Remote Sensing Monitoring of Mines in the Yili River Basin

Yang Li[1, a], Xiaoli Liu[2], Haitao Yang[1,b,*], Xixuan Zhou[1], Qian Wang[3]

[1]Aerospace Engineering University, Center for Space Security Studies, Beijing, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Strategic Support Force, Integrated Training Battalion, Beijing, China

[a]lyly149@126.com, [b]yanghtt@126.com

[*] Corresponding author

Keywords: mining area; object-oriented; random forest; Yili River basin; ecological environment

Abstract: The increasing demand of mining resources, the more prominent the contradictory characteristics of frequent mining and ecological environmental protection, for which how to effectively extract large-scale mining points is not only a prerequisite for building a basic database of mines, but also a data support for monitoring the ecological environment of mining areas. In this paper, based on object-oriented technology and random forest, we use Sentinel-2 as the data source, image segmentation to build a multi-dimensional classification feature library of spectrum, topography, texture and geometry, and establish a sample database of mining areas in the Yili River basin with the assistance of mineral rights data; finally, we implement random forest model construction, parameter optimization and classification in Python. The results show that (1) the object-oriented technology can weaken the "pretzel effect" to a certain extent, and the scale, shape, and tightness factors are set to 100, 0.1, and 0.5 to ensure the integrity of the mine area; (2) the optimized random forest parameters NST and MF are 5 and 1000 respectively to meet the relationship between model accuracy and efficiency, and the feature (2) The importance assessment shows that there is a large correlation between the mine area and the spectrum and topography, and only the standard deviation feature has a large correlation among the texture features; (3) There are 395 mine sites in the Yili River basin, with a total area of 143.220km2, which are concentrated in the middle and low elevations and close to the river, spatially reflecting "group", in terms of quantity, the mine area is mostly in the range of 0-0.5km2., and the non-metallic mining sites are the most abundant, with an average distance of 9364.69m from the river, and the metal mining sites are the closest to the river.

1. Introduction

The first national comprehensive risk census of natural disasters (2020-2022) in the Ministry of Emergency Management proposed to conduct a base inventory of mining areas in the country to improve natural disaster prevention and control capabilities by mapping the number, size, and spatial location information of mines nationwide and forming a basic database of mines. Mines are important energy resources on earth, and mining as an economic activity, while making outstanding contributions to socio-economic development, has also become a hot zone of natural environmental degradation [1], such as illegal mines, which not only cause serious damage to the environment, but also often lead to casualties due to illegal operations [2], so by specifically quantifying the scope and period of large-scale mining activities, the establishment of a basic database of mines is essential to securing the ecology and environment of the earth as well as the security of social and economic properties is of great significance [3].

The identification of mining areas is crucial in the establishment of the basic mine database. In the current study, there is the use of medium-resolution sensors based on the spectral and textural features of the features, using fuzzy clustering to aggregate the mine image elements with similar

  DOI: 10.25236/icceme.2023.039

features within a certain distance in the spatial range, so as to extract the boundary information [4]; there is also the use of reflectance datasets to extract NDVI, NBR, NDMI, and combine RED and NIR bands, using multivariate normal distribution fitting to find out the extent of mine areas [5]; when in densely vegetated areas, Julzarika derived the land expansion of the mine area based on the time-series NDVI [6]. In addition, high-resolution images have been used in mine detection, Pagot et al [7] used IKONOS images to extract diamond mines based on object-oriented techniques and maximum likelihood methods, Demirel et al [8] used IKONOS images to detect mine changes based on machine learning techniques, and also scholars [9] combined the role of human vision and Snake model to extract high-resolution images of of mine boundaries for curve evolution, which was then used to identify mine areas in other regions. All of these attempts show the powerful capability of remote sensing in mapping mining activities, however, in contrast, few studies have used remote sensing to map the extent of large regional mining areas [10-12] because of the challenges of text statistics, accuracy and computational effort.

In order to solve the problems faced by remote sensing technology in detecting large scale mining areas using medium resolution images with insufficient accuracy and difficulty in solving data volume using high resolution images, this study takes the Yili River basin in Xinjiang as the study area, based on object-oriented technology and machine learning methods, and combines medium and high resolution Sentinel-2 remote sensing data to mine the Sentinel by building a library of spectral, texture, topographic, and geometric features The aim is to form a basic database of mines, and to provide data support for monitoring the ecological environment of mines.

2. Study area overview and data sources

2.1 Study Area Overview

The Yili River originates in the western part of the Tianshan Mountains, flows through Yili, Xinjiang, China, and finally joins Lake Balkhash in Kazakhstan, which is a cross-border endorheic river between China and Kazakhstan with a basin area of about 15.75×105 km2 [13], of which the basin area in China accounts for about 1/3. The overall characteristics are four mountains sandwiched by three valleys and one basin, high in the east and low in the west, wide in the west and narrow in the east, wedge-shaped [14], and the basin is geologically complex and rich in mineralization conditions, with 86 kinds of minerals in 9 categories: 5 kinds of energy minerals, 16 kinds of nonferrous metal minerals, 4 kinds of ferrous metal minerals, 13 kinds of rare earth minerals, and 4 kinds of nonmetallic minerals. 13 kinds, 4 kinds of non-metallic ores, 10 kinds of chemical raw material ores, 3 kinds of special material ores, 29 kinds of building material ores, etc. It contains dolomite, beryllium and potassium feldspar, which have the highest reserves in China, and important resources coal, oil and gold [15] The study area is outlined in Figure 1.

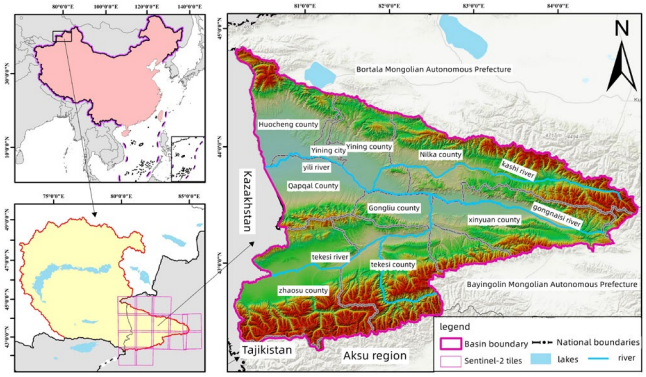

Fig.1 Study Area Overview

2.2 Data source

Sentinel Satellite Data. The Sentinel-2A (S2A) satellite, successfully launched on June 23, 2015, is the second satellite in the Copernicus program and the first satellite of the spectral imaging

mission with high spatial resolution and multispectral features, with spatial resolution up to 10 m. In this study, the S2A data were selected for imaging between June and November 2019. In this study, the S2A data were selected to be imaged from June to November 2019, with less than 5% of clouds.

Terrain Data SRTM. SRTM is made by NASA and NIMA joint measurement processing, and SRTM1 of 30m is selected for this study to obtain topographic features.

Auxiliary Data. The auxiliary data contains the mineral rights data of the Yili River basin, which are outlined by visual interpretation methods as training samples and test samples, where the training samples are used to build the model and the test samples are used for result comparison.

2.3 Data pre-processing

Sentinel-2 LIC data are geometrically refined orthophotos. To obtain atmospheric bottom reflectance data (L2A level) then atmospheric correction is required using the atmospheric correction module provided by Sentinel 2A. To enhance the comparative information between the original and post-mining mine surface, a set of auxiliary variables are generated in addition to the original bands of Sentinel data, including NDVI, NDWI, elevation, and slope.

3. Research Methodology

3.1 Image Segmentation

The advantage of object-oriented over image-oriented classification is that the basic processing unit is the object, which can reduce pretzel noise and improve classification accuracy. Image segmentation is the key step of object-oriented classification, and the segmentation method often uses multi-scale segmentation, and the optimal segmentation scale is determined using the multi-band image scale parameter estimation tool (ESP2) [16], and considering that the method requires several iterations when dealing with large regions, it slices and selects the relatively feature-rich areas to use ESP2 (Table 1), so as to improve the computational efficiency and reduce the computational cost.

Tab.1 Segmentation scale parameter table

Parameter	Numerical value
Shape	0.1
Compactness	0.5
Starting segmentation scale	10

3.2 Feature Constructio

The extracted features in this study (Table 2) contain spectral features: band means of R, G, B, NIR, NDVI, NDWI, as well as brightness and maximum difference; topographic features: elevation, slope; texture features: generated using the Gray Level Co-generation Matrix [17] (GLCM); and geometric features: generated by Cognition calculations.

Tab.2 Classification of detection features in mining area

Feature Category	Feature Name	Abbreviations	Number
Spectral characteristics	Band average	R; G; B; NIR; NDVI; NDWI	6
	Brightness	Brightness	1
	Maximum difference	MaxDiff	1
Terrain characteristics	Elevation	Elevation	1
	Slope	Slope	1
Texture features	Homogeneity	Homogeneity	5
	Dissimilarity	Dissimilarity	5
	Entropy	Entropy	5
	Correlation	Correlation	5
	Contrast degree	Contrast	5

	Second order angular moment	Angular second moment	5
	Mean value	Mean	5
	Standard deviation	Standard deviation	5
Geometric features	Shape index	Shape index (SI)	1
	Aspect Ratio	Length-width (LW)	1
	Roundness	Roundness	1
	Tightness	Compactness	1
	Area	Area	1
	Length	Length	1

Equations for luminance and maximum difference:

$$B = \frac{1}{n_{vis}} \sum_{i=1}^{n_{vis}} \overline{C}_{i(vis)} \tag{1}$$

$$\text{MaxDiff} = \frac{\left| \min\left(\overline{C}_{i(vis)}\right) - \max\left(\overline{C}_{i(vis)}\right) \right|}{B} \tag{2}$$

Where: n: band, C: band eigenvalue.

3.3 Random Forest Classifier

Random Forest [18] is an integrated learning classifier that can run on large datasets and is easy to use. First, bootstrap is used to have put-back extraction to form a subset of samples, and the rest become OOB samples, which are used to evaluate the generalization ability of the method; then CART is used to construct decision trees, after which m features and corresponding The forest is constructed by using CART, and after that, m features and corresponding amount of information are obtained at each node in it to construct the forest; finally, the classification results of each decision tree are integrated (simple voting method).

In this paper, the random forest parameters [19] are optimized: (1) Number of Sub Decision Tree (NST); (2) Maximum number of features allowed to be used in a single decision tree (Max Features, MF); (3) Minimum Sample Leaf size (Min Sample Leaf, MSL); (4) Minimum number of features allowed to be used in each division according to the attributes Minimum number of samples per division when dividing nodes (Min Samples Split, MSS); (5) Criterion supported by the decision tree (Criterion); (6) Whether to calculate out-of-bag error rate (OOB Score).

The accuracy comparison of the classification results is done by calculating the area of the overlapping area between the extracted range and the true range, and when the overlapping area is greater than 70%, the area is marked as correct, using the confusion matrix and the Kappa coefficient [20] (Table 3).

Tab.3 confusion matrix and the Kappa coefficient

Indicators	male type
Accuracy rate(ACC)	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision rate(PPV)	$Precision = \frac{TP}{TP + FP}$
Sensitivity/Recall Rate(TPR)	$\text{Sensitivity / Recall} = \frac{TP}{TP + FN}$
F-value	$F = \frac{2PR}{P + R}$
Kappa Coefficient	$\text{Kappa} = \frac{P_0 - P_e}{1 - P_e}$

4. Results and Analysis

4.1 Distribution of mining area

The distribution of mining areas extracted by the above method is shown in Figure 2 (the boundaries of mining areas are bolded to show obvious), most of the mining areas are close to the Yili River and its tributaries, and in the flat land in the northwest of the basin (Figure a), there are mining areas with large areas, while the eastern part of the basin as well as the southwest are distributed with small and dense mining sites (Figures b and c); as can be seen from Table 4, the object-oriented random forest performs well with an overall accuracy of 93.03%. In order to assess the accuracy of the results at a finer scale, nine sub-regions were randomly selected to observe the degree of overlap between the actual mine area and the detected mine boundary by overlaying Google HD images. (1) Google images are not used in segmentation, so the spatial extent of the mines does not exactly match; (2) for small mines, due to the relatively insignificant feature changes, the boundaries are easily detected. However, the algorithm can generally distinguish the mine area from other features, and the classification results obtained can better capture the shape of the mine area, and can more accurately detect the location of the mine area, including the accumulation area and excavation area.

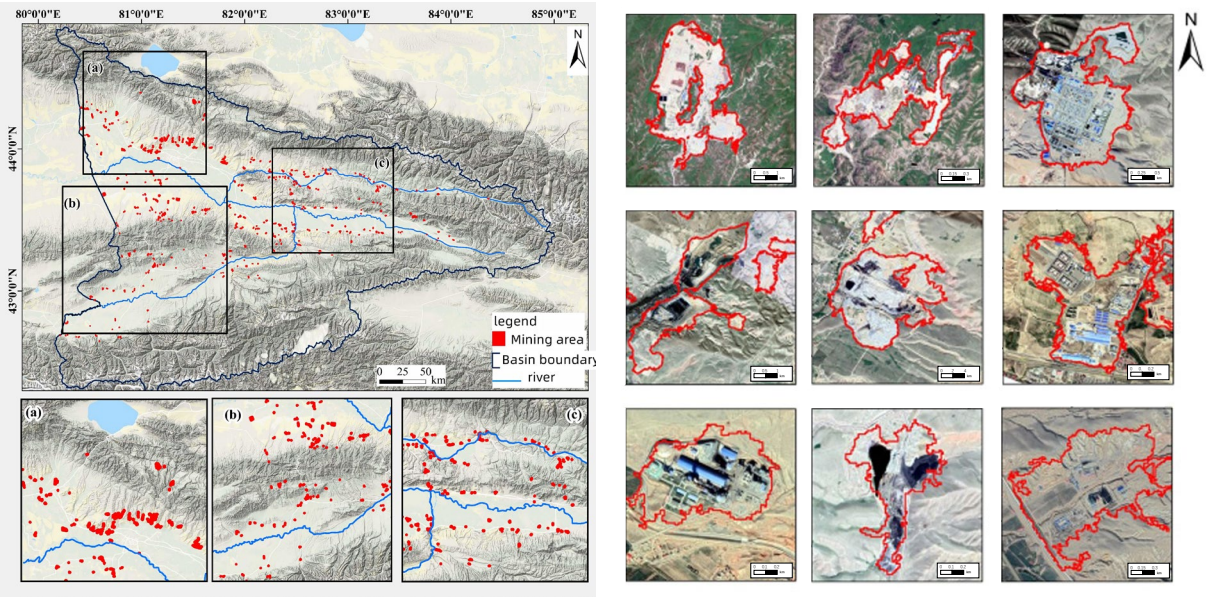

Fig.2 Overall distribution and local display of the mining areas

Tab.4 Verification table of precision of mining area detection

		Forecast classification	
		Non-mine mining area	Mining area
Actual classification	Non-mine mining area	872	35
	Mining area	53	303
Accuracy rate(ACC)		93.03%	
Precision rate(PPV)		89.64%	
Sensitivity/Recall Rate(TPR)		85.11%	
F-value		87.31%	

4.2 Quantitative characteristics of mining area

According to statistics, there are 395 mining areas in the Yili River basin, with a total area of 143.220 km2 and an average area of 0.363km2. According to the area occupied by mining areas, they are divided into four classes (Figure 3, Table 5): the largest number of mining areas (190) are in

the area of 0-0.5 km2, accounting for 48.1% of the overall number of mining areas, the smallest number (10) are >2.0 km2, accounting for only 2.5% In terms of area, the largest proportion is still 0.1-0.5 km2, amounting to 31.2%, with an average area of 0.235 km2, and the smallest proportion is <0.1 km2, with only 5.2%, with an average area of 0.056 km2. In summary, the distribution of mining areas in the basin is characterized by the majority of small and medium-sized mining areas and the random distribution of large mining areas.

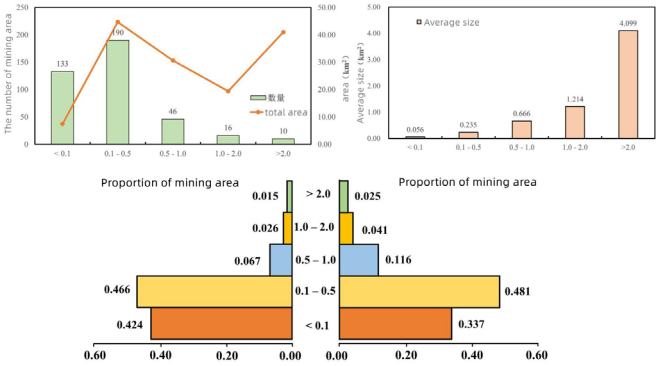

Fig.3 Comparison maps of mining numbers and areas

Tab.5 Statistical table of the number of mining areas

	The number of mining area	Number share	Total area /km^2	Area share	Average area/km^2
< 0.1 km^2	133	0.337	7.491	0.052	0.056
0.1 – 0.5 km^2	190	0.481	0.116	0.312	0.025
0.5 – 1.0 km^2	46	0.116	30.635	0.214	0.666
1.0 -2.0 km^2	16	0.041	19.420	0.136	1.214
> 2.0 km^2	10	0.025	40.986	0.286	4.099

In this paper, according to the characteristics of mineral resources in mining areas and their uses, they are divided into metal mining areas, non-metal mining areas and coal mining areas (energy mining areas). From Figure 4, Table 6, it can be seen that most of the mining areas in the Yili River basin belong to non-metallic mining areas, accounting for 82.03%, with a total mining area of 57.31 km2, which is located in the second, with an average area of 0.177 km2, indicating that the non-metallic mining areas in the mining areas of the Yili River basin are small in area but large in number and prone to a large number of small private industrial enterprises, for metal and coal mining areas, a total of 71, accounting for 17.97%, the total mining area accounts for 59.98%, the average area in 0.883 km2 and 1.776 km2.

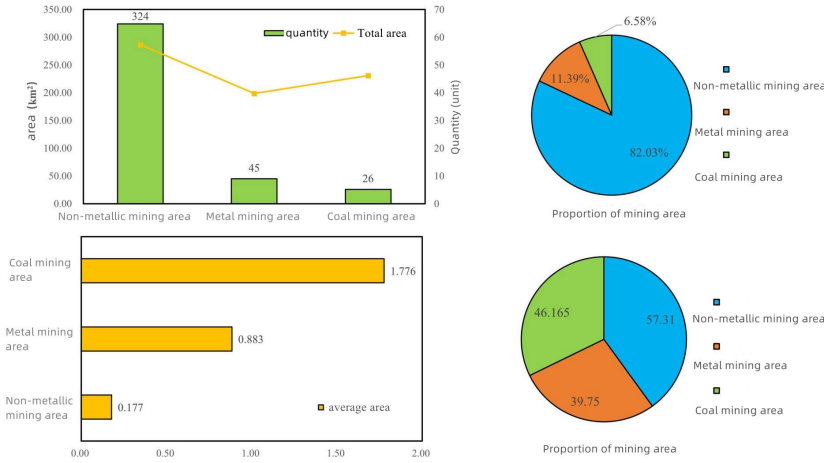

Fig.4 Comprehensive comparison diagrams of different types of mining areas

Tab.6 Statistical table of different types of mining areas

	Total area /km^2	Average area /km^2	Percentage of area	Quantity/pc	Quantity share
Non-metallic mining area	57.31	0.177	40.02%	324	82.03%
Metal mining area	39.75	0.883	27.75%	45	11.39%
Coal mining area	46.17	1.176	32.23%	26	6.58%

4.3 Spatial characteristics of the mine site

4.3.1 Spatial aggregation characteristics.

The overall distribution of mining areas in the Yili River basin is uneven, gradually decreasing from west to east (Figure 5), with the Yili River as the centerline, and a decreasing trend from the middle to both sides, where the densely distributed areas are in the middle and lower reaches of the Yili River, including the southeast area of Huocheng County, Yining City, the southwest area of Yining County and the north area of Chabchal County; the downstream area of the Kashgar River, mainly the southwest area of Nilek County; the Turks River downstream, including the northern part of Turks County and the northwestern part of Gongliu County. In addition, there are two obvious clusters in the distribution of mining areas in the Yili River basin: the junction of "Huocheng County - Yining City - Yining County - Chabchal County" and the junction of "Nilek County - Gongliu County". mining areas, and Huocheng County has the largest number of metal mining areas. In summary, the spatial distribution of mining areas in the Yili River basin is "group": uneven overall distribution, high concentration in the local area, horizontal continuity and vertical succession.

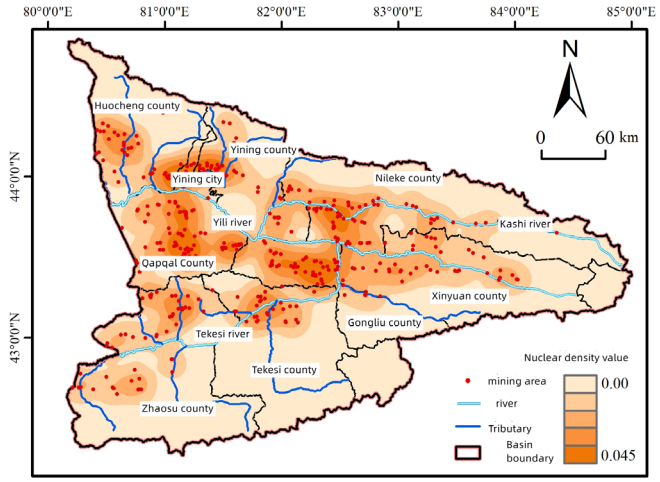

Fig.5 Distribution of kernel density in Yili River basin

Tab.7 Distribution of mining areas in counties of Yili River

	Number of mining area	Number of metal mining sites	Number of non-metallic mine sites	Number of coal mining areas
Yining City	14	4	8	2
Huocheng County	39	8	25	6
Yining County	34	6	23	5
Chabchal County	74	6	61	7
Nilek County	62	5	53	4
Xinyuan County	43	4	39	0
Gongliu County	51	5	45	1
Turks County	28	4	24	0
Zhaosu County	50	3	46	1

4.3.2 Elevation and slope characteristics.

The average elevation of mine site distribution is 1177.67m and the average slope is 6.22°, which is lower than the average elevation of 2055.81m and the average slope of 14.74° in the Yili River basin (Figure 6). Among them, 78.9% of the mining sites are located in the altitude range of 0-1572m, while 90.6% of the mining sites are located in the slope range of 0-15°, i.e. the distribution of mining sites in the Yili River Basin tends to gather in the middle and low altitude gentle terrain in general. Among them, the average elevation of metal mining areas is 1188.42m and the average slope is 6.41°, while the average elevation of coal mining areas is 1210.83m and the average slope is located at 6.58°, which is 10.75m, 33.16m and 0.19°, 0.36° higher than the average. The results show that the overall elevation of the distribution of metal mining areas and coal mining areas is higher and the slope is larger, so for these industrial enterprises it is more important to prevent and control the arrival of natural disasters such as landslides and mudslides.

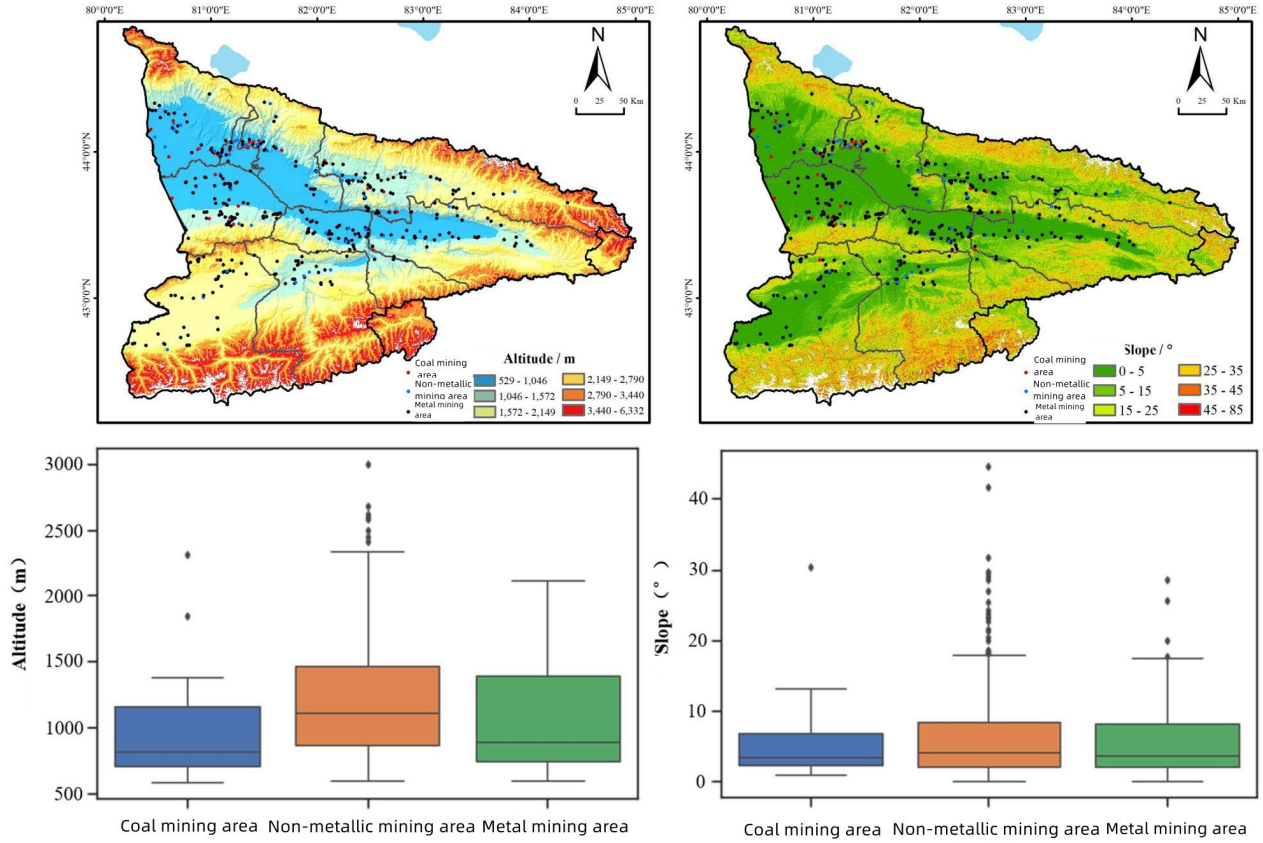

Fig.6 Distribution of mining areas at different elevations and slopes in the Yili River Basin

4.3.3 Distance of the mine site from the river.

The overall spatial distribution of mining areas in Yili River basin is greatly affected by topography, resulting in their tendency to cluster in the flat terrain around rivers and mountains. It can be seen from FIG. 7 and Table 8 that: The distance from the mining area to the river is mostly 8646.83 to 14090.78m, with an average distance of 9364.69m. The distance from the non-metallic mining area to the river is 9716.17m, the distance from the metal mining area to the river is 7523.01m, and the distance from the coal mining area to the river is 8185.75m. The distance from the metal mining area to the river is the closest. Because of its waste polluting mine production life is also the most serious of the three types of mining area, so once the waste into the river, the river have a huge pollution is beyond doubt, so in the daily production management, mine head should not only strictly on slag processing, and related departments should also focus on mining engineering safety, In order to prevent the occurrence of secondary geological disasters, mining area more serious river ecological pollution.

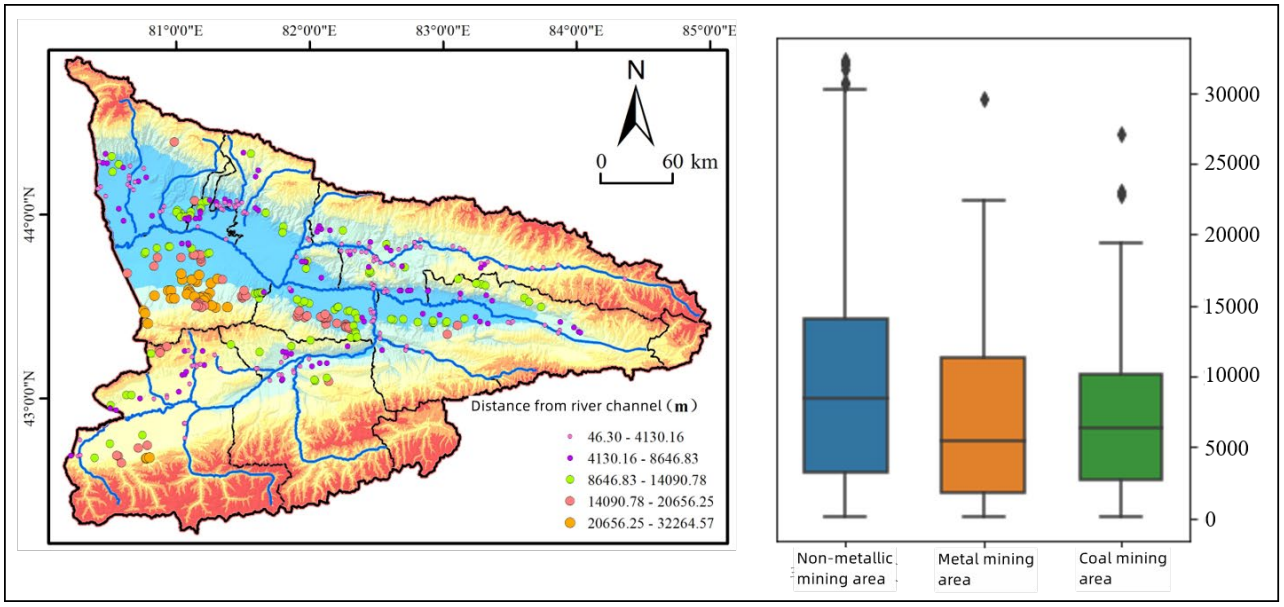

Fig.7 Diagram of distance distribution between mining area and river

Tab.8 Distance between different types of mining areas and rivers

	Non-metallic mining area	Metal mining area	Coal mining area	The overall mining area
Mean distance from the river(m)	9716.17	7523.01	8185.75	9364.69

5. Discussion

5.1 Influence of sample point and segmentation scale on prospecting accuracy in mining area

In mining area identification, the quantity and proportion of sample bank will affect the RF simulation results. In this study, considering the sample size and proportion, four sample schemes as shown in Table 9 were established, and OOB error [21] was used as an evaluation index to illustrate the impact of samples on the model. Specifically, the difference between sample i and sample ii lies in the difference of the total amount, and the difference between sample ii and sample iii lies in the difference of the number of non-mining areas. The OOB errors obtained are 13.12% and 9.03%, respectively. 8.71%, it can be seen that the prediction accuracy of the model increases with the increase of the number of sample points, but compared with sample iv and sample ii, the OOB error increases. The difference between them lies in the difference of sample proportion. Therefore, the proportion of mining and non-mining samples should also be considered when setting up the sample library [22]. Therefore, to sum up, sample iii is finally selected as its sample library. Therefore, in the future work, we can choose to adjust the distribution of sample points in the training set to further improve the accuracy of mining detection results.

Tab.9 Different sample point combinations and OOB errors

sample	Number of mining sites	Number of non-mining mining sites	total	OOB Error (%)
sampleI	226	450	676	13.12
sampleII	522	1178	1700	9.03
sampleIII	522	1478	2000	8.71
sampleIV	616	1109	1725	11.50

Under the condition that the shape and compactiity parameters were consistent, the segmentation scale parameter [23] was changed, and three different schemes were set: 50, 100 and 300. The total

amount of corresponding objects was 82.29×105, 22.21×105 and 2.49×105, and the OOB error was 9.46%, 9.46% and 9.62% (Table 10). The number of objects generated by scale 50 is about four times that of scale 100, but the OOB errors of scale 50 and scale 100 are almost the same, so the relatively appropriate scale parameter 100 is chosen.

Tab.10 Parameter list of different segmentation scales

Segmentation scale	Total number of segmented objects	Shape	Compactness	OOB
50	82.29 x 10^5	0.1	0.5	9.46%
100	22.21 x 10^5	0.1	0.5	9.46%
300	2.49 x 10^5	0.1	0.5	9.62%

5.2 Random forest parameter optimization

The same sample base was applied to linear regression (LR), support vector machine (SVM), decision tree (DT) and random forest (RF), and the ROC-AUC curve was used to characterize the mining detection performance of the model under the premise that the initial parameters were not changed (FIG. 8). The results showed that: Decision tree has the worst simulation ability and the lowest accuracy (AUC =0.85), while linear regression and random forest have similar performance, and combined with the stability and computational efficiency of the model, random forest model performs better.

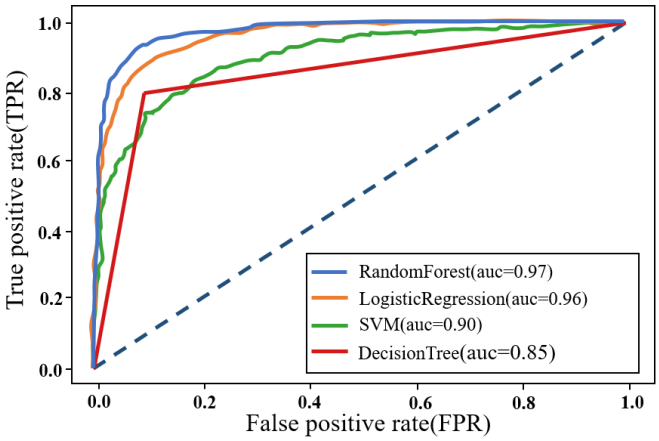

Fig.8 ROC-AUC curves of different models

In the optimization of random forest model parameters, grid search cross validation (k=10) is usually used to perform repeated combinations of different parameter values. The results in Figure A-C show that the MF parameter performs well when the number of features is base 2 (Log2), indicating that the construction of sub-decision tree with more sparse features is conducive to random forest classification. The smaller the MSL value and MSS value (MSL=10, MSS=60), the higher the accuracy of random forest. When the NST value is kept constant, the higher the MF value is (MF=5), the higher the random forest accuracy is. When NST is in the range of 1-10, the accuracy of the classification model fluctuates. When the value is in the range of 10-100, the error rate of random forest model decreases faster, and the error rate of random forest model decreases faster in NST> After 100, the accuracy of the model does not improve much, and after 1000, the accuracy tends to be stable. Therefore, to ensure the credibility of the model prediction, 1000 is chosen as the NST value. For the criteria supported by the decision tree (FIG. 4D), the entropy approach is more accurate than the Gini coefficient approach.

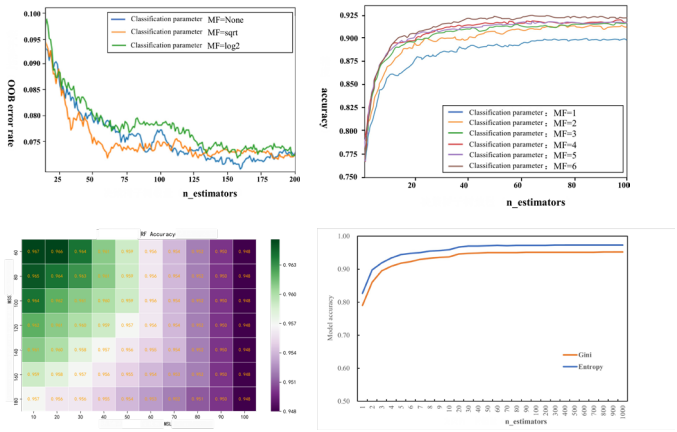

Fig.9 Random forest parameters optimization

5.3 Importance assessment of characteristic variables

Random forest uses OOB error to calculate the importance of characteristic variables, which is represented byX^j importance$V(X)^j$[23,24]:

$$V(X)^j = \frac{1}{N}\sum_{t=1}^{N}\left(e_t^j - e_t\right) \tag{3}$$

Where e_t is the out-of-pocket error of out-of-pocket data estimation, e_t^j is the new out-of-pocket error after randomly changing the value of the jth characteristic variable X^j of out-of-pocket data.The greater the out-of-pocket error caused by the variable, the lower the accuracy, and the greater MDA[25], indicating that the X^j variable is more important.In this paper, the method of cross-validation is still adopted, and the MDA value is calculated 20 times repeatedly. Finally, the results are arranged from the largest to the smallest, and the top 11 feature variables are selected (Table 11). In the table it can be seen that the mining area and the terrain, has greater correlation spectrum and texture feature, including the factors associated with topographic features were selected, and in all eight spectral characteristics has six sorts, in contrast, texture features only the standard deviation of the mining area to identify related, therefore, in the large-scale mining area detection, spectrum and terrain characteristics plays an important role, Texture and geometric features are discarded due to low correlation.

Tab.11 Features importance

Characteristics of the type	Characteristics of the name	MDA
Spectral characteristics	Max_diff	56.011
Terrain features	Elevation	47.135
Spectral characteristics	NDVI	36.280
Spectral characteristics	NDWI	31.793
Spectral characteristics	Blue	26.752
Spectral characteristics	NIR	24.931
Terrain features	Slope	24.563
Spectral characteristics	Brightness	24.146
Terrain features	GLCM.StdDev.90°	20.153
Terrain features	GLCM.StdDev.all.dir	17.492
Terrain features	GLCM.StdDev.45°	14.328

6. Conclusion

Mineral resources are the effective support of human life, but the economic development of mining areas is easy to conflict with environmental protection. Therefore, it is of great significance

to explore an effective and accurate algorithm to identify the location of large-scale mining areas, which is important for the formation of mine database and the investigation of illegal mining activities in mining areas. In this paper, the Yili River basin in Xinjiang is taken as the study area, and 60 spectral, topographic, texture and geometric classification features are constructed based on sentinel and topographic data. Then, the distribution of mining areas in the basin is obtained by object-oriented random forest algorithm. The results show that: in the object-oriented random forest algorithm, the "salt and pepper effect" can be weakened, and the mining area has a large correlation with the spectrum and topography, and only the standard deviation feature is related to the texture feature. There are 395 mining sites in the Yili River Basin, with a total area of 143.220km2, which are concentrated in the middle and low altitudes and close to the river channel, reflecting the "group nature" in space. In terms of the number of mining areas, the area of 0-0.5km2 is the majority, and non-metallic mining areas are the most abundant, and metal mining areas are the closest to the river channel. In the future research work, the algorithm can be further optimized to extract the mining areas of the country and form a richer mine database.

References

[1] Sun Z H, Xie X D, Wang P, et al. Heavy metal pollution caused by small-scale metal ore mining activities: A case study from a polymetallic mine in South China[J]. Science of the Total Environment, 2018, 639: 217-227.

[2] Gallwey J, Roblati C, Coggan J, et al. A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining[J]. Remote Sensing of Environment, 2020, 248, 111970.

[3] Wu Q, Song C, Liu K, et al. Integration of TanDEM-X and SRTM DEMs and spectral imagery to improve the large-scale detection of opencast mining areas[J]. Remote Sensing, 2020, 12: 1451.

[4] Soulard C E, Acevedo W, Stehman S V, et al. Mapping extent and change in surface mines within the United States for 2001 to 2006[J]. Land Degradation and Development, 2016, 27(2): 248-257.

[5] Katherine L C, Connette G, Bernd A, et al. Assessment of mining extent and expansion in Myanmar based on freely-available satellite imagery[J]. Remote Sensing, 2016, 8, 912.

[6] Atriyon J. Mining land identification in Wetar Island using remote sensing data[J]. Journal of Degraded and Mining Lands Management, 2018, 6, 1513-1518.

[7] Pagot E, Pesaresi M, Buda D, et al. Development of an object-oriented classification model using very high resolution satellite imagery for monitoring diamond mining activity[J]. International Journal of Remote Sensing, 2008, 29(2), 499-512.

[8] Demirel N, Emil M K, Duzgun H S. Surface coal mine area monitoring using multi-temporal high-resolution satellite imagery[J]. International Journal of Coal Geology, 2011, 86(1), 3-11.

[9] Song X L, He G J, Zhang Z M, et al. Visual attention model based mining area recognition on massive high-resolution remote sensing images[J]. Cluster Computing 2015, 18(2), 541-548.

[10] Li J, Zipper C E, Donovan P F, et al. Reconstructing disturbance history for an intensively mined region by time-series analysis of Landsat imagery[J]. Environment Monitoring and Assessment, 2015, 187(9): 4766.

[11] Sen S, Zipper C E, Wynne R H, et al. Identifying revegetated mines as disturbance/recovery trajectories using an interannual Landsat chronosequence[J]. Photogrammetric Engineering Remote Sensing, 2012, 78(3): 223–235.

[12] Townsend P A, Helmers D P, Kingdon C C, et al. Changes in the extent of surface mining and reclamation in the Central Appalachians detected using a 1976–2006 Landsat time series[J]. Remote Sensing of Environment, 2009, 113(1): 62–72.

[13] Wang Yujuan, Guo Dongmei. Eco-environment Evolution and its Driving Forces in Yili River Basin of China-Kazha Border River[J]. Journal of Eurasian Economy, 2016, 7(4): 100-128.

[14] Gao Yaming. Estimation of Ecological Service Function Value of Yili River Basin in Xinjiang[D]. Urumchi: Xinjiang Normal University, 2017.

[15] Song Jiaren. Annals of Kazak Autonomous Prefecture of Yili[M]. Xinjiang: Xinjiang People's Publishing House, 2004: 08-155.

[16] Dragut L, Csillik O, Eisank C, et al. Automated parameterisation for multi-scale image segmentation on multiple layers[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 88: 119-127.

[17] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1973, SMC-3(6): 610-621.

[18] Breiman L. Random Forests[J]. Machine Learning, 2001, 45, 5–32.

[19] Zhou Tianning, Ming Dongping, Zhao Rui. Land Cover Classification based on Algorithm of Parameter Optimization Random Forest[J]. Science of Surveying and Mapping, 2017, 42(2): 88-94.

[20] Bekkar M, Djemaa H K, Alitouche T A. Evaluation measures for models assessment over imbalanced datasets[J]. Journal of Information Engineering and Application, 2013, 3: 27–38.

[21] Puissant A, Rougier S, Stumpf A. Object-oriented mapping of urban trees using Random Forest classifiers[J]. International Journal of Applied Earth Observation and Geoinformation, 2014, 26: 235–245.

[22] Duro D C, Franklin S E, Dube M G. Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests[J]. International Journal of Remote Sensing, 2012, 33, 4502–4526.

[23] Stumpf, A, Kerle N. Object-oriented mapping of landslides using Random Forests[J]. Remote Sensing of Environment, 2011, 115: 2564–2577.

[24] Van B, Comber A, Lamb A. Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data[J]. Remote Sensing of Environment, 2014, 149: 118-129.

[25] Zhu Z, Woodcock C E, Rogan J, et al. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peril-urban land cover classification using Landsat and SAR data[J]. Remote Sensing of Environment. 2012, 117: 72-82.