), 408–425.
34. Xu, H.; Sun, Y.; Li, X. Unmixing of remote sensing images based on weighted posterior probability support vector machines. J. Geo-Inf. Sci. **2013**, 15(2), 249–254.
35. You, X.; Wang, R.; Tao, D. Diverse expected gradient active learning for relative attributes. IEEE T. Image Process. **2014**, 23(7), 3203–3217.
36. Arabasadi, Z.; Alizadehsani, R.; Roshanzamir, M.; Moosaei, H.; Yarifard, A.A. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. Comput. Meth. Prog. Bio. **2017**, 141(C), 19–26.
37. Crisman, K.D.; Park, J. Zero-cycles with modulus associated to hyperplane arrangements on affine spaces. Manuscripta Math. **2017**, 155(3), 1–31.
38. Gong, D.; Sun, J.; Miao, Z. A set-based genetic algorithm for interval many-objective optimization problems. IEEE T. Evolut. Comput. **2018**, 22(1), 47–60.
39. Bai, X.; Yan, W.; Ge, S.S.; Cao, M. An integrated multi-population genetic algorithm for multi-vehicle task assignment in a drift field. Inform. Sciences **2018**, 453, 209–255.
40. Bandyopadhyay, S.; Pal, S.K. Classification and Learning Using Genetic Algorithms: Applications in Bioinformatics and Web Intelligence (Natural Computing Series). Classification and learning using genetic algorithms : applications in bioinformatics and web intelligence. **2007**.
41. Gad, S.; Kusky, T. Lithological mapping in the Eastern Desert of Egypt, the Barramiya area using Landsat thematic mapper (TM). J. Afr. Earth Sci. **2006**, 44, 196–202.
42. Aydal, D.; Ardal, E.; Dumanlilar, O. Application of the Crosta technique for alteration mapping of granitoidic rocks using ETM+ data: case study from eastern Tauride belt (SE Turkey). Int. J. Remote Sens. **2007**, 28, 3895–3913.
43. Massironi, M.L.; Bertoldi, P.; Calafa, D.; Visona, D.; Bistacchi, A.; Giardino, C.; Schiavo, A. Interpretation and processing of ASTER data for geological mapping and granitoids detection in the Saghro massif (eastern Anti-Atlas, Morocco). Geosphere **2008**. 4(4), 736–759.
44. Wei, L.F.; Wang, H.B. Change detection from high-resolution remote sensing image based on MSE model. Spectrosc. Spect. Anal. **2013**, 33(3):728.

45. Schorr, K.; Da, C.F. Quantitative determination of enhydrin in leaf rinse extracts and in glandular trichomes of Smallanthus sonchifolius (Asteraceae) by reversed-phase high-performance liquid chromatography. Phytochem. Analysis **2010**, 16(3), 161–165.
46. Bertoldi, L.; Massironi, M.; Visona, D.; Carosi, R.; Montomoli, C.; Gubert, F., Nalettoc, G.; Pelizzo, M.G. Mapping the Buraburi granite in the Himalaya of Western Nepal: remote sensing analysis in a collisional belt with vegetation cover and extreme variation of topography. Remote Sens. Environ. **2011**, 115, 1129–1144.
47. Pour, A.B.; Hashim, M. The application of ASTER remote sensing data to porphyry copper and epithermal gold deposits. Ore Geol. Rev. **2012**, 44(1), 1–9.
48. Zhu, G.; Dan, G.B. Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel. Remote Sens. Environ. **2002**, 80(2), 233–240.
49. Cristianini, N.; Shawe-Taylor, J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge, UK: Cambridge University Press, **2004**.
50. Jin, X.; Song, K.; Du, J.; Liu, H.; Wen, Z. Comparison of different satellite bands and vegetation indices for estimation of soil organic matter based on simulated spectral configuration. Agri. Forest Meteorol. **2017**, 244-245:57–71.
51. Hochreiter, S.; Obermayer, K. Support Vector Machines for Dyadic Data. Neural Comput. **2014**, 18(6), 1472–1510.
52. Feng, R.; Wang L.; Zhong, Y. Joint local block grouping with noise-adjusted principal component analysis for hyperspectral remote-sensing imagery sparse unmixing. Remote Sens. **2019**, 11(10), 1-25.
53. Samadzadegan, F.; Hasani, H.; Schenk, T. Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using Ant Colony Optimization. Can. J. Remote Sens. **2012**, 38(2), 139–156.
54. Liang, S. Classification of hyperspectral imagery based on ant colony compositely optimizing SVM in spatial and spectral features. Spectrosc. Spect. Anal. **2013**, 33(8), 2192.
55. Dolques, X.; Ber, F.L.; Huchard, M.; Grac, C. Performance-friendly rule extraction in large water data-sets with AOC posets and relational concept analysis. Int. J. Gen. Syst. **2016**, 45(2), 1–24.
56. Alsaeedan, W.; Menai, M.E.B.; Al-Ahmadi, S. A Hybrid genetic-ant colony optimization algorithm for the word sense disambiguation problem. Inform. Sciences **2017**, 417, 20–38.
57. Wang, M.; Wan, Y.; Ye, Z.; Lai, X. Remote sensing image classification based on the optimal support vector machine and modified binary coded ant colony optimization algorithm. Inform. Sciences **2017**, 402, 50–68.