Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

Compiling a Preliminary Regional Geological Map Using Sentinel-2 Satellite Imagery and the Random Forest Algorithm in the East of Iran

Hamid Geranian* and Mohammad Amir Alimi

Department of Mining Engineering, Birjand University of Technology, Birjand, Iran

Article Info

Received 16 September 2024

Received in Revised form 23 October 2024

Accepted 8 November 2024

Published online 8 November 2024

DOI: 10.22044/jme.2024.15083.2882

Keywords

Geological map

Random forest algorithm

SMOTE algorithm

Image processing

Sentinel- 2 image processing

Abstract

This study employs Sentinel-2 satellite images along with the random forest algorithm to create a regional geological map. For this purpose, the input data include the images for 10 Sentinel-2 bands of the Khosuf-I region, while the class labels include a geological map of Khosuf-I divided into three and fifteen rock units. The classification accuracy of the resulting model is 90.97 and 84.85% for the three-class training and testing data, and 94.76 and 63.92% for the fifteen-class training and testing data, respectively. These models are then applied to the Sentinel-2 satellite images' data of the Birjand-IV region to prepare two preliminary geological maps. The Birjand-IV region's three-class geology map reveals that igneous rocks are present in the northern and southern regions, while sedimentary rocks occupy the middle section and metamorphic rocks are found within the region's igneous masses. Similarly, the fifteen-class geology map of Birjand-IV indicates that andesite, dacite, intermediate tuff rock units, and metamorphic rocks characterize the northern region. Conversely, the southern part of the region is mainly composed of ophiolite, flysch sediments, basaltic and ultra-basic volcanic rocks, and limestone and shale interlayers. Field studies in three areas confirm the accuracy of the preliminary geology maps.

1. Introduction

Techniques for mineral potential mapping include geological information and maps, geochemical, geophysical, remote sensing, and exploration drilling data analysis [1,2]. Preparing regional-scale geological maps for exploration is a crucial task for the government [3] and the Geological Survey and Mineral Exploration of Iran have been actively working on creating these maps for over half a century at 1:250,000 and 1:100,000 scales. In recent years, the organization has included the creation of maps at scales of 1:50,000 and 1:20,000 in their agenda. However, given the small coverage area of these maps and the large quantity required, it is necessary to use new technologies to reduce the time and cost of preparing each map. Techniques such as geological information and maps, geochemical analysis, geophysical analysis, remote sensing, and exploration drilling data analysis can all contribute to mineral potential mapping. One potential technology for creating preliminary geological maps is using satellite images with appropriate spatial resolution, which can replace previously employed aerial photographs.

Satellite images are widely used in exploration projects due to their high spatial and spectral resolution, availability, the existence of numerous image processing software, cost-effectiveness, and speed of use [4-7]. Satellite images can be used to take images of the study area, highlight line structures, determine rock units, separate alteration zones, and identify promising mineral areas [8-12].

Rock units can be identified using the relationship between various satellite image bands and rock units. In general, the techniques for identifying and separating rock units using satellite

Corresponding author: h.geranian@birjandut.ac.ir (H. Geranian)

images fall into two categories: (1) spectrum matching and (2) subpixel methods [13-16]. Spectrum-matching techniques determine the degree of spectral similarity between the reference and image spectrum. The reference spectrum can be from a library or a field spectrum. Spectral Angle Mapper (SAM), Spectral Feature Fitting (SFF), and supervised classification methods, such as Logistic Regression (LR), Neural Networks (NN), Support Vector Machine (SVM), Random Forest (RF), and Discriminant Analysis (DA), have been used in this category [17-26]. Under certain geological conditions, such as geological simplicity, ample sample quantity, and a pure spectrum, these methods can generally produce satisfactory results [14,16,18,27].

Subpixel techniques involve combining several satellite image bands to extract pure spectral pixels with unique characteristics from each image. This process is similar to the transformation methods used in multivariate data sets, which can remove redundancy in spectral information and reduce computational needs for further processing. Principal Component Analysis (PCA), Independent Component Analysis (ICA), Minimum Noise Fraction (MNF), and band ratios are among subpixel techniques that different researchers have used to identify rock units [11,25,28-35]. However, they are most effective when used in a small area with appropriate validation methods [11,14,25,33].

In recent years, the Sentinel series images, particularly Sentinel-2, with higher spatial resolution and more bands than Landsat and Aster images, have developed better applications in mineral exploration, especially in identifying iron oxide minerals and geological mapping [25,36-41]. This paper aims to create a preliminary geological map with a scale of 1:50,000 using Sentinel-2 images, which can assist the Geological Survey and Mineral Exploration of Iran in reducing costs and time. Another goal of this paper is to investigate the possibility of modeling and separating rock units in an area with a high number of rock units using a supervised learning method such as an RF algorithm with a high number of trees. Although a classification algorithm performance can vary significantly depending on the dataset and modeling conditions, but the studies shown that the RF algorithm is straightforward to train, computationally efficient and highly stable with respect to variations in classification model parameter values. This algorithm can also have high accurate like the SVM and NN algorithms [18]. Since geological maps with this scale are prepared with an accuracy of 0.2 mm [42], equating to 10 m in the field in the 1:50,000 scale map, Sentinel-2 satellite images were selected due to their spatial resolution of 10 m and possessing more bands than Aster and Landsat satellites in the VNIR range. The paper used a supervised classification method, a 1:50,000 scale geological map, and Sentinel-2 satellite images to build a model, which was then used to prepare the preliminary geological map of the surrounding areas.

2. Geological setting of the study district

2.1. Geological map status in the South Khorasan province

As part of its organizational duties, the Geological Survey and Mineral Exploration of Iran has been tasked with preparing geological maps for regional explorations. For the past fifty years, this organization has primarily focused on creating small-scale geological maps, which include (1) quadrangle maps with a scale of 1:250,000 by 1.5° of longitude and 1° of latitude with an area of approximately 15,000 km^2 and (2) sheet maps with a scale of 1:100,000 by 0.5° of longitude and 0.5° of latitude with an area of about 2500 km^2. Iran is covered by 121 quadrangle and 659 sheet geological maps, which cover approximately 90% and 80% percent of Iran, respectively, to date [43]. However, in recent years, the Geological Survey and Mineral Exploration of Iran has shifted its attention to preparing medium-scale geological maps with scales of 1:50,000 and 1:20,000. The initial phase entails creating 1:50,000 geological maps that cover 2662 areas throughout Iran.

The province of South Khorasan, covering an area of 150,800 km^2, is located east of Iran (Figure 1a). This province, which produces 40 ore minerals, is the fifth most mineral-rich province in Iran. Although this province has approximately 2.5 billion tons of geological resources, representing approximately 10% of Iran's mineral resources [44], there is no world-class deposit due to the absence of systematic exploration activities. The Geological Survey and Mineral Exploration of Iran has created geological maps at a 1:50,000 scale for the province of South Khorasan to address this issue. Figure 1b depicts the quadrangle and sheet maps of this province. The geological data for the 1:50,000 scale Khosuf-I map has been collected and is currently being printed. It was used as a model for preparing the preliminary geological map of Birjand-IV at 1:50,000 scale using Sentinel-2 images and a classification algorithm (Figure 1c). Figure 2 depicts the Sentinel-2 images of the districts on these two maps. All satellite images and geological maps in this paper are drawn with ArcMap 10.8.2 software.

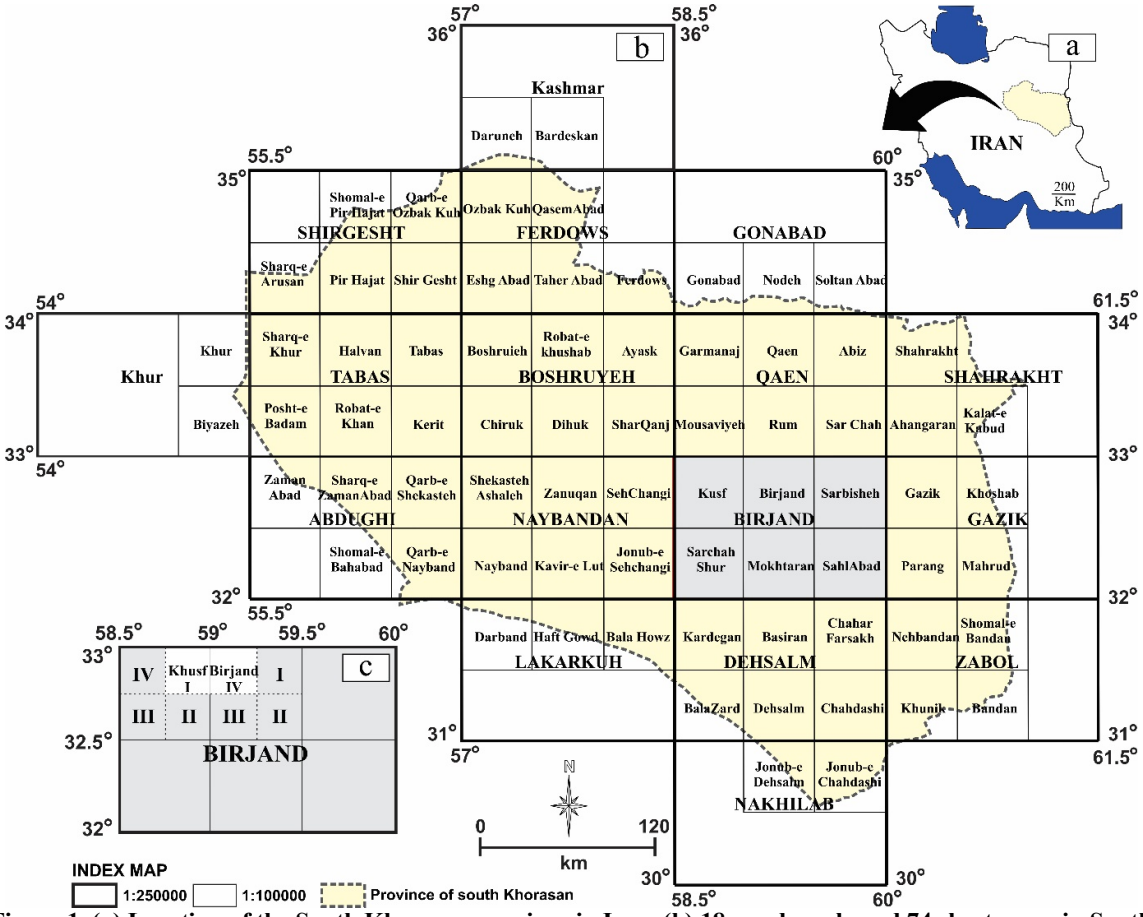

Figure 1. (a) Location of the South Khorasan province in Iran, (b) 18 quadrangle and 74 sheet maps in South Khorasan province and (c) 1:50,000 scale Khosuf-I and Birjand-IV maps in the 1:250,000 scale Birjand map.

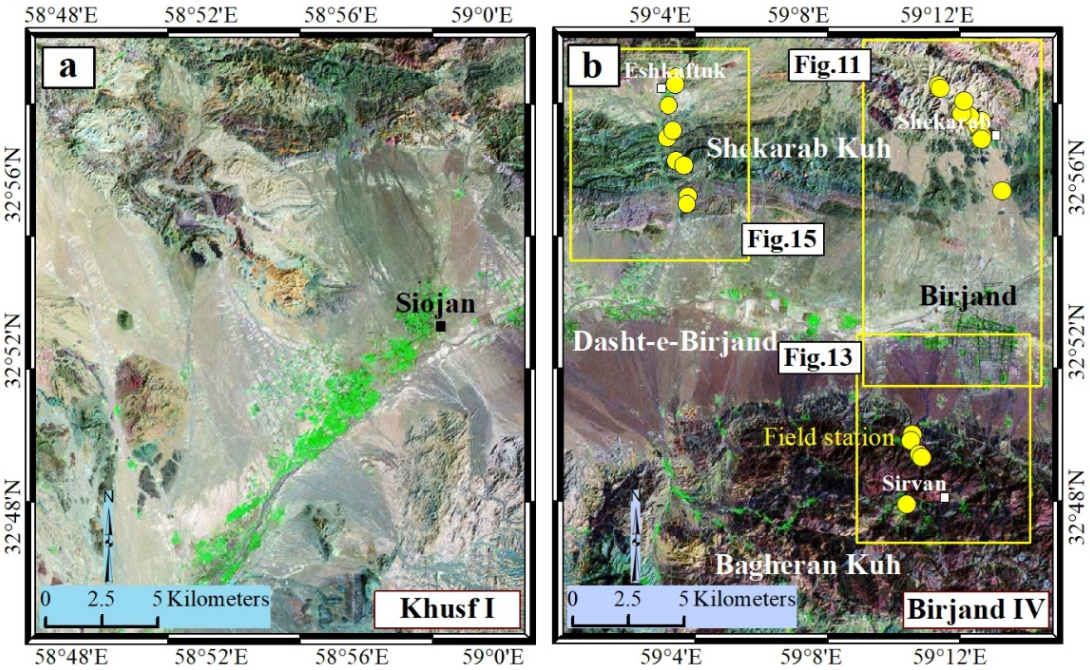

Figure 2. Satellite images from Sentinel-2 of the study area: (a) Khosuf-I map, (b) Birjand-IV map (Images are RGB color composite of the bands 4, 3, and 2 and yellow points and boxes represent field stations that described in section 6).

2.2. Tectonic and geology of the study district

The Eastern Iranian range, forms a Z-shaped sigmoid that it has extend between the two Lut and Afghan blocks. The north eastern curvature of the orogen in Birjand area has penetrated into the Lut block in an interfinger pattern (Figure 3). The South Khorasan province is structurally situated in the Flysch basin of eastern Iran, specifically in the Lut and Tabas blocks Flysch facies include a sequence of fine-grained sediments such as shale, marl, mudstone, and siltstone with coarse-grained sediments such as sandstone or detrital limestone, which is introduced as a tectonic facies concurrent with orogeny and associated with deep waters. Due to the vast size of the study area, the Cretaceous and Tertiary lithologies are incredibly diverse. Thick flysch deposits, sandstone, and shale accumulations cover the ophiolitic basement associated with the Young Tethys oceanic crust. This region of flysch-melange ophiolite is known as the Sistan suture zone [45,46].

The study area is characterized by tectonic activity, specifically the presence of deformed accretionary wedges belonging to the Ratuk, Neh, and Sefidabeh forearc basin complexes. These wedges extend from Birjand to Zahedan and are situated in front of a Neo-Tethys branch (Figure 3). The Ratuk complex is an older part of the accretionary wedges and can be found along the northwest trend. It forms the underlying part of the Sefidabeh basin and contains fragmented ophiolites with Cenomanian to Maastrichtian pelagic sediments and meta-volcanic rocks that have undergone amphibolitic facies metamorphism [46]. The Neh Complex includes Cretaceous ophiolitic melange, marine sediments, and turbiditic rocks that date back to the Cenomanian to Paleocene period and have undergone low-grade facies metamorphism [46]. The ophiolites of this complex can be seen in areas such as the curved splay Birjand (Bagheran Mountain), and parts of them are mixed with the Lut block due to tectonic activity. The Sefidabeh forearc basin covers the Ratuk and Neh complexes and the southwestern margin of the Afghan block. The basin deposits include clastic-destructive formations and carbonate rocks that date back to the Cenomanian to Eocene period (Figure 3).

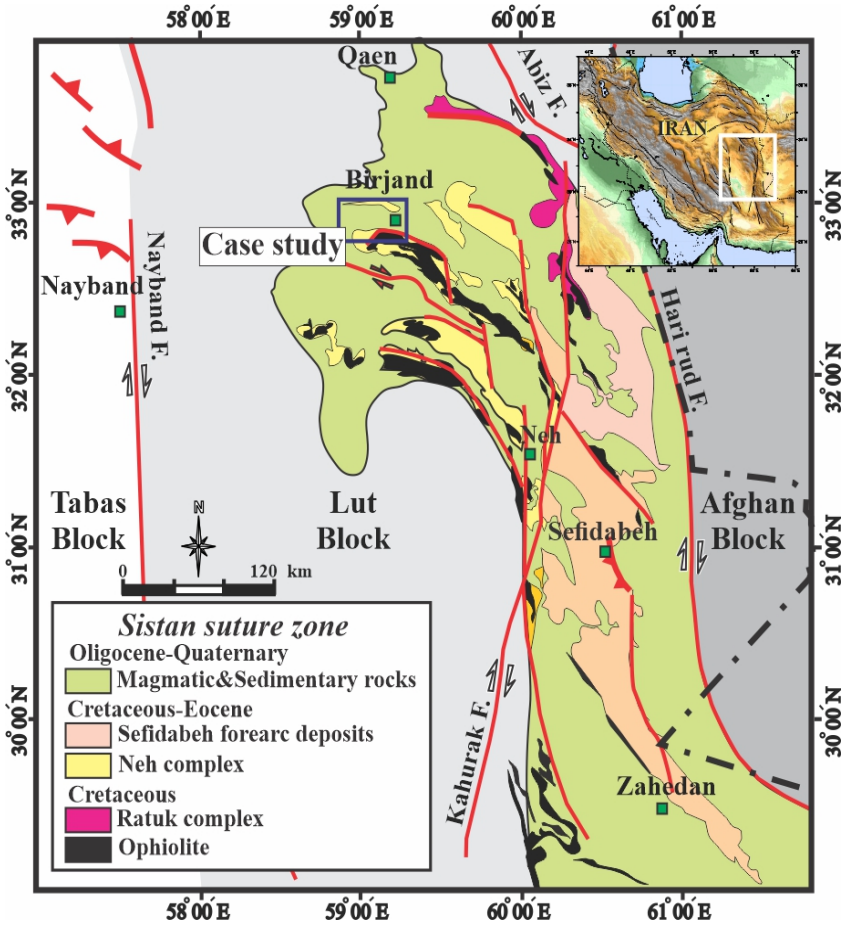

Figure 3. Distribution map of lithologic units in eastern Iran.

In the tectonic development of the study district, the rifting between the Afghan and Lut continental blocks during the Early to Middle Cretaceous period resulted in the formation of the Sistan Ocean in the study district [47-49]. However, there is still some debate among researchers regarding how and when the ocean was closed. Various models have been proposed, including subduction under the Afghan block in the east [45,46], subduction under the Lut block in the west [49,50], and two-way subduction [51]. Ultimately, the ocean closed, and the Lut and Afghan blocks collided during the Late Cretaceous [49,52] to the Middle Eocene [45,46]. The Middle Eocene-Early Oligocene was marked by extensive calc-alkaline magmatism in eastern Iran, indicating an extensional post-collisional period [50]. This extensional regime resulted in lithospheric thinning that expanded the alkaline basalts from the Middle Miocene to the Quaternary [50,53].

2.3. Rock units of the Khosuf-I geological map

The Khosuf-I region is situated in eastern Iran, bordering Birjand city to the west. It is surrounded by Qain, Birjand, Sarchah Shur, and Sechangi in the north, east, south, and west, respectively (Figure 1b). This area lies in the northern part of the Central Lut block. Figure 4 depicts a geological map of the Khosuf-I region at a scale of 1:50,000 based on three rock units (Figure 4a), fifteen rock units (Figure 4b), and the percentage of coverage for each rock unit. This region has extensive exposures of sedimentary rocks ranging from the Late Cretaceous to the Early Tertiary and Cenozoic volcanic rocks.

Calcareous sandstones and silty marls with tens of meters of thickness are frequently red and occasionally green-gray in the map's north and west (Figure 4b). These flysch sediments date from the early Eocene. In the region's northernmost elevations, nomulitic limestones are exposed. Under the upper volcanic rock sequence, folded Paleogene tuffs and marls sequences can be observed. The southwestern plains and the region's center are dominated by these rocks, and the nearly horizontal Quaternary alluvial deposits are placed directly on top of them (Figure 4b). Conglomerates appeared dispersed in the north and southwest, and they may contain sand layers, tuff, brecciated tuff, siltstone, or silty shales. The andesite-tuff unit is comprised of tuff and andesite intervals that extend northward.

In addition to the previously mentioned rock units, the region of Khosuf-I also contains dacite andesite lava flows. In the southwestern region, dacites are overlain by conglomerates and igneous rocks, and they occasionally cut through Paleocene-Eocene flysch. In addition, there is a rhyolite massif to the region's west. Colorful Neogene marls are found irregularly distributed on older units, with a thickness of 80 to 100 m, and sometimes within layers of gravelly sandstones. Neogene conglomerates are adjacent to the units above in the region's north. Overlying older layers are extensive and unconformable quaternary sediments in the form of alluvial sediments, terraces, alluvial fans, and loose and uncemented alluviums of the riverbeds. The extensive outcrops of phyllites, sericite, and albite slates in the region's south indicate that the region has undergone low-grade metamorphism. The narrow northwestern sequences of the middle ophiolite belt of eastern Iran extend from Birjand to the northeastern areas of the Khosuf-I map and completely disappear from this region to the west. In the northeastern regions, red and brown listvenites and dispersed masses of ultrabasic rocks, predominantly peridotite and serpentinized gabbro, have been observed, frequently mixed and displaced by faults (Figure 4b).

3. Random Forest (RF) Algorithm

The RF algorithm, introduced by Breiman [54], is a classification method that utilizes ensemble learning. The algorithm employs multiple decision trees to form a forest and classify each sample. Each decision tree has a tree structure resembling a flowchart, where an internal node (non-leaf node) represents a variable test, and each branch denotes an outcome of the test. Each end node (leaf node) also contains a class label, k and the root node is the topmost node in a tree. Most decision tree-building algorithms adopt a top-down approach, beginning with a set of training data with class labels. Furthermore, throughout the development of the decision tree, the training data is recursively divided into smaller subsets [55-57].

There are multiple ways to generate a decision tree. The RF algorithm utilizes the CART method and the Gini index. For instance, if training data (D) is divided by a variable (A) into two subsets of D_1 and D_2, the Gini index value for such a subset is determined per [56,58]:

$$Gini_A(D) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \qquad (1)$$

where $|D_1|$, $|D_2|$ and $|D|$ denote the sample numbers of the data set D_1, D_2 and D, respectively,

and the Gini value of each subset is obtained from the following equation [56,58]:

$$Gini(D_j) = 1 - \sum_{i=1}^{m} p_i^2 \;\; ; \;\; j = 1,2 \qquad (2)$$

where p_i is the probability that a sample from D_j belongs to each class, and m is the number of classes. The variable with the lowest Gini value is chosen as the splitting or testing variable in the CART method.

When using the RF algorithm to build a decision tree, variables are selected using a random vector to determine the split and node formation. This vector is sampled independently and has the same distribution as all the trees in the forest. In addition, the training set is produced by a sampling method for each decision tree, which may result in some samples appearing more than once and others not being present at all. Each tree votes during classification, and the final result is determined by a majority vote [54,56,59].

Preventing over-fitting, improving classification accuracy, working with a data set with missing values, no need to normalize data, suitable performance with big data, identifying influential variables in classification, high generalizability and working with a class-imbalanced data set are all advantages of the RF algorithm. On the other hand, this algorithm has disadvantages such as the requirement for high-power computing, large training data, time-consumption and low interpretability [54,59]. Studies have demonstrated that the RF algorithm outperforms other classification algorithms, particularly in classifying satellite images [18]. This algorithm is employed in this paper for the reasons stated above, as well as for its superior performance when dealing with large and class-imbalanced data sets (conditions present in the study area).

4. Data Preparation

Sentinel-2 was the second satellite in the Sentinel series launched by the European Space Agency in 2015 (Sentinel 2A and Sentinel 2B). The images from this satellite have thirteen spectral bands in the visible, near-infrared and short-wave infrared ranges, a temporal resolution of five days at the equator and three days at mid-latitudes, regular global coverage of the earth's surface from 56° S to 84° N, spatial resolutions of 10, 20, and 60 m, and a field of view of 290 km [60]. Table 1 lists the band specifications for this satellite. The Sentinel 2A satellite over the study area captured the image utilized in this paper on November 26, 2021.

Table 1. Specifications of Sentinel-2 satellite image bands (https://sentinel.esa.int).

Band name	Band number	Sentinel-2A		Sentinel-2B		Resolution (meters)
		Central Wavelength (nm)	Bandwidth (nm)	Central wavelength (nm)-2	Bandwidth (nm)-2	
Coastal Aerosol	1	443.9	20	442.3	20	60
Blue	2	496.6	65	492.1	65	10
Green	3	560.0	35	559	35	10
Red	4	664.5	30	665	30	10
Vegetation Red Edge	5	703.9	15	703.8	15	20
Vegetation Red Edge	6	740.2	15	739.1	15	20
Vegetation Red Edge	7	782.5	20	779.7	20	20
NIR	8	835.1	115	833	115	10
Narrow NIR	9	864.8	20	864	20	20
Water Vapor	10	945.0	20	943.2	20	60
SWIR – Cirrus	11	1373.5	30	1376.9	30	60
SWIR	12	1613.7	90	1610.4	90	20
SWIR	13	2202.4	180	2185.7	180	20

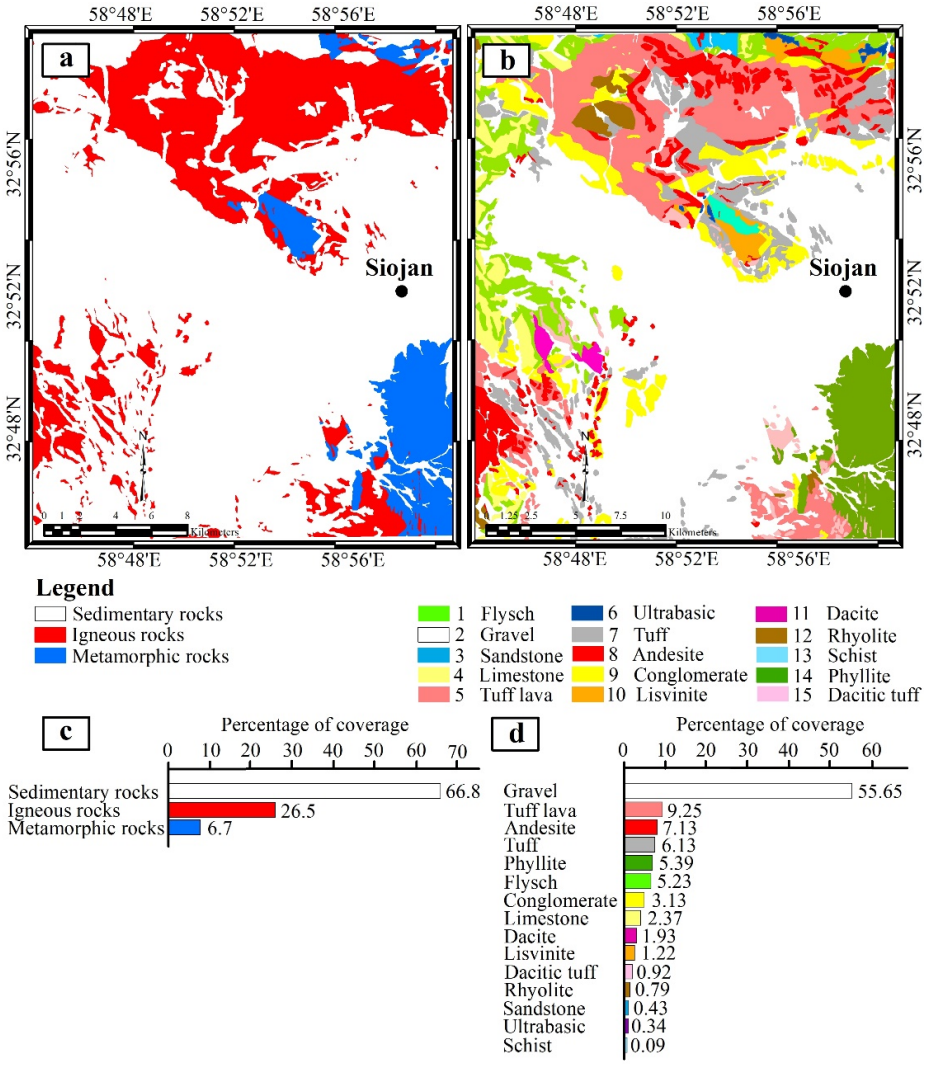

Figure 4. Simplified geological map of the Khosuf-I region at a scale of 1:50,000 based on (a) three rock units; (b) fifteen rock units and the percentage of coverage of each rock unit in the study region (the data of this map is based on 1:100,000 scale Khosuf sheet and the field surveys).

The images used in this study belong to the 1C level and have undergone atmospheric and radiometric corrections using the sen2cor tool in the SNAP software. The Resampling menu of SNAP software was then employed to resample all image bands to a spatial resolution of 10 m. The images were then mosaicked for connection, and the Khosuf-I and Birjand-IV regions were separated. To remove cloud and water vapor studies, bands 1, 10, and 11 were excluded from the data set. Finally, the primary data set was derived by extracting the numerical values of the pixels from the remaining ten band images. As previously mentioned, the resolution of each pixel is 10 meters. Therefore, rock units must occupy at least one-pixel area to be recognized in the data set. The classification model was constructed using the Khosuf-I data matrix with dimensions 6714631x10. The Birjand-IV data matrix, with dimensions 6756720x10, was used to determine the class labels and prepare a preliminary geological map.

The of Khosuf-I region's geological map was used to determine the class labels of its matrix data. The sedimentary (class 1), igneous (class 2), and metamorphic (class 3) rock types are depicted on the 1:50,000 geological map of the Khosuf-I region in Figure 4a.

According to this map, sedimentary, igneous, and metamorphic rocks cover 66.8, 26.5, and 6.7% of this area, respectively (Figure 4c). By superimposing this map (Figure 4a) with the Khosuf-I satellite image (Figure 2a), the class label of each pixel can be determined (i.e., the class label

of each row or sample of the Khosuf-I data matrix). In addition, the rock units from the geology map of Khosuf-I were utilized. The Khosuf-I geology map, in Figure 4b, is divided into 15 types of rock units, including flysch (class 1), gravel (class 2), sandstone (class 3), limestone (class 4), tuff lava (class 5), ultrabasic (class 6), tuff (class 7), andesite (class 8), conglomerate (class 9), listvenite (class 10), dacite (class 11), rhyolitic (class 12), schist (class 13), phyllite (class 14), and dacitic tuff (class 15) rocks.

In this region, the gravel rock unit has the highest coverage (55.65%), while the schist rock unit has the lowest (0.09%). Figure 4d depicts the percentage of other covered rock units. By superimposing Figure 4b with the Khosuf-I satellite image (Figure 2a), it is possible to determine each pixel's class labels based on classifying rock unit types. Therefore, two matrices containing three and fifteen classes were utilized data modeling with the RF algorithm.

5. Data Classification with the RF Algorithm

The first step in data classification is to check its significance level, which can be accomplished using the Mahalanobits distance test or analyzing variance (ANOVA) test [61,62]. Figures 5 and 6 provide valuable information regarding the central location of the classes in the Khosuf-I data matrix for three and fifteen classes, respectively. According to these figures, the spectral graphs of the rock units are distinct, and the radiometric value (reflection of electromagnetic waves) increases with increasing wavelength. Sedimentary and igneous rocks have the highest and lowest radiometric values in different bands, while metamorphic rocks fall in between [19,23,25,29,36]. Consequently, these spectral graphs can classify and separate rock units appropriately.

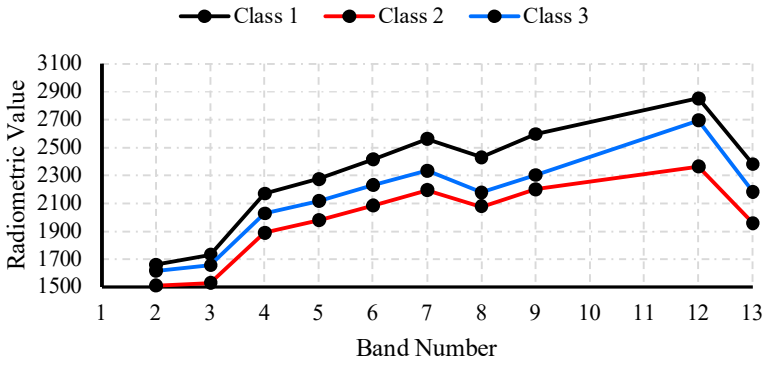

Figure 5. Distance between the centers of the classes in the three-class matrix of Khosuf-I data.

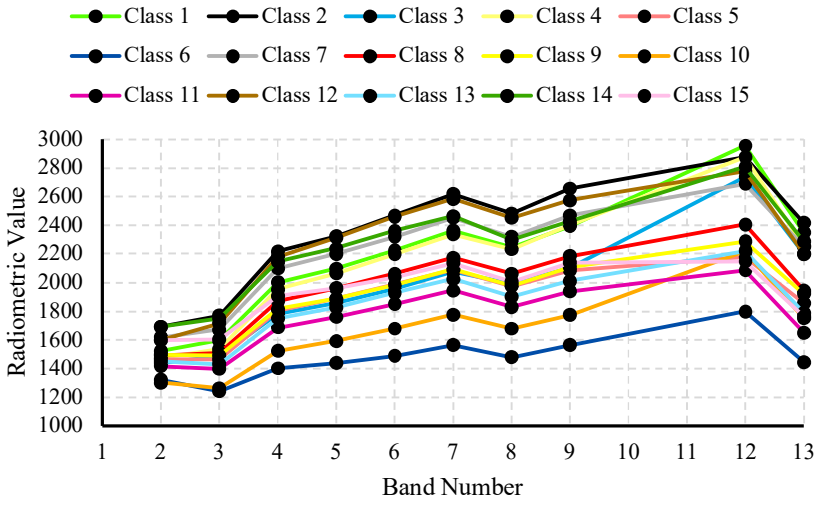

Figure 6. Distance between the centers of the classes in the fifteen-class matrix of Khosuf-I data.

The next step is to examine the number of class samples. Figures 4c and 4d illustrate the unequal distribution of class samples. For matrices with three and fifteen classes, the majority class samples are 10 and 600 times that of minority class samples, respectively. The disparity in class sample sizes will diminish the performance of the classification algorithm and consequently reduce classification accuracy [56,63]. There are two oversampling and undersampling methods for overcoming and minimizing this issue. Under the oversampling method, the number of minority class samples increases until it equals the number of majority class samples. In contrast, in the undersampling method, the number of samples from the majority class was randomly reduced until it reached the minority class [63-65].

The Synthetic Minority Over-Sampling Technique (SMOTE) is one of the most widely used approaches in this field, developed by Chawla et al. [66]. The number of minority class samples is increased, in this technique, through the following steps [64,66]:

1. Taking a variable vector consisting of the difference between *k* nearest samples around one of the samples of the minority class.

2. This vector is multiplied by a random number between 0 and 1.

3. The obtained value is added to the main variable (the primary sample of the minority class) to obtain a new sample.

The SMOTE technique produces new data of the negative class by operating in the variable space rather than the data space. This augmentation in the minority instances enhances the decision accuracy of the trained algorithms. The random forest method is an ensemble classification method that uses a set of trees to make decisions. Increasing the number of samples using the SMOTE method can increase the diversity of the RF algorithm, that will result in an increase in classification accuracy [67,68]. As a result, the SMOTE technique was used for both data matrices to balance the number of samples from all classes (the three- and fifteen-classes data matrices of Khosuf-I). The RF algorithm was then applied to them. Each matrix's data set was divided into two parts to evaluate the classification model. Subsequently, 80% of the data set was used as training data to build a model, and the remaining 20% was used as testing data to ensure the model's accuracy [56,61].

When implementing the RF algorithm, it is crucial to consider the number of trees used in the process. This number typically ranges from 1 to 100 and can be optimized by fixing the classification error rate [54,59]. Figure 7 shows the classification error rate based on the number of different trees on the three-class data matrix of Khosuf-I. In addition, the figure shows that the class error remains constant as the number of trees increases from 15 onwards. Therefore, the optimal number of trees is 15.

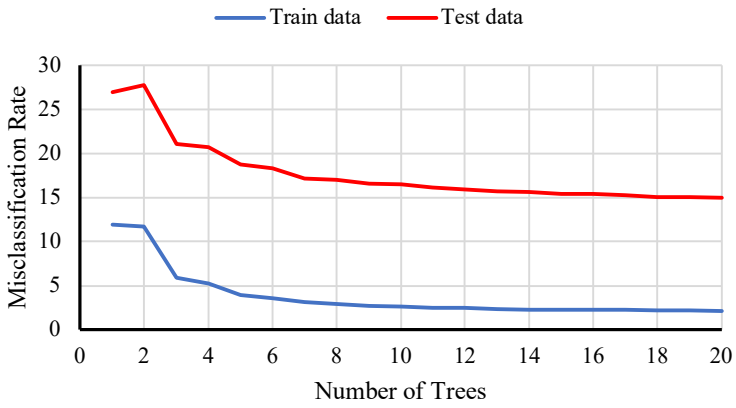

Figure 7. Classification error rate plot vs. the number of trees in the Khosuf-I three-class data matrix.

The RF algorithm with 15 trees was applied to the three- and fifteen-class training data of Khosuf-I to build the classification models. The obtained models were then utilized to determine the training and testing data classes. Table 2 displays the classification accuracy of the RF algorithm on both the training and testing data of Khosuf-I. The model achieved an accuracy of over 95% on the

training data, indicating a high level of accuracy. On the testing data of the three-class data matrix, the RF algorithm model achieved an accuracy of approximately 85%. This can be attributed to the significant differences in the spectral bands' graph of sedimentary, igneous, and metamorphic rocks, as shown in Figure 5.

The RF algorithm's classification accuracy on the testing data in the fifteen-class data matrix is acceptable (approximately 64%). The highest accuracy belongs to phyllite and dacitic tuff rock units, classes 14 and 15, respectively, while the lowest accuracy belongs to tuff and andesite rock units, classes 7 and 8, respectively. This difference in classification accuracy can be attributed to a meaningful distance between the spectral band graphs of phyllite and dacitic tuff rock units and other rock units and a distance close to this for tuff and andesite rock units and other rock units (Figure 6). Moreover, the increase in the number of classes have a little influence on reducing the classification model's accuracy of the fifteen-class data matrix.

Table 2. Percentage of correct classification of training and testing data by the RF algorithm.

Class number	Three-class data		Fifteen-class data	
	Train data	Test data	Train data	Test data
1	97.74	76.62	95.68	63.20
2	97.80	83.56	93.45	51.85
3	97.97	94.38	96.24	78.97
4			95.09	67.41
5			94.14	57.12
6			94.14	65.08
7			92.61	46.13
8			91.26	39.62
9			92.00	48.67
10			95.82	67.84
11			94.94	57.49
12			96.75	79.55
13			95.89	69.88
14			96.68	83.06
15			96.45	83.02
Total proportion	97.90	84.85	94.76	63.92
Kappa Score	0.99	0.77	0.96	0.62

In general, it can be stipulated that this algorithm's classification accuracy on testing data is also acceptable and high; this demonstrates the high generalization of these models [18,61]. However, the Kappa scores for the three- and fifteen-class data sets were also calculated (Table 2). These scores indicate that the random forest algorithm performed well on the training data and adequately on the test data (for more details, see Ref. 69). Therefore, the obtained models are applicable for the classification of Birjand-IV data.

Moreover, the sensitivity analysis method can determine the significance and influence of the variables (10 selected bands) in the classification model (for more details, see Refs. 61,70,71). The results of this sensitivity analysis on the three-class (Figure 8a) and fifteen-class (Figure 8b) training data set are depicted in Figure 8. Bands with longer wavelengths (i.e., bands 8, 9, and 12 except for band 13) are more significant in the classification model than bands with shorter wavelengths (i.e., bands 3, 4, and 5 except for band 2). Comparing Figure 8 with Figures 5 and 6 demonstrates that increasing the difference between the radiometric value in each band for distinct classes can increase the significance of that band in the classification model.

6. Preparation of preliminary geological maps for the region of Birjand-IV

After building three-class and fifteen-class models with the Khosuf-I data, these models were applied to the Birjand-IV data set to generate two preliminary geological maps for this region. The primary maps for three and fifteen rock units in the Birjand-IV region are depicted in Figures 9a and 9b, respectively, based on data modeling obtained from the random forest algorithm. Adjusting noise pixels to improve the map's quality and better delineate the boundary between rock units is necessary. Pixels whose class does not match the class of adjacent pixels are known as noise pixels. To this end, a method for image analysis employing various filters, including the median and average moving filters [72,73], has been proposed. The preliminary geological maps obtained after median and average moving filtering are depicted in Figures 9c and 9d. These figures

demonstrate that the separation of the rock units was feasible.

Figure 9c demonstrates that the majority of the rock unit in the region is igneous and is exposed as two large masses, one in the northern part (Shekarab Mountain area) and the other in the southern part of the region (Bagheran Mountain area). In addition, sedimentary rocks (Birjand alluvial plain) can be found in the central part of the region. Metamorphic rocks are located inside igneous masses and mostly in the southern part of the region. This division of rock units is consistent with the previous studies and the small-scale geological maps of the region [46,74].

The Shekarab Mountain area consists of andesite, dacite, and intermediate tuff rock units, as well as metamorphic rocks that are bordered to the south by conglomerate sediments, marl, and alluvium, as depicted on the map with more detailed rock units (Figure 9d). At the same time, the rock units of the Bagheran Mountain area include ophiolite, flysch sediment units, basaltic, and ultrabasic volcanic rocks. Large masses of ultramafic rocks are interlayered with sedimentary rocks (limestones and shales). Flysch rocks are also found at the boundary with Birjand plain sediments [49].

Field studies were conducted in the three Shekarab, Bagheran, and Eshkaftuk areas to assess the accuracy of the obtained preliminary geological maps and compare them to surface conditions. These studies aim to identify rock units and determine their boundaries and expansion in these areas, the specifics of which are listed below.

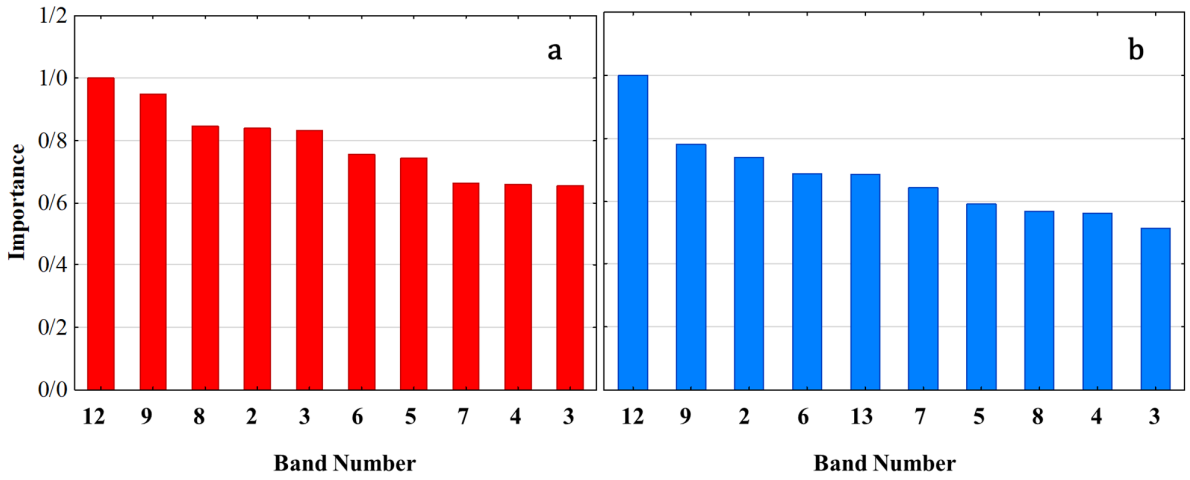

Figure 8. Importance of various bands in the classification model for (a) the three-class and (b) the fifteen-class training data sets.

6.1. Shekarab area

The Shekarab area is located in the northeast of the Birjand-IV geological map (Figure 2b). The mountains in this area, which extend from east to west, include various folded and thrust rock units from the Upper Cretaceous period (ophiolitic melange rock unit) as well as the Paleogene period (the andesite rock unit). These rock units are bordered by conglomerate sediments, Neogene marl, and Quaternary alluvium on the southern slope.

Field observations (points in Figures 2b and 11a) indicate that andesitic volcanic rocks with an east-west trend and an Eocene age, as well as marl, sandstone, and conglomerate units, outcrop in front of the Shekarab Mountains (Figure 10a). Veins of brown, yellow, or white listvenites, are found in ultramafic rocks that were formed by serpentinite carbonate (Figure 10b). On the peaks of the Shekarab Mountain area are phyllite and schist metamorphic rocks, as well as ultrabasic rocks with microscopic folds (Figures 10c and 10d). A portion of the Middle borders this complex to the region's Late Cretaceous ophiolitic melange rock unit (peridotite and spilite rock units). On the Google map of the Shekarabarea, rock units are separated based on field observations, and a simplified geological map (i.e., a lithological map) of this area has been created. This map and the geological maps resulting from the modeling are shown in Figure 11.

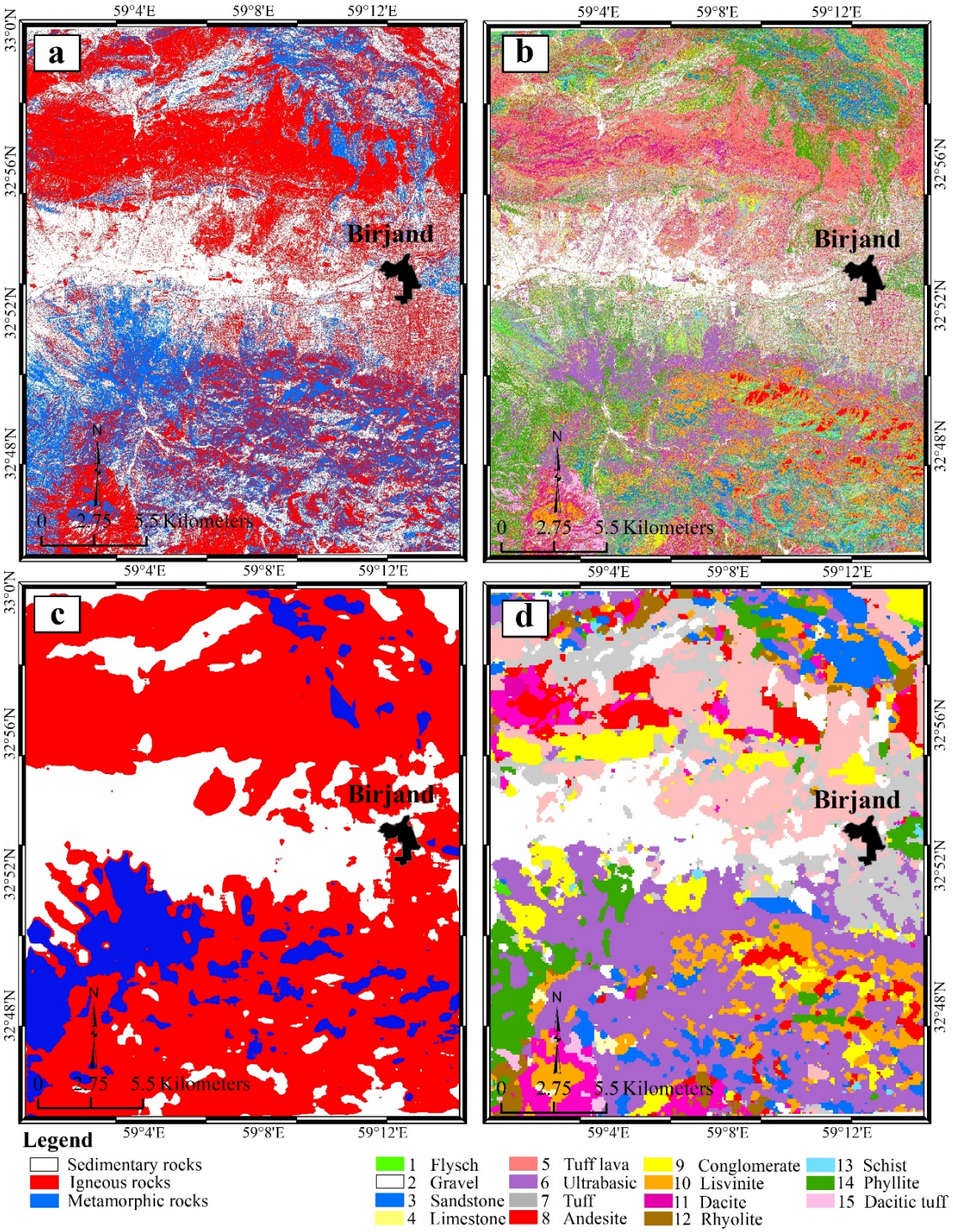

Figure 9. The 1:50,000 scale geological map of the Birjand-IV region derived from the modeling technique, (a) primary three-classes, (b) primary fifteen-classes, (c) preliminary three-classes, and (d) preliminary fifteen-classes maps.

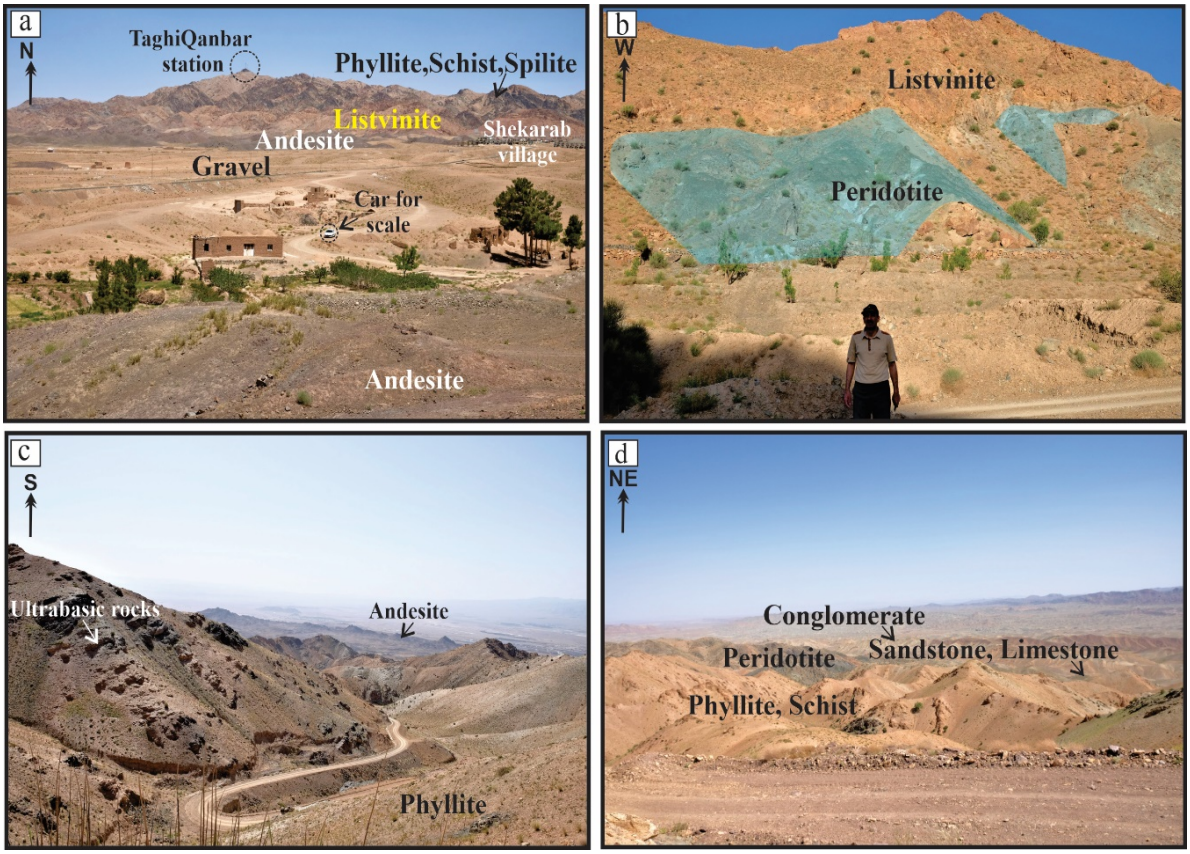

Figure 10. Rock units of the Shekarab area, (a) a view of the Shekarab Mountain area with different colored rock units, (b) peridotite rock masses in the listvenite unit, (c) an ultramafic rock unit with a dark color tone in the phyllite rock unit, and (d) parts of the ophiolitic sequence in the Shekarab Mountain area.

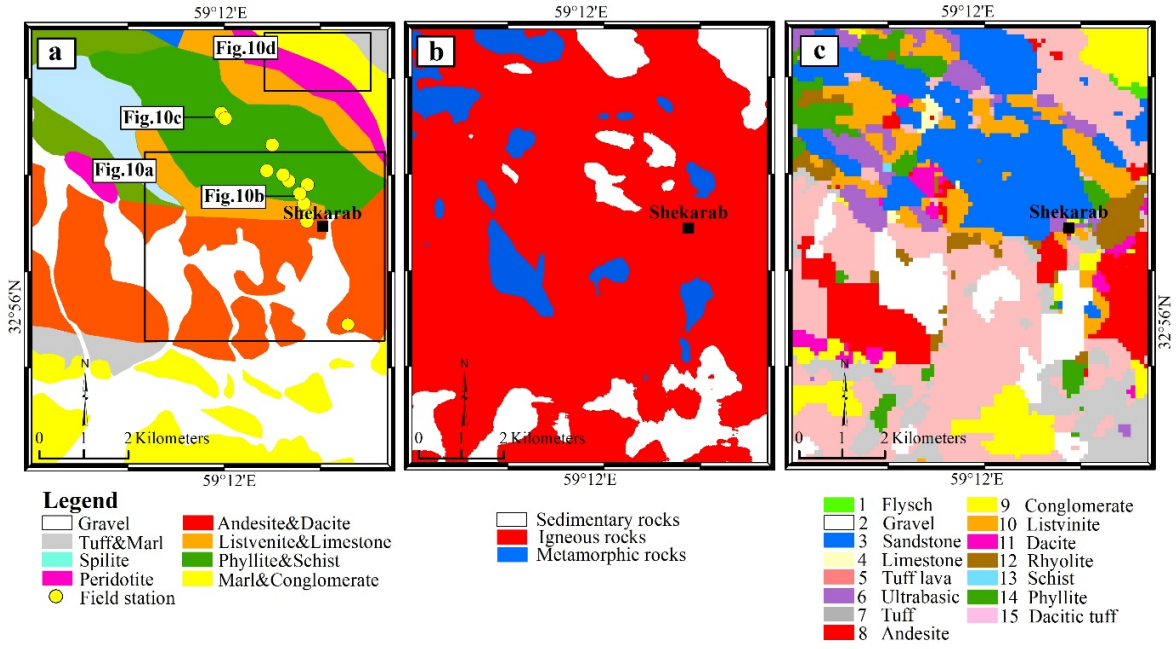

Figure 11. Lithological map prepared using field studies in the (a) Shekarab area, (b) a part of 1:50,000 scale Birjand-IV geological map in the Shekarab area obtained by filtered three-classes, and (c) fifteen-classes modeling method.

Table 3 compares the rock units in the maps of Figure 11. Most phyllite and schist rock units have been introduced as sandstone and, in some cases, phyllite fragments in the northeastern and western parts of the region (Figures 11a and 11c). The tuff and marl rock units conform to the model more closely. Andesites have a significant expansion in the central part of the region, adjacent to conglomerate, marl, and tuff rock units. In addition to identifying andesitic and conglomerate rock units in this region, tuff and lava tuff rock units have also been modeled (Figure 11a and 11c).

Regarding rock type and boundary separation, the modeled results for listvenite, andesite, and gravel rock units are generally acceptable. The Birjand plain is covered with gravel sediments that the Birjand anticline, composed of conglomerate and marl rock units, is exposed in its central part (Figure 11). This part comprises modeling gravel, lava tuff, and conglomerate rock units. There is a good fit between the model and the terrain in some areas, but there is a mismatch in the south and adjacent to the Bagheran Mountain area. This discrepancy in identification appears to be due to the erosion of the ophiolitic rock unit of the Bagheran Mountains and their placement on gravelly sediments in the plain (Figures 10 and 11). Clearly, the proximity of spectral graphs of rock units (Figure 6) can also lead to modeling errors.

Table 3. Comparison of rock units in the Shekarab area based on field studies and modeling.

Row	Location	Field study	Modeling
1	Northeast	Conglomerate, listvenite, phyllite, schist and peridotite	Sandstone, listvenite and dacite tuff
2	Northwest	Phyllite, schist, listvenite, split, tuff and marl,	Dacite, listvenite, ultrabasic and tuff
3	Central	Andesite	Andesite, dacite, tuff and lava tuff
4	Central	Conglomerate and marl	Conglomerate, tuff and lava tuff
5	South	Conglomerate, marl and gravel	Conglomerate, lava tuff, gravel and ultrabasic

6.2. Bagheran area

The Bagheran area is located in the southeast of the geological map of Birjand-IV (Figure 2b). The Bagheran Mountain area consists of 40-kilometer-long, northeast-to-southwest-curving mountains. Field studies have been conducted in the central part of these mountains. The main rock units of the mountains are Upper Cretaceous Ophiolitic series rock units. At the boundary between the Birjand plain and the mountains, Paleocene flysch sediments are flanked by ophiolites (Figure 12a). Green shale, sandstone, siltstone, and phyllite shale rocks are mixed with spilite or ultrabasic rock units in varying proportions in some areas of Flysches. (Figure 12b). Spilites are also important lithological components of the area, which in some places are associated with pink radiolarian sand limestones (Figure 12c).

Ultrabasic rocks are one of the constituent units of the ophiolite complex, which is widely dispersed throughout the Bagheran Mountain area and consists of the peridotite, harzburgite, and serpentinite minerals. There is a colored melange rock unit to the west of Sirvan village, and in the north, exposed metamorphic rocks consisting primarily of phyllite and schist rock units (epidote, chlorite, sericite, and talc schist minerals). These rocks are located near Cretaceous sandstone and shales. Figure 13a depicts a map of the main rock units of the Bagheran area as determined by these field studies. This map also depicts the locations of the field sampling points and the images in Figure 12. Figures 13b and 13c display modeling maps of the rock units in the Bagheran area.

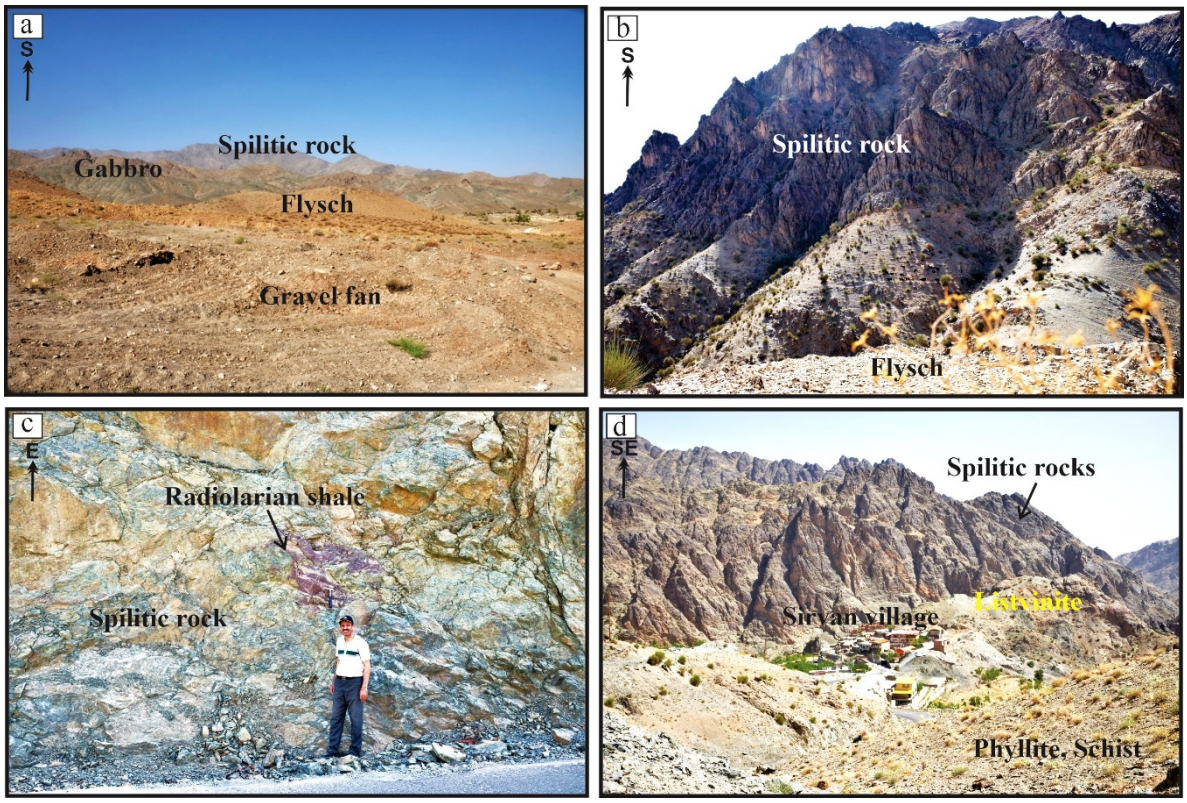

Figure 12. Rock units of the Bagheran area, (a) view of the Bagheran Mountain area with rock units of various colors, (b) boundary between flysch rock of spilite rock unit, (c) fragments of purple radiolarian shales in the spilite rock unit, and (d) listvenite, phyllite, and schist rock units in the vicinity of the spilite rock mass in Sirvan village.

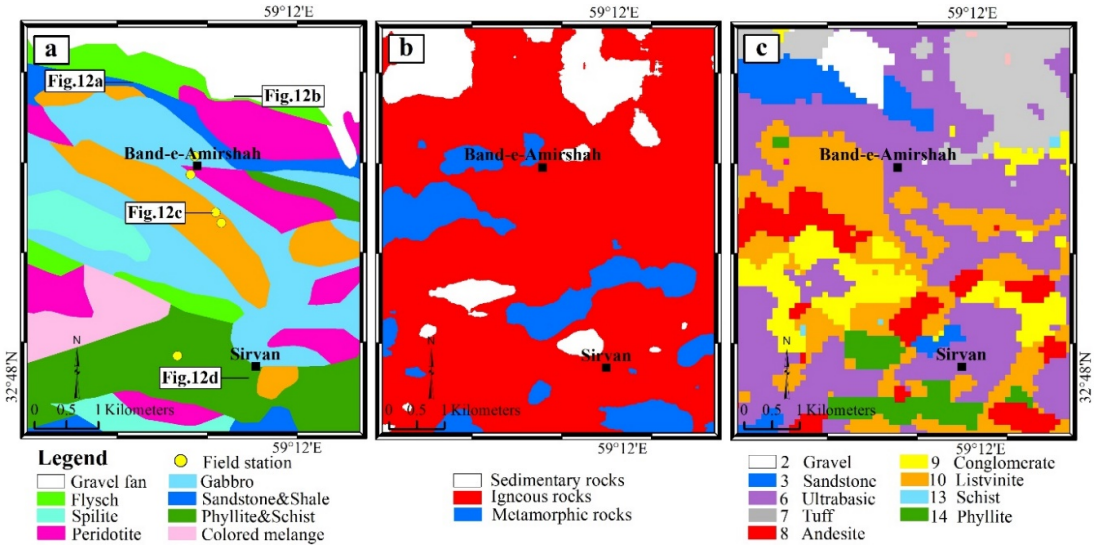

Figure 13. Lithological map prepared using field studies in the (a) Bagheran area, (b) a part of 1:50,000 scale Birjand-IV geological map in the Bagheran area obtained by filtered three-classes, and (c) fifteen-classes modeling method.

Table 4 compares the lithological map's rock units to those of modeling maps in the Bagheran area. In the northern part of the region, in the Birjand plain, sedimentary rocks eroded from the Bagheran mountains cover alluvial cones adjacent to the mountains (Figure 13). Modeling in this region reveals gravel, tuff, lava tuff, conglomerate, and ultrabasic rock units, which are actually

alluvial cone gravel sediments; the inconsistency in some areas can be attributed to ultrabasic rock surface cover. On the modeling maps, peridotite units are highlighted alongside ultrabasic units, flysch alongside conglomerate and sandstone, and spilite alongside andesite and conglomerate rock units. This section correctly separates listonite. The relative accuracy of the sequence of sandstone and shale and interlayers of lime with conglomerate and listonite separates the flysch rocks of the region (Figures 12 and 13).

Table 4. Comparison of rock units in the Bagheran area based on field studies and modeling.

Row	Location	Field study	Modeling
1	North	Gravel, peridotite, flysch, sandstone and shale	Gravel, ultrabasic, conglomerate and listvenite
2	Central	Split and peridotite	Andesite, conglomerate and ultrabasic
3	South	Phyllite, schist, sandstone, shale, peridotite and ophiolite melange	Phyllite, schist, andesite and ultrabasic

Furthermore, spilite rocks have been introduced as andesite, which contradicts reality because there was no spilite class on the modeling map of fifteen classes. In the south, phyllite and schist rock units have been partially identified (Figures 13a and 13c). However, ultrabasic rock units are separated from peridotites and ophiolites. These two rock units are also absent from the fifteen-class modeling map (Figure 13c). Therefore, it was impossible to distinguish between these rock units using modeling in this region.

6.3. Eshkaftuk area

The Eshkaftuk area is located to the northwest of the geological map of Birjand-IV (Figure 2b). A row of Neogene sedimentary rocks, including conglomerate, gypsum marl, tuff, and tuff marl, is exposed as one travels from the Birjand plain to the Bagheran Mountain area. The development and expansion of the Eocene tuff series, which includes red, green, and yellow tuffs, as well as tuffaceous marls and tuff breccias, are located in the continuation of the Paleogene volcanic rocks. This complex is covered by Eocene-Oligocene andesite and dacite rock units, along with volcanic breccia, tuff breccia, and andesitic agglomerates in certain areas (Figure 14a). Pliocene sandstone and phyllite shale rock units are visible in the vicinity of Upper Cretaceous spilite and ultrabasic (peridotite) rocks to the north of Eshkaftuk village (Figure 14b). Figure 15a displays the lithological map of the principal rock units in the Eshkaftuk area for comparison with the modeling maps (Figures 15b and 15c) of this region.

The comparison results between the lithology map derived from field studies and the modeling maps are displayed in Table 5. The sandstone rock unit in the north part of the region has been accurately modeled (Figure 15). In this section, tuff and gravel rock units are clearly distinguished and exhibited. Even though andesite and tuff rock units are visible in the region's center, the boundary between these rock units must be revised (Figures 15a and 15c). The range of gravelly sediments of the Birjand plain has been correctly identified and differentiated, and the conglomerate, agglomerate with conglomerate, and tuff rock units have been accurately modeled in the southern region (Figure 15).

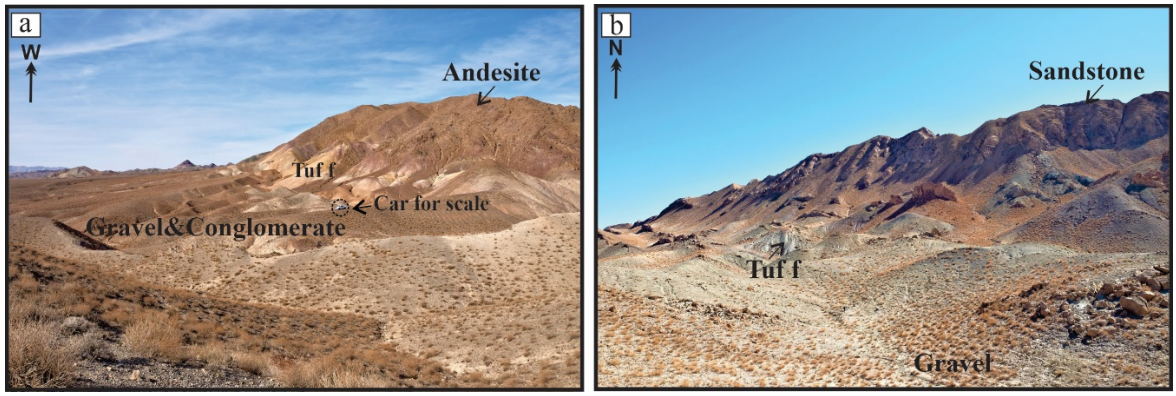

Figure 14. Rock units of the Eshkaftuk area, (a) the boundary between andesitic volcanic rocks and alluvial gravels with light-colored tuff rock units, and (b) the separation of light-colored tuff rock units from sandstone and shale with a dark color tone.

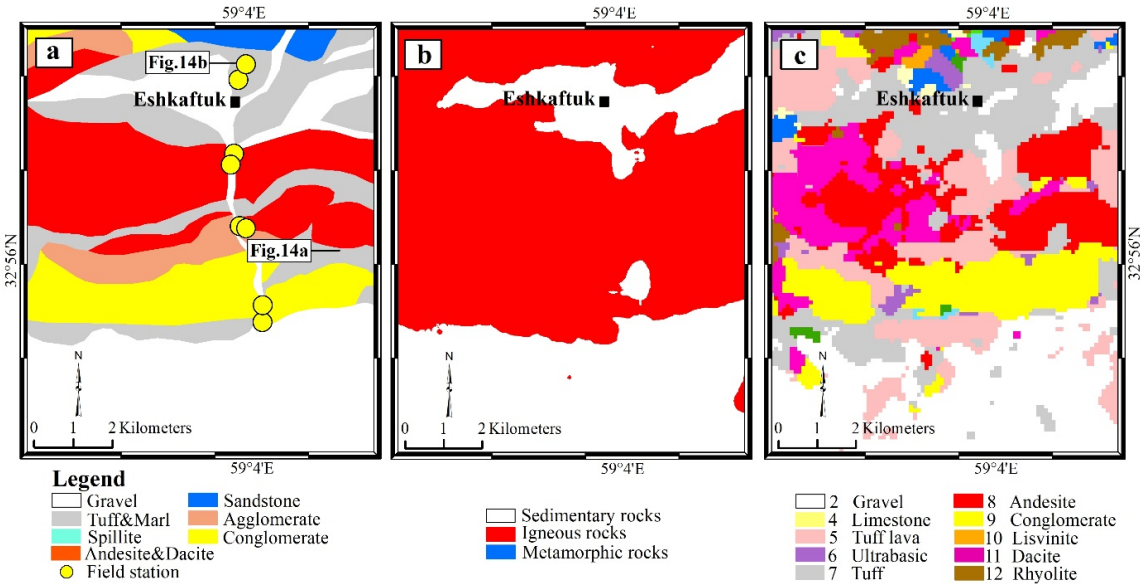

Figure 15. Lithological map prepared using field studies in the (a) Eshkaftuk area, (b) a part of 1:50,000 scale Birjand-IV geological map in the Eshkaftuk area obtained by filtered three-classes and, (c) fifteen-classes modeling method.

Table 5. Comparison of rock units in the Eshkaftuk area based on field studies and modeling.

Row	Location	Field study	Modeling
1	North	Sandstone, gravel, conglomerate and tuff	Sandstone, tuff, conglomerate and ultrabasic
2	Central	Andesite and tuff	Andesite-dacite and tuff
3	South	Conglomerate, agglomerate and tuff	Conglomerate and tuff

6.4. Adaptation of lithological data with rock unit thematic mapping

Shales and fine-grained sedimentary rocks, such as mudstones, have a generally dark tone in the study region. In contrast, coarse-grained sedimentary rocks, such as sandstones and conglomerates, have a light tone. The flysch facies are also observed in light green sandstone and violet-gray shale in the study region. Sandstones and flysch rock units have been identified in this paper as andesite and dacite rock units, possibly due to their similarity in color to andesite rock. In the studied region, dark-toned andesitic and dacite lavas are found in close proximity to light-colored marl and tuff rock units. Therefore, the classification of andesite and dacite rock units is somewhat acceptable. However, some parts of these rocks have been separated as lava tuff.

In the study region, carbonate rocks have a light tone. Listvenites can also be found in a light brown tone. These rock units have been highlighted in the modeling but cannot be distinguished without field studies.

Mafic and basaltic lava flows are darker than rhyolitic and siliceous lava flows in the study region. In the Bagheran Mountain area's ophiolite rock unit, dark green to dark brown dunite outcrops with green-yellow minerals can be observed. As a result of alteration, harzburgite's black-to-dark-green tone has been transformed into serpentine. Green to black is the weathering color of serpentinite, which is highly sheared and crushed. On the modeling maps, these rock units have been identified as ultrabasics. In some instances, however, they have been introduced as sandstone or andesite, which requires revision. In addition, the syphilitic massifs in the Bagheran Mountain area have been classified as andesites, presumably due to their dark green tone.

Alternated microgabbro rocks near fault planes are black to dark gray in tone. Due to faulting and tectonization, the progression of alteration, and the small size of this rock's outcrop in some areas, it is impossible to separate this rock from ophiolitic rock units. Consequently, it is known as a colored mixture. The ophiolitic complex contains gray-green diabase rocks. All these units are properly highlighted as ultrabasics in the modeling maps.

This study demonstrated that the preliminary geological map can be generated by a modeling algorithm. However, rock units are effectively highlighted; units with similar lithological characteristics (similar color tone) are presented similarly. This is evident in the rock units of

andesite, dacite, sandstone, and phyllite. This modeling identifies and separates basic and ultrabasic rock units more accurately. It appears that the selection of appropriate classes is required for satellite image processing. For instance, it is possible to combine the classes of andesite with dacite and tuff and lava tuff with marl to produce a simpler and more accurate modeling map. The preparation of the three-classes geological map will also help in identifying and separating the lithological units in the final map. In such a way, the rock units with misclassification can be corrected using igneous, sedimentary and metamorphic geological map, due to its higher accuracy. For instance, the classification of sandstone and shale from andesite, sandstone from ultrabasic and or phyllite and schist from andesite and ultrabasic can be done by this map. In any case, it is necessary to correct the mapping of rock units on the preliminary map, and field observations should be also used to create the final geological map of a region.

7. Conclusions

Due to the critical nature of geological maps with a scale of 1:500,000 for various applications in Iran, especially for mineral exploration, they must be created in less time and at a lower cost. With a spatial resolution of 10 meters, a large number of bands, and open-source SNAP software, Sentinel-2 satellite images were used instead of aerial photographs for this study. In addition, the random forest algorithm was employed for modeling due to its superior performance with large data sets and class-unbalanced data sets. Sentinel-2 band images and the Khosuf-1 geological map were utilized as input data and class labels, respectively, for modeling. The obtained model is then implemented on the Sentinel-2 band images of the Birjand-IV region to produce a geological map. Two geological maps were produced for the Birjand-IV region as a result of modeling using three-class and fifteen-class data sets.

The three-class geology map of the Birjand-IV region revealed that igneous rocks are exposed in the region's northern and southern parts. While sedimentary rocks are located in the region's center under study, metamorphic rocks are found within the igneous masses. The fifteen-class geology map of the Birjand-IV region revealed that the northern part of the study region is composed of andesite, dacite, and intermediate tuff rock units, as well as metamorphic rocks, while the southern part is bordered by conglomerate, marl, and alluvium sediments. In contrast, the rock units of the region's southern part include ophiolite, flysch sediments, basaltic, and ultrabasic volcanic rocks. Limestone and shale rock units are interlayered with ultramafic rock masses.

This study has also shown that (1) Sentinel 2 satellite images, due to having 8 bands in the VNIR spectral, can be useful in separating rock units for geological mapping, (2) increasing the difference between the radiometric value in each band for distinct classes and the non-overlapping spectral graphs can be effective in better classification and separation of each rock unit and (3) having only two bands in the SWIR region, as a limitation of Sentinel-2 Satellite Imagery, causes not able to detect molecular absorption of many silicate, carbonate and hydroxide minerals. However, the use of satellite images with hyper-Spectral and higher spatial resolution can overcome these limitations, especially in the preparation of local geological maps.

Comparing the field studies to the obtained geology maps revealed that the three-class and fifteen-class geology maps have high and acceptable accuracy, respectively. Two factors were influential in preparing a preliminary geological map of a region using satellite images: (1) the difference in the appearance color and band spectrum of rock units and (2) the optimal selection of the number of classes. However, using satellite images to create a preliminary map with high speed, low cost, and acceptable accuracy can aid in creating a final geological map with medium scales.

References

[1]. Alexandre, P. (2020). Novel Methods and Applications for Mineral Exploration, *MDPI Publisher*, 264 p.

[2]. Haldar, S.K. (2018). Mineral Exploration: Principles and Applications, 2nd Edition, *Elsevier*, 378 p.

[3]. Hill, P.R., Lebel, D., Hitzman, M., Smelrop, M. & Thorleifson, H. (2020). The Changing Role of Geological Surveys, Geological Society, *London, Special Publications 2020*, 499, 1–15.

[4]. Sinha, K., Sharma, P., Sharma, A., Singh, K. & Hassan, M. (2024). Analysis of Land Subsidence in Joshimath Township using GIS and Remote Sensing. *Journal of Mining and Environment*, 13(3), 817-843.

[5]. Gururani, D.M., Kumar, Y., Abed, S.A., Kumar, V., Vishwakarma, D.K., Al-Ansari, N., Singh, K., Kuriqi, A. & Mattar, M.A. (2023). Mapping Prospects for

Artificial Groundwater Recharge Utilizing Remote Sensing and GIS Methods. *Water*, 15, 3904.

[6]. Dhakal, D., Abdi, S., Singh, K. & Sharma, A. (2023). Applications of GIS and Remote Sensing in Highway Project-A Review. *Journal of Mining and Environment*, 14(2), 375-388.

[7]. Bhardwaj, V. & Singh, K. (2023). Landslide susceptibility assessment using remote sensing and GIS-a review. *Journal of Mining and Environment*, 14(1), 133-154.

[8]. Beiranvand Pour, A., Zoheir, B., Pradhan, B. & Hashim, M. (2021). Multispectral and Hyperspectral Remote Sensing Data for Mineral Exploration and Environmental Monitoring of Mined Areas. *Remote Sensing*, 13, 519.

[9]. Himanshu, G. (2015). Hyperspectral Remote Sensing Application in Mineral Exploration, *LAP Lambert Academic Publishing*, 160 p.

[10]. Jain, R. & Bhu, H. (2024). Potentials of spaceborne imaging spectrometer PRISMA, ASTER, and ALOS-1 PALSAR datasets for mineral mapping and geological controls of mineral deposit in Jahazpur, Rajasthan, India. *Remote Sensing*, 34, 101193.

[11]. Rajendran, S. & Nasir, S. (2019). ASTER capability in mapping of mineral resources of arid region: A review on mapping of mineral resources of the Sultanate of Oman. *Ore Geology Review*, 108, 33-53.

[12]. Wang, G. & Weng, Q. (2020). Remote Sensing of Natural Resources, *CRC Press*, 580 p.

[13]. Ahmad, A. & Nair, A.M. (2024). Lithological mapping of charnockites using spectral mixture analysis. *Remote Sensing*, 34, 101165.

[14]. Rajan Girija, R. & Mayappan, S. (2019). Mapping of mineral resources and lithological units: a review of remote sensing techniques. *International Journal of Image Data Fusion*, 10, 79-106.

[15]. Van der Meer, F.D. (2006). The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 8, 3–17.

[16]. Van der Meer, F.D., Van der Werff, H.M.A., Van Ruitenbeek, F.J.A., Hecker, C.A., Bakker, W.H., Noomen, M.F., Van der Meijde, M., Carranza, E.J.M., Boudewijn de Smeth, J. & Woldai, T. (2012). Multi-and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14, 112–128.

[17]. Chen, X., Warner, T.A. & Campagna, D.J. (2007). Integrating visible, near-infrared and short-wave infrared hyperspectral and multispectral thermal imagery for geological mapping at Cuprite, Nevada. *Remote Sensing of Environment*, 110, 344-356.

[18]. Cracknell, M.J. & Reading, A.M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Computers & Geosciences*, 63, 22–33.

[19]. Girouard, G.G., Bannari, A., El Harti, A. & Desrochers, A. (2004). Validated Spectral Angle Mapper Algorithm for Geological Mapping: Comparative Study between Quickbird and Landsat-TM, *Geological Mapping Project*.

[20]. Leverington, D.W. (2010). Discrimination of sedimentary lithologies using Hyperion and Landsat Thematic Mapper data: a case study at Melville Island, Canadian High Arctic. *International Journal of Remote Sensing*, 31, 233–260.

[21]. Leverington, D.W. & Moon, W.M. (2012). Landsat-TM-based discrimination of lithological units associated with the Purtuniq Ophiolite, Quebec, Canada. *Remote Sensing*, 4, 1208–1231.

[22]. Oommen, T., Misra, D., Twarakavi, N.K.C., Prakash, A., Sahoo, B. & Bandopadhyay, S. (2008). An objective analysis of support vector machine based classification for remote sensing. *Mathematical Geosciences*, 40, 409–424.

[23]. Pal, M., Rasmussen, T. & Porwal, A. (2020). Optimized Lithological Mapping from Multispectral and Hyperspectral Remote Sensing Images Using Fused Multi-Classifiers. *Remote Sensing*, 12, 177.

[24]. Waske, B., Benediktsson, J.A., Árnason, K. & Sveinsson, J.R. (2009). Mapping of hyperspectral AVIRIS data using machine-learning algorithms. *Canadian Journal of Remote Sensing*, 35, 106–116.

[25]. Ye, B., Tian, S., Cheng, Q. & Ge, Y. (2020). Application of Lithological Mapping Based on Advanced Hyperspectral Imager (AHSI) Imagery Onboard Gaofen-5 (GF-5) Satellite. *Remote Sensing*, 12, 3990.

[26]. Yu, L., Porwal, A., Holden, E.J. & Dentith, M.C. (2012). Towards automatic lithological classification from remote sensing data using support vector machines. *Computers & Geosciences*, 45, 229–239.

[27]. Fatima, K., Khattak, U.K. & Kausar, A.B. (2013). Selection of appropriate classification technique for lithological mapping of Gali Jagir area, Pakistan. *International Journal of Earth Sciences and Engineering*, 7, 964–971.

[28]. Adiri, Z., El Harti, A., Jellouli, A., Maacha, L. & Mostafa Bachaoui, E. (2016). Lithological mapping using Landsat 8 OLI and Terra ASTER multispectral data in the Bas Drâa inlier, Moroccan Anti Atlas. *Journal of Applied Remote Sensing*, 10, 016005.

[29]. AlMakki Mohamed, M.T., Al-Naimi, L.S., Mgbeojedo, T.I. & Agoha, C.C. (2021). Geological mapping and mineral prospectivity using remote sensing and GIS in parts of Hamissana, Northeast Sudan.

Journal of Petroleum Exploration and Production Technology, 11, 1123–1138.

[30]. Amaral, C.H.d., Almeida, T.I.R.d., Souza Filho, C.R.d., Roberts, D.A., Fraser, S.J., Alves, M.N. & Botelho, M. (2018). Characterization of indicator tree species in neotropical environments and implications for geological mapping. *Remote Sensing of Environment*, 216, 385-400.

[31]. Gomez, C., Delacourt, C., Allemand, P., Ledru, P. & Wackerle, R. (2005). Using ASTER remote sensing data set for geological mapping, in Namibia. *Physics and Chemistry of the Earth*, 30, 97–108.

[32]. Khalifa, A., Bashir, B., Çakir, Z., Kaya, S., Alsalman, A. & Henaish, A. (2021). Paradigm of Geological Mapping of the Adıyaman Fault Zone of Eastern Turkey Using Landsat 8 Remotely Sensed Data Coupled with PCA, ICA, and MNFA Techniques. *International Journal of Geoinformatics*, 10, 368.

[33]. Mahmoudishadi, S., Malian, A. & Hosseinali, F. (2017). Comparing independent component analysis with principal component analysis in detecting alterations of porphyry copper deposit (Case study: Ardestan area, central Iran). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLII-4/W4, Tehran, Iran.

[34]. Mohan, M. & Meyyappan, M. (2022). Mapping of mafic-ultramafic rocks in SMUC-SGT, India using ASTER & Sentinel-2A satellite images. *Remote Sensing*, 28, 100826.

[35]. Peyghambari, S. & Zhang, Y. (2021). Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review. *Journal of Applied Remote Sensing*, 15, 031501.

[36]. Al-Nahmia, F., Saddiqia, O., Hilalia, A., Rhinanea, H., Baiddera, L., El arabia, H. & Khanbari, K. (2017). Application of remote sensing in geological mapping, Case study Al Maghrabah area- Hajjah region, Yemen. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume IV-4/W4, Karabuk, Turkey.

[37] Andreani, L., Herrmann, E., Lorenz, S., Zimmermann, R., Kirsch, M., Brazzo, N. & Gloaguen, R. (2021). Sentinel-2 as a tool for mapping iron-bearing alteration minerals: a case study from the Iberian Pyrite Belt (Southern Spain). *EGU General Assembly*, 10719.

[38]. Ge, W., Cheng, Q., Jing, L., Wang, F., Zhao, M. & Ding, H. (2020). Assessment of the Capability of Sentinel-2 Imagery for Iron-Bearing Minerals Mapping: A Case Study in the Cuprite Area, Nevada. *Remote Sensing*, 12, 3028.

[39]. Van der Meer, F.D., Van der Werff, H.M.A. & Van Ruitenbeek, F.J.A. (2014). Potential of ESA's Sentinel-2 for geological applications. *Remote Sensing of Environment*, 148, 124-133.

[40]. Van der Werff, H. & Van der Meer, F.D. (2015). Sentinel-2 for Mapping Iron Absorption Feature Parameters. *Remote Sensing*, 7, 12635-12653.

[41]. Van Der Werff, H. & Van Der Meer, F.D. (2016). Sentinel-2A MSI and Landsat 8 OLI provide data continuity for geological remote sensing. *Remote Sensing*, 8, 883.

[42]. ODSS, (2011). Instructions for Compiling and Digitizing Large Scale Geological- Exploration Maps, Office of Deputy for Strategic Supervision, Bureau of Technical Execution System, *Ministry of Industries and Mines of Iran*, Report No. 532, 77 p. (In Persian).

[43]. Shuai, A., Ardebili, O., Badrikallo, N., Jadidi, R., Strabi-Ashtiani, M. & Awadnamur, M. (2019). Atlas of maps of the Geological Survey and Mineral Explorations of Country, *Geological Survey & Mineral Explorations of Iran*, Tehran, 218 p., (In Persian).

[44]. Geranian, H. & Khajeh Miry, Z. (2018). Introducing mineral potentials and investment solutions in the mining industry of South Khorasan province, *The 1st National Conference on Investment Opportunities and Limitations in South Khorasan Industry*, 58-64, (In Persian).

[45]. Camp, V.E. & Griffis, R.J. (1982). Character, genesis and tectonic setting of igneous rocks in the Sistan suture zone, eastern Iran. *Lithos*, 15, 221-239.

[46]. Tirrul, R., Bell, I.R., Griffis, R.J. & Camp, V.E. (1983). The Sistan suture zone of eastern Iran. *Geological Society of America Bulletin*, 94, 134-150.

[47]. Babazadeh, S.A. & De Wever, P. (2004). Early Cretaceous radiolarian assemblages from radiolarites in the Sistan Suture (eastern Iran). *Geodiversitas*, 26, 185-206.

[48]. Delaloye, M & Desmons, J. (1980). Ophiolites and mélange terranes in Iran: a geochronological study and its paleotectonic implications. *Tectonophysics*, 68, 83–111.

[49]. Zarrinkoub, M.H., Pang, K.N., Chung, S.L., Khatib, M.M., Mohammadi, S.S., Chiu, H.Y. & Lee, H.Y. (2012). Zircon U\Pb age and geochemical constraints on the origin of the Birjand ophiolite, Sistan suture zone, eastern Iran. *Lithos*, 154, 392–405.

[50]. Pang, K.N., Chung, S.L., Zarrinkoub, M.H., Khatib, M.M., Mohammadi, S.S., Chiu, H.Y., Chu, C.H., Lee, H.Y. & Lo, C.H. (2013). Eocene-Oligocene post-collisional magmatism in the Lut-Sistan region, eastern Iran: magma genesis and tectonic implications. *Lithos*, 180-181, 234-251.

[51]. Arjmandzadeh, R., Karimpour, M.H., Mazaheri, S.A., Santos, J.F., Medina, J.M. & Homam, S.M. (2011). Sr\Nd isotope geochemistry and petrogenesis of

the Chah-Shaljami granitoids (Lut Block, Eastern Iran). *Journal of Asian Earth Sciences*, 41, 283–296.

[52]. Angiboust, S., Agard, P., De Hoog, J.C.M., Omrani, J. & Plunder, A. (2013). Insights on deep, accretionary subduction processes from the Sistan ophiolitic "melange" (Eastern Iran). *Lithos*, 156–159, 139–158.

[53]. Walker, R., Gans, P., Allen, M.B., Jackson, J., Khatib, M., Marsh, N. & Zarrinkoub, M. (2009). Late Cenozoic volcanism and rates of active faulting in eastern Iran. *Geophysical Journal International*, 177, 783–805.

[54]. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.

[55]. Charbuty, B. & Abdulazeez, A.M. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2, 20-28.

[56]. Han, J., Kamber, M. & Pei, J. (2011). Data mining: concepts and techniques, 3rd Edition, *Morgan Kaufmann*, 744 p.

[57]. Patel, H. & Prajapati, P. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. *International Journal of Computer Systems Science and Engineering*, 6, 74-78.

[58]. Tangirala, S. (2020). Evaluating the Impact of GINI Index and Information Gain on Classification using Decision Tree Classifier Algorithm. *International Journal of Advanced Computer Science and Applications*, 11, 612-619.

[59]. Schonlau, M. & Zou, R.Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal: Promoting communications on statistics and Stata*, 20, 3-29.

[60]. Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V.R., Murayama, Y. & Ranagalage, M. (2020). Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sensing*, 12, 2291.

[61]. Aggarwal, C.C. (2014). Data Classification: Algorithms and Applications, *Chapman and Hall*, 707 p.

[62]. Brereton, R.G. & Lloyd, G.R. (2016). Re-evaluating the role of the Mahalanobis distance measure. *Journal of Chemometrics*, 30, 134–143.

[63]. Qiong, G., Xian-Ming, W., Zhao, W., Bing, N. & Chun-Sheng, X. (2016). An Improved SMOTE Algorithm Based on Genetic Algorithm for Imbalanced Data Classification. *Journal of Digital Information Management*, 14, 92-103.

[64]. Brownlee, J. (2021). Imbalanced Classification with Python, *Machine Learning Mastery*, 463 p.

[65]. Sisodia, D.S. & Verma, U. (2019). Distinct multiple learner-based ensemble SMOTE bagging (ML-ESB) method for classification of binary clad imbalance problems. *International Journal of Technology*, 10, 721-730.

[66]. Chawla, N.V., Bowyer, K.W., Hall, L.O. & Kegelmeyer, W.P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357.

[67]. Abedin, M.Z., Guotai, C., Hajek, P., & Zhang, T. (2023). Combining weighted SMOTE with ensemble learning for the class-imbalanced prediction of small business credit risk. *Complex & Intelligent Systems*, 9, 3559–3579.

[68]. Wongvorachan, T., He, S., & Bulut, O. (2023). A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining. *Information*, 14, 54.

[69]. Grandini, M., Bagli, E. & Visani, G. (2020). Metrics for multi-class classification: An overview, *Computer Science, Mathematics*, arXiv:2008.05756v1.

[70]. Ardjmand, E., Millie, D.F., Ghalehkhondabi, I., Young, W.A. & Weckman, G.R. (2021). A State-Based Sensitivity Analysis for Distinguishing the Global Importance of Predictor Variables in Artificial Neural Networks. *Advances in Artificial Neural Systems*, Article ID 2303181.

[71]. Saarela, M. & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*, 3, 272.

[72]. Nasar Mursal, A.S. & Ibrahim, H. (2020). Median Filtering Using First-Order and Second-Order Neighborhood Pixels to Reduce Fixed Value Impulse Noise from Grayscale Digital Images. *Electronics*, 9, 2034.

[73]. Al-amri, S., Kalyankar, N.V. & Khamitkar, S.D. (2010). A Comparative Study of Removal Noise from Remote Sensing Image. *IJCSI International Journal of Computer Science Issues*, 7, 32-36.

[74]. Eftekharnezhad, J. (1980). Separating Different Parts of Iran in Vies of Constructional Position Related Too Sedimentary Basins, *Petrol Association Publication*, No: 82.

انجمن مهندسی معدن ایران
نشریه مهندسی معدن و محیط زیست
نشانی نشریه: www.jme.shahroodut.ac.ir
دانشگاه صنعتی شاهرود

تهیه پیش نقشه زمین¬شناسی ناحیه¬ای به کمک تصاویر ماهواره¬ای سنتینل ۲ و الگوریتم جنگل تصادفی در شرق ایران

حمید گرانیان* و محمد امیر علیمی

گروه مهندسی معدن، دانشگاه صنعتی بیرجند، بیرجند، ایران

اطلاعات مقاله

تاریخ ارسال: ۲۰۲۴/۰۹/۱۶
تاریخ داوری: ۲۰۲۴/۱۰/۲۳
تاریخ پذیرش: ۲۰۲۴/۱۱/۰۸

DOI: 10.22044/jme.2024.15083.2882

کلمات کلیدی

نقشه زمین¬شناسی
الگوریتم جنگل تصادفی
الگوریتم SMOTE
پردازش تصویر
پردازش تصاویر سنتینل ۲

چکیده

در این تحقیق از تصاویر ماهواره¬ای سنتینل ۲ به همراه الگوریتم جنگل تصادفی برای تهیه یک نقشه زمین شناسی ناحیه¬ای استفاده شده است. برای این منظور تصاویر ۱۰ باند ماهواری سنتینل ۲ محدوده خوسف-۱ به عنوان داده¬های مستقل و نقشه¬ی زمین¬شناسی خوسف-۱ تفکیک شده به سه و پانزده واحد سنگی به عنوان کلاس داده¬ها بکار رفته تا دو مدل کلاسه¬بندی توسط الگوریتم جنگل تصادفی با ۱۵ درخت بدست آید. مدل¬های بدست آمده از دقت کلاسه¬بندی ۹۷/۹۰ و ۸۴/۸۵ درصد برای داده¬های آموزشی و آزمایشی سه کلاسه و ۹۴/۷۶ و ۶۳/۹۲ درصد برای داده¬های آموزشی و آزمایشی پانزده کلاسه برخودار هستند. سپس دو مدل بدست آمده بر روی داده¬های تصاویر ماهواره¬ای سنتیل ۲ محدوده بیرجند-۴ (محدوده کناری خوسف-۱) پیاده شده تا دو پیش نقشه زمین¬شناسی سه و پانزده واحد سنگی برای این محدوده تهیه گردد. نقشه سه کلاسه بیرجند-۴ نشان می¬دهد که سنگ¬های آذرین در بخش¬های شمالی و جنوبی، سنگ¬های رسوبی در بخش میانی و سنگ¬های دگرگونی در داخل توده¬های آذرین منطقه واقع شده‌اند. همچنین نقشه پانزده کلاسه بیرجند-۴ نشان می¬دهد که بخش شمال منطقه را واحدهای سنگی آندزیتی، داسیتی و توف¬های حدواسط به همراه سنگ¬های دگرگونی پوشش می¬دهند. در حالیکه واحدهای سنگی بخش جنوب منطقه بیشتر شامل توالی از افیولیت، رسوبات فلیشی و سنگ‌های آتشفشانی و ولکانیکی بازی و فوق‌بازی به همراه میان لایه¬های آهکی و شیلی هستند. مطالعات صحرایی در سه محدوده نشان می¬دهد که پیش نقشه-های بدست آمده از دقت مناسبی برخودار هستند.

نویسنده مسئول مکاتبات: h.geranian@birjandut.ac.ir (ح. گرانیان)