Geological Mapping via Convolutional Neural Network Based on Remote Sensing and Geochemical Survey Data in Vegetation Coverage Areas

Ting Pan, Renguang Zuo, and Ziye Wang

***Abstract*—Geological mapping in vegetation coverage areas is a challenging task. In this article, convolutional neural networks (CNNs) were employed for geological mapping in a vegetation coverage area based on remote sensing images and geochemical survey data. The Gram–Schmidt fusion technology was first applied to fuse Sentinel-2A and ASTER remote sensing images to enhance the spatial resolution and enrich spectral information of remote sensing data. The fused remote sensing images were then organically integrated with geochemical survey data according to the correlations between the geochemical element contents and spectral reflectance of the objects. A case study of mapping six lithologic units in Jilinbaolige, Inner Mongolia, China was implemented using a CNN model based on the fused data. The classification map obtained an overall accuracy of 83.0%, which exhibited a better performance in contrast to random forest model. The results showed that CNNs can take full advantage of the spatial features of fused data and solve the problems of the "salt and pepper phenomenon" against the shallow machine learning algorithms, and the fusion of remote sensing and geochemical data can provide rich diagnostic information for geological mapping.**

***Index Terms*—Convolutional neural network, data fusion, geological mapping, random forests.**

I. Introduction

GEOLOGICAL mapping plays a critical role in geological survey. It is one of the most basic and important means for mineral prospecting and exploration. There are two difficulties associated with identifying lithological units: data and methods. Rocks and minerals have specific geochemical compositions, spectral absorption, and reflection properties, which can be identified using geochemical and remote sensing data. Remote sensing images contain information about the electromagnetic waves of the target, which have the advantages of multiscale, multitemporal, and high-spatial-resolution, particularly in areas that are hard to access [1].

However, multispectral remote sensing technology cannot detect subsurface geological bodies. In addition, it is difficult to solve the problem of misclassification caused by mixed pixels and spectral variations. Although the scale of regional geochemical survey data is small, they record the geochemical characteristics of geological units, which offers a way to solve the abovementioned problems. Because the mineral characteristic spectrum is a response to mineral chemical composition, there is a close relationship between them. Both geochemical survey data and multispectral remote sensing data reflect the different attributes of the same ground object, which are the physical and chemical characteristics, respectively. The distribution of geochemical elements involves physical processes such as electronic transitions and atomic vibrations, which leads to microscopic changes such as spectral reflection and radiation. These changes can be recorded in remote sensing images [2]. Accordingly, for the same ground object, there is a spatial and genetic correlation between the spectra in the remote sensing imagery and geochemical element concentrations. Multisource data fusion provides an effective way to utilize the advantages of different kinds of data sources by combining the information regarding different features of the same target, thus providing abundant information for geological mapping from multiple perspectives [1], [3], [4].

Numerous approaches have been developed to identify lithologic units for geological mapping, ranging from traditional statistical methods (such as principal component analysis) [5] to machine learning (ML) algorithms, such as random forests (RF) [6], [7], support vector machines (SVM) [8], [9], and lasso regression [10]. For example, Harris and Grunsky [7] used the RF classification algorithm for geological mapping based on geochemical survey data, aeromagnetic data, and γ-ray spectrum. Yu et al. [11] used SVM to classify ASTER data for automatic lithology classification, and revealed the importance of spatial detail features such as texture structure and topography. Li et al. [12] used the maximum likelihood method to classify TM remote sensing data in combination with spatial texture features, and improved the classification accuracy from 54.3% to 83.2%. The combination of MLs and multisource data is an easy and inexpensive method for geological mapping [1]. However, the abovementioned methods make decisions for each pixel in the classification map, which may ignore the spatial characteristics of the neighboring data, and thus, may cause the salt-and-pepper phenomenon in the classification results. Deep learning (DL) algorithms can overcome the limitations of these methods. DL is a neural network model with multiple hidden layers that can improve the accuracy of classification

Manuscript received 25 July 2022; revised 4 January 2023 and 11 February 2023; accepted 16 March 2023. Date of publication 22 March 2023; date of current version 11 April 2023. This work was supported by the National Natural Science Foundation of China under Grant 42172326. *(Corresponding authors: Renguang Zuo; Ziye Wang.)*

The authors are with the State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China (e-mail: 2740430275@qq.com; zrguang@cug.edu.cn; ziyewang@cug.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3260584

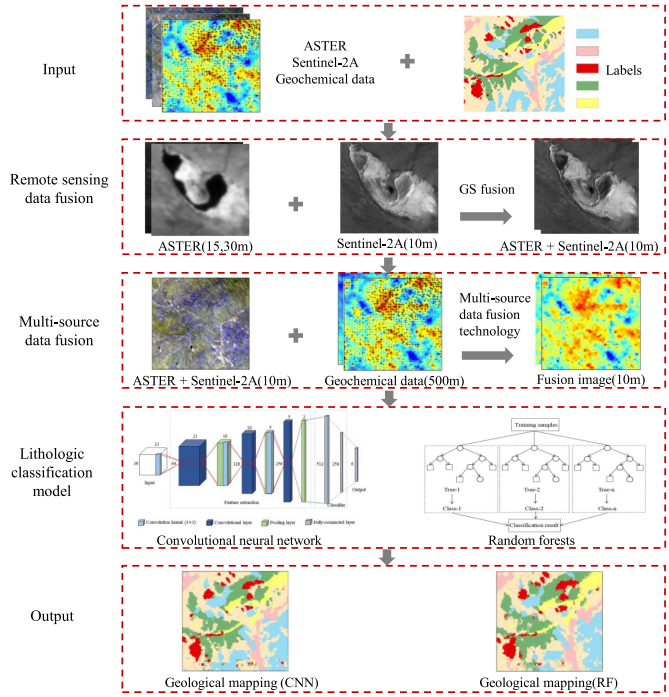

Fig. 1. Basic workflow of this article.

and prediction by learning and extracting deep abstract features of the data [13]. Convolutional neural networks (CNNs), as a famous DL algorithm, can simulate the structure and function of human brain neurons [14]. Moreover, CNNs are typical imagewise classifiers that have gained increasing attention because their performance is much better than that of traditional ML algorithms [15].

This article provides an alternative approach for geological mapping in vegetation coverage areas using multisource data fusion technology and a lithology identification model based on a CNN, aiming to solve two issues: 1) a single type of data resource cannot accurately diagnose the characteristics of all lithological units; 2) commonly used ML methods only perform pixel-wise classification that ignores spatial features. In this regard, Gram–Schmidt (GS) multisource data fusion technology was first applied to fuse Sentinel-2A and ASTER remote sensing images to improve the spatial resolution of the remote sensing data and enrich the spectral information. The fused remote sensing images were then organically integrated with the geochemical survey data according to the correlations between geochemical element concentrations and spectral reflectance, which not only enriches the spatial details of geochemical survey data but also complements the advantages of different data sources, thus providing rich diagnostic information for geological mapping. A CNN model (see Fig. 1) was built to distinguish the geological units based on the fused data. The results were compared with an RF to demonstrate the advantage of a CNN for geological mapping using a case study in Jilinbaolige, Inner Mongolia, China. In summary, this article is valuable and practical for geological mapping in vegetation coverage areas.

II. Methods

A. GS Fusion Technology

The GS transform [16], a classical pixelwise fusion method, eliminates redundant information using orthogonal transformation of multidimensional images or matrices [17], and improves the problem of excessive concentration of information. The fused high-resolution image obtained using the GS transform retains the spectral information to the greatest extent and has a color retention similar to that of the original multispectral image. The registered remote sensing images are synthesized into a new image by the GS fusion algorithm, and as much information as possible is retained to improve the image quality and provide clearer details of the ground objects. The basic steps of GS fusion are as follows [18], [19], [20], [21].

1) A panchromatic image is simulated using the original low-spatial-resolution image.
2) The panchromatic image is adopted as the first component of the GS transformation of the low-spatial-resolution image, and the specific transformation equation is

$$\mathrm{GS}_T(i,j) = (B_T(i,j) - \mu_T) - \sum_{l=1}^{T-1} (\emptyset(B_T, \mathrm{GS}_l) \cdot \mathrm{GS}_l(i,j)) \quad (1)$$

$$\mu_T = \frac{\sum_{j=1}^{N}\sum_{i=1}^{M} B_T(i,j)}{M \times N} \quad (2)$$

$$\emptyset(B_T, \mathrm{GS}_l) = \left[\frac{\sigma(B_T, \mathrm{GS}_l)}{\sigma(GS_l, \mathrm{GS}_l)^2}\right] \quad (3)$$

where GS_T represents the Tth orthogonal component after the GS transformation, B_T is the T band of the original image, μ_T is the mean value of the gray value of the T-band pixel of the original image, $\emptyset(B_T, \mathrm{GS}_l)$ is the covariance between the T band and GS_l of the original image, i and j, M, and N represent the number of rows and columns of the original image and the entire image, respectively.
3) The first component, namely the GS_1 component, is substituted by a high-spatial-resolution image after the GS transformation.
4) GS inverse transform is performed on the replaced dataset using the following equation to obtain the fused image:

$$B_T(i,j) = (\mathrm{GS}_T(i,j) + \mu_T) + \sum_{l=1}^{T-1} (\emptyset(B_T, \mathrm{GS}_l) \cdot \mathrm{GS}_l(i,j)). \quad (4)$$

In the equation, all the symbols have the same meanings as abovementioned.

B. Multisource Data Fusion

Geochemical survey data is fused with remote sensing images using a multisource data fusion method [1], [22], [23]. The geochemical element layers and multispectral bands have the same spatial coordinate system and cover the same area. Ding

et al. [23] proposed a detailed workflow for fusing geochemical and remote sensing data as follows.

1) The same-size single-element geochemical layer (*geo_l*) and multispectral remote sensing images (*MS*) were clipped, and spatial correction and registration were implemented.
2) The Laplacian pyramid algorithm was adopted to decompose the *MS* into two parts: high-frequency (MS_H_i) and low-frequency components (MS_L_i), which represent the spatial details and background, respectively [24], [25].
3) The geochemical layer of each element (geo_l_j) and low-frequency component of multispectral bands (MS_L_i) were resampled to the same resolution as MS_H_i using the cubic convolution method [26].
4) The relationship between the geochemical concentration and spectral reflectance was established. The following relational function can be built, and the correlation coefficient f_{ij} can be obtained:

$$geo\_lj = fij \cdot MS\_Li \tag{5}$$

where *i* is the band index and *j* is the element index.

5) The geochemical and high-frequency components MS_H_i are reconstructed according to the correlation coefficient f_{ij}, and the high-frequency information layer *geo_h* can be written as

$$geo\_h = fij \cdot MS\_Hi. \tag{6}$$

6) The high-frequency information fusion layer that represents the spatial details was injected into the resampled geochemical layer to obtain the fused high-resolution geochemical layer geo_*f*

$$geo\_f = geo\_lj + geo\_h. \tag{7}$$

C. Convolutional Neural Network

A CNN is a feedforward neural network with convolution and pooling operations [14]. CNNs adopt the network construction mode of local perception and weight sharing to reduce the complexity of the neural network and improve its generalization ability [27], [28]. A typical CNN primarily involves input, convolution, pooling, fully connected, and output layers. These layers have their own mapping relationships, extracting the characteristics of the input data and transmitting them to the complete CNN structure.

The convolutional layer, as the core of a CNN, controls feature extraction through a convolution operation between a convolution kernel of a certain size and the input matrix. The convolutional layer is characterized by local perception and weight sharing. Local perception (receptive field) implies that each neuron in the hidden layer does not have to connect all neurons [29], [30], [31]. The pooling layer, also known as the downsampling layer, is designed to decrease the size of the output feature map, and reduce the space occupation of the subsequent feature map [32]. In this article, the maximum pool layer is adopted. After several convolution and pooling processes, a CNN can construct fully connected layers to complete the final classification or predictive task using a Softmax classifier [33].

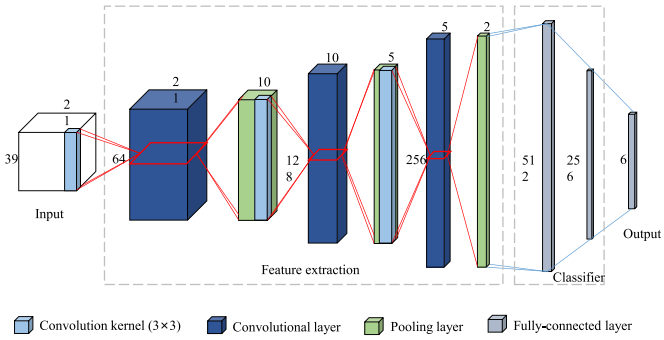

Fig. 2. Basic structure of a CNN.

The learning rate controls the rate of parameter adjustment during network training, which affects whether the network can converge or not. Epochs directly affect the training efficiency of the model. If the values of epoch are too small, underfitting will occur, resulting in insufficient network training; if the values are too large, overfitting will occur, resulting in insufficient model generalization ability. Batch size refers to the number of training samples in one iteration. Batch size is a trade-OFF between computing power and efficiency. Increasing the batch size helps the stability of convergence within a certain range, however, the performance of the model would decrease with the increase of batch size. Therefore, it is necessary to set reasonable parameter values according to the training situation.

A network structure of 10-layer CNN model based on the AlexNet is constructed for geological mapping, including input layer, three convolutional layers with kernel size of 3×3, three pooling layers, two fully connected layers, and output layer (see Fig. 2). Standardization is used in the network to avoid problem with the gradient of the model, and can accelerate the convergence of the model. The last layer of the fully connected layer is a Softmax classifier with an output dimension of six corresponding to six lithological categories in the article area.

D. Random Forests

RF is a machine learning classification algorithm that combines ensemble learning with several decision trees [34]. A RF classifies multidimensional data sources by generating decision trees from randomly selected subsets in the training set using the bagging method [35]. For the training sample *x*, the *k*th decision tree is

$$fk(x) = f(x, \theta_k) \tag{8}$$

where θ_k represents a random vector with an independent distribution. A decision tree is a tree structure classifier that transmits classification results through the root, branch, and leaf nodes. At each nodal point of the decision tree, *m* feature variables were randomly chosen and one branch was selected to grow without pruning. Here, the key factor is determining how to divide the input training set into homogeneous subsets, i.e., the optimal segmentation criteria, such that each subset has the best classification process. Breiman proposed the use of Gini impurity to determine the optimal segmentation threshold [36].

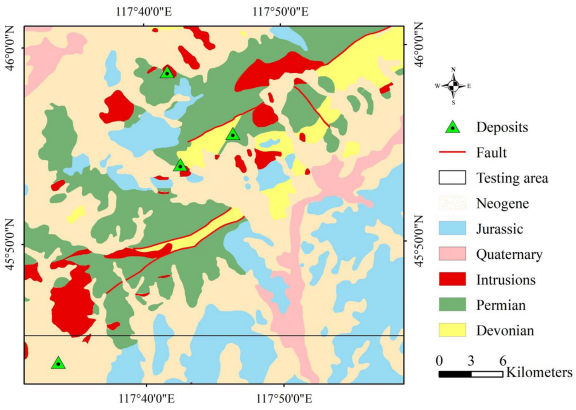

Fig. 3. Simplified geological map of Jilinbaolige, inner Mongolia, China (after from [42]). The training (upper) and testing (lower) area are divided by a black line.

The Gini impurity can be regarded as the probability of misclassifying a random event into its opposite events or minimizing the probability of misclassification, which is expressed as I_G

$$I_G(f) = \sum_{i=1}^{m} f_i(1 - f_i) \tag{9}$$

where f_i represents the probability of the occurrence of class i samples. The smaller the Gini impurity, the better is the classification effect [37]. According to this criterion, the optimal splitting method for each tree was obtained until all the samples in each subset were correctly classified. For all the decision trees, the final classification results were obtained by voting. Here, the probability P_q of the training sample being assessed as q by the random forest is

$$P_q = \frac{1}{T} \sum_{q=1}^{T} z_q^i(x) \tag{10}$$

where T represents the number of decision trees, q is the forecast category, and $z_q^i(x)$ is the ith decision tree that identifies the training sample as the qth predictive result.

III. Study Area and Data

A. Geological Setting

The study area was located in the northeastern part of Inner Mongolia, China, where most of the land was covered by grassland. The study area is a key Ag polymetallic metallogenic belt in northern China [38]. The geological formations exposed in the study area range from the Paleozoic to the Mesozoic, including predominantly Jurassic, Permian, and Devonian strata. The Devonian strata are mainly composed of slate, sandstone, siltstone, and pyroclastic rocks [39]. The Permian strata mainly include tuff, tuffaceous sandstone, siltstone, slate, and andesitice porphyrite. Jurassic strata are composed of tuff, basalt, rhyolite, and breccia lava. Ternary and Quaternary sediments also occupy a large proportion, mainly sand, clay, and mud [40]. Intrusions include a few small intrusions with moderate acid and alkalescency, covered by sediments, distributed throughout the area, mostly stretching along the NE-trending belt [41], [42], [43] (see Fig. 3). For more detailed geological data, see Liu [44] and Yang [45]. Table I summarizes the detailed petrographic characteristics of the six lithologic units that were widely emerged in the study area.

TABLE I
Description of the Lithological Units

Lithological categories	Petrography	Areas (km^2)
Neogene	Brown red and orange mudstone	531.6
Jurassic	Tuff, basalt, rhyolite, breccia lava and breccia tuff with rhyolite	238.5
Quaternary	Silt and gravel layer	65.8
Intrusions	Granite and porphyry	68.8
Permian	Tuff, tuffaceous sandstone, siltstone, slate and andesite porphyrite	250.1
Devonian	Tuffaceous sandstone, fine sandstone, tuff and slate	72.7

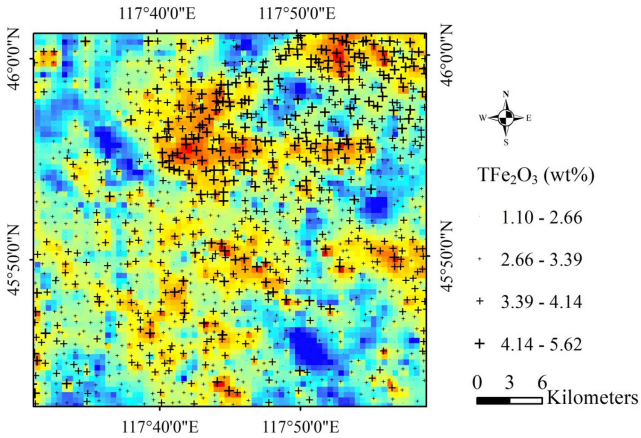

Fig. 4. Geochemical samples and concentration distribution of TFe_2O_3.

B. Data

A total of 1412 stream sediment geochemical samples with a scale of 1:200 000 were collected (sampling density of 1-2 per km^2) in the study area [42]. Sampling locations were irregularly distributed across the study area (see Fig. 4). Each sample identifies 39 major and trace elements using inductively coupled plasma-mass spectrometry (for Bi, Cd, Co, Cu, Mo, Nb, Ni, Pb, Th, U, W, Y, and Zn), X-ray fluorescence (for Ba, Cr, Mn, P, Ti, Zr, SiO_2, Al_2O_3, TFe_2O_3, and K_2O), and other methods (for Be, La, Li, Sr, V, MgO, CaO, Na_2O, Ag, B, Sn, As, Sb, Hg, F, and Au) [43]. The basic statistics of 39 elements were shown in Table II. These data were subsequently explored by Wang and Zuo [41], [43]. Geochemical data are typically compositional data because the content of each element is related to the content of the whole sample, which may lead to false correlations between elements [46], [47]. Therefore, it is necessary to preprocess the composition data to eliminate the influence of closure effects. The central log-ratio transformation [48], [49] was adopted to handle the geochemical compositional data with the support of the open-source software CoDaPack [50]. Geochemical images with a spatial resolution of 500 m $\times$ 500 m were obtained using the inverse distance weighted interpolation method (see Fig. 4).

TABLE II
BASIC STATISTICS OF 39 ELEMENTS

Element	Maximum	Minimum	Average	Standard deviation	Variance
Au	14.30	0.06	1.00	0.78	0.61
Ag	185.08	21.91	79.39	21.52	462.99
As	53.16	0.56	8.16	2.94	8.64
B	151.00	11.10	32.02	10.41	108.46
Ba	783.00	278.70	565.25	39.93	1594.02
Be	3.06	1.35	2.15	0.25	0.06
Bi	2.79	0.10	0.27	0.12	0.01
Cd	316.57	18.80	97.34	33.68	1134.64
Co	21.49	2.41	7.97	2.53	6.39
Cr	74.97	11.30	43.00	10.76	115.84
Cu	42.66	4.39	14.71	5.07	25.72
Hg	55.04	5.05	15.93	5.29	28.03
La	42.62	9.17	23.66	5.75	33.09
Li	55.29	11.27	23.95	5.83	33.94
Mn	1269.70	159.60	460.01	131.67	17337.12
Mo	4.31	0.19	0.64	0.25	0.06
Nb	17.25	4.80	11.49	1.95	3.79
Ni	36.47	4.13	17.76	5.95	35.46
P	1271.70	186.30	543.31	139.23	19384.37
Pb	47.78	13.61	20.44	3.51	12.33
Rb	130.50	62.10	102.76	9.80	96.10
Sb	6.68	0.34	0.76	0.28	0.08
Sn	30.29	1.00	2.46	1.63	2.67
Sr	1800.90	128.50	258.93	134.82	18176.53
Th	13.48	3.05	8.04	1.92	3.70
Ti	7033.70	1269.10	3215.63	549.95	302444.55
U	31.40	0.66	1.77	1.46	2.12
V	105.80	21.40	59.74	12.86	165.27
W	3.67	0.56	1.35	0.32	0.11
Y	40.98	8.34	22.02	4.38	19.21
Zn	163.46	16.60	47.24	13.18	173.58
Zr	639.20	71.20	301.68	84.83	7196.86
SiO_2	72.62	30.14	61.03	7.39	54.60
Al_2O_3	15.32	6.15	12.48	1.29	1.68
TFe_2O_3	5.62	1.10	3.30	0.77	0.59
MgO	10.56	0.41	1.51	1.01	1.02
CaO	23.47	0.84	3.91	4.15	17.21
Na_2O	6.26	0.61	2.01	0.44	0.20
K_2O	3.56	1.36	2.72	0.36	0.13

Note: the units of Ag and Cd are 10^{-9}; the units of SiO_2, Al_2O_3, TFe_2O_3, MgO, CaO, Na_2O, and K_2O are wt%; and the units of other elements are 10^{-6}.

Two types of remote sensing images, ASTER and Sentinel-2A, collected on August 28, 2003, and June 1, 2020, respectively, were obtained from the US Geological Survey (https://earthexplorer.usgs.gov). Sentinel-2A remote sensing images included 11 visible and near infrared bands (VNIR) and two short-wave infrared bands (SWIR), with spatial resolutions of 10 m, 20 m, and 60 m [51]. The ASTER remote sensing images in the study area included three VNIR bands with resolutions of 15 m and six SWIR bands with resolutions of 30 m. These two types of remote sensing data not only have similar spectral response functions in the same spectrum but can also integrate their respective advantages [52]. Therefore, a combination of these two images was used for geological mapping.

The preprocessing of multispectral remote sensing data includes image mosaic and clipping, radiometric calibration, and atmospheric corrections. The ASTER images were corrected using the FLAASH atmospheric correction module in ENVI 5.3 software [53]. Atmospheric correction of the Sentinel-2A images was carried out using the Sentinel Application Platform and Sen2cor software package supplied by the European Space Agency [54].

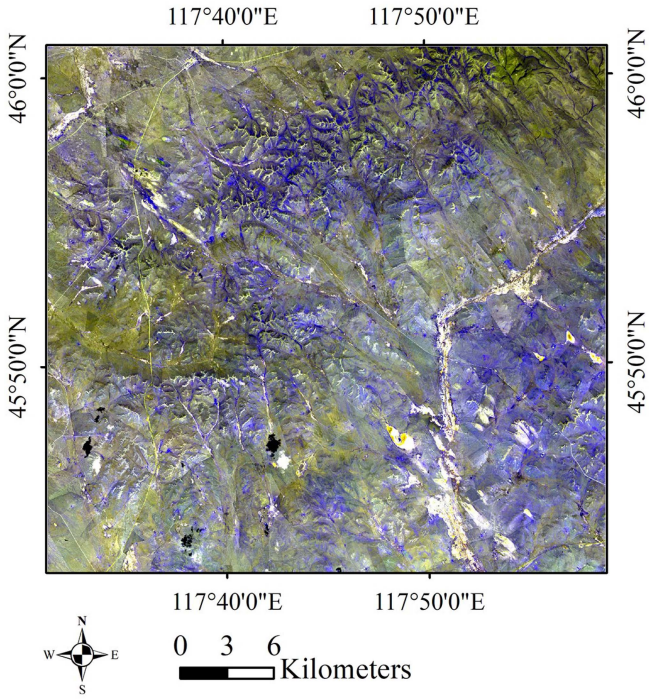

Fig. 5. Fused remote sensing images of the study area (false color composite of bands 3, 2, and 1).

IV. RESULTS AND DISCUSSION

A. Data Fusion

ASTER images have a high spectral resolution, providing more diagnostic information to identify ground objects [55], and Sentinel-2A images provide higher spatial resolution and stronger information interpretation ability. The GS method was adopted to fuse ASTER and Sentinel-2A images to improve the overall spatial resolution and enrich the diagnostic information (see Fig. 5). For example, the spatial resolution of ASTER Band 5 was increased from 30 m to 10 m, and rich texture information and clear edge details were obtained (see Fig. 6).

Fig. 7 displays a fused image of the TFe_2O_3 and remote sensing images obtained using the multisource data fusion method. The spatial resolution of TFe_2O_3 improved from 500 m (see Fig. 4) to 10 m and the fused geochemical pattern of TFe_2O_3 was consistent with that of the original TFe_2O_3. This method not only complements some areas without geochemical survey data but also retains the original element distribution pattern and rich spatial details and texture structure in the remote sensing images.

B. Training Samples

The study area was divided into a training area (upper) and a testing area (lower) at a ratio of 6:1 (see Fig. 3). At the center of each lithologic unit of the training area, 2500 evenly distributed points were selected. The samples with spatial characteristics

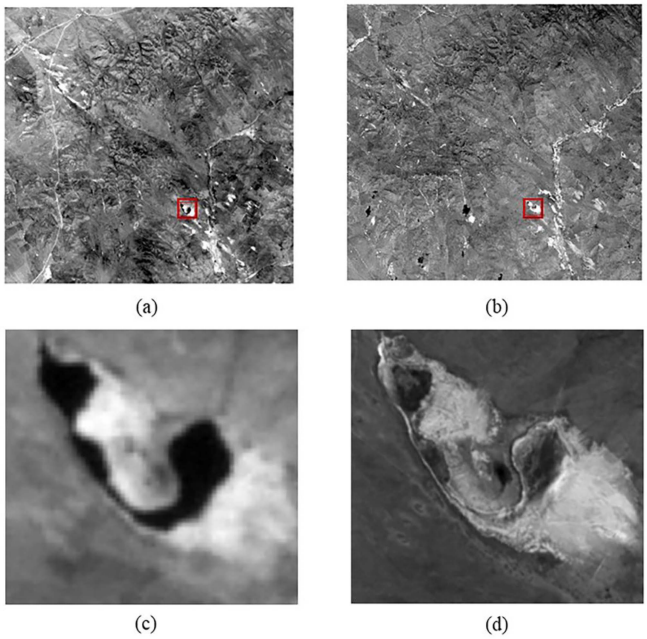

Fig. 6. ASTER Band 5: (a) Original image, (b) fused image obtained by the GS fusion technology, (c) partial magnification of (a) (200 × 200 pixels), and (d) partial magnification of (b).

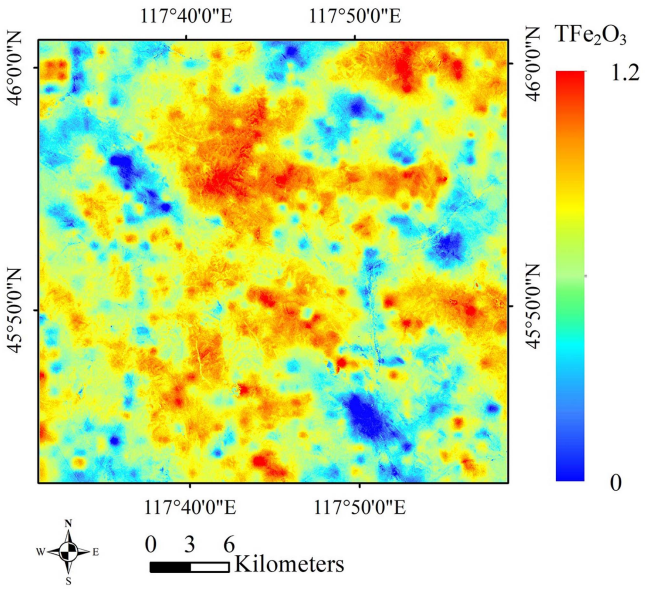

Fig. 7. Fusion image of TFe_2O_3.

were created with each point as the center using the sliding window technique, and the samples were labeled by the center point category (see Fig. 8). Among them, 2000 and 500 samples were selected as the training and validation datasets, respectively.

The size of window determines the extraction of spatial pattern features and the design of network structure. In this article, the smallest identifiable lithologic unit size was 21 × 21, therefore, the window size was set to 21 × 21 with a step size of 1 after comparing the classification accuracy under different window sizes.

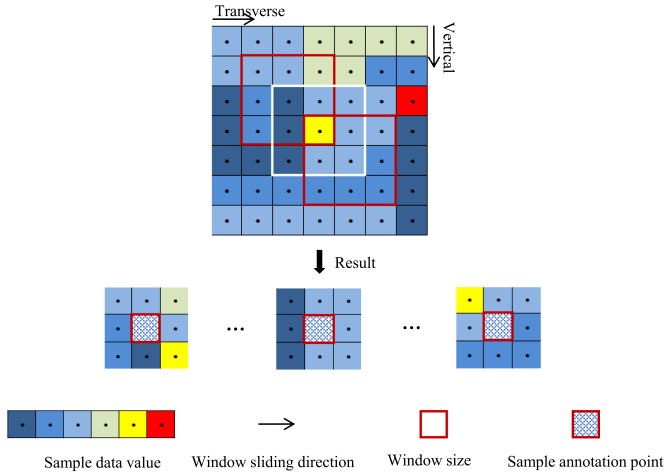

Fig. 8. Diagram showing a sliding window technique for creating training samples.

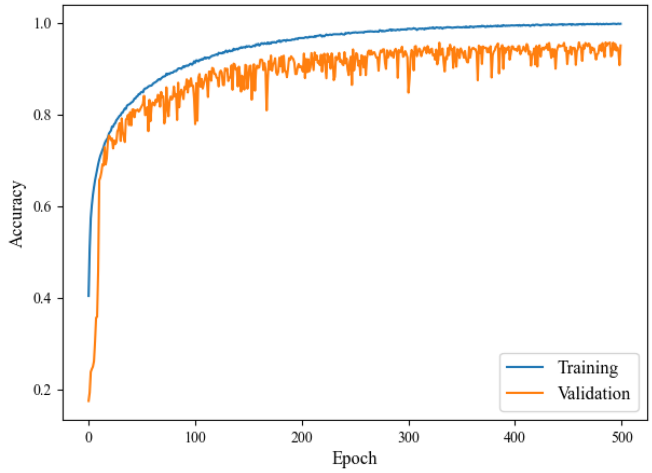

Fig. 9. Classification accuracy of training and validation datasets.

C. Geological Mapping

The CNN model was trained based on the training and validation datasets, and the trained CNN model was then used to classify the fused data samples using a sliding window of 21 × 21 pixels in the testing area. The input dimension of each sample data was 21 × 21 × 39, where 21 and 21 were the length and width of the sample, and 39 represented the number of channels of the sample fused from the geochemical data and the remote sensing bands. The parameters were optimized based on the grid search method to construct the optimal network model structure, in which the RMSProp algorithm [56] was adopted for the gradient descent, the cross-entropy was selected as the loss function. The hyperparameters of batch size, epochs, and learning rate were set to 128, 500, and 10^{-5}, respectively. The loss function versus classification accuracy of the training and validation datasets are shown in Figs. 9 and 10, indicating that the CNN model was well trained effectively. The resulting classification map (see Fig. 11) obtained by the CNN model indicated that most of the lithologic units were classified in

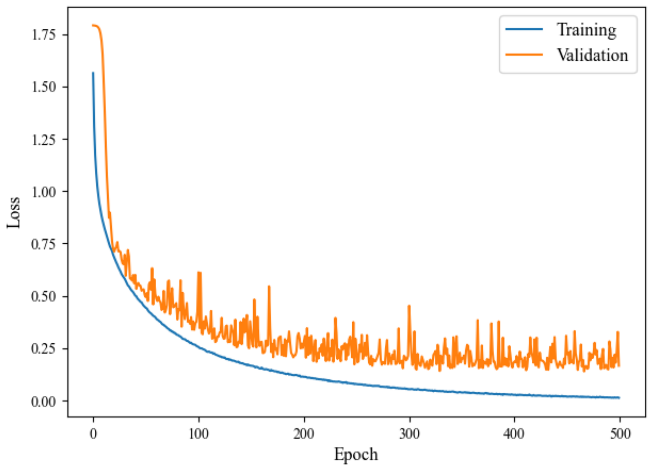

Fig. 10. Loss function of training and validation datasets.

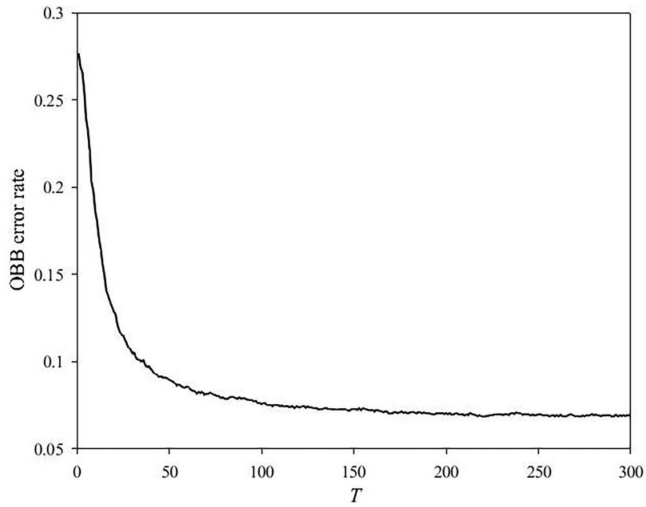

Fig. 12. Number of decision trees in random forest.

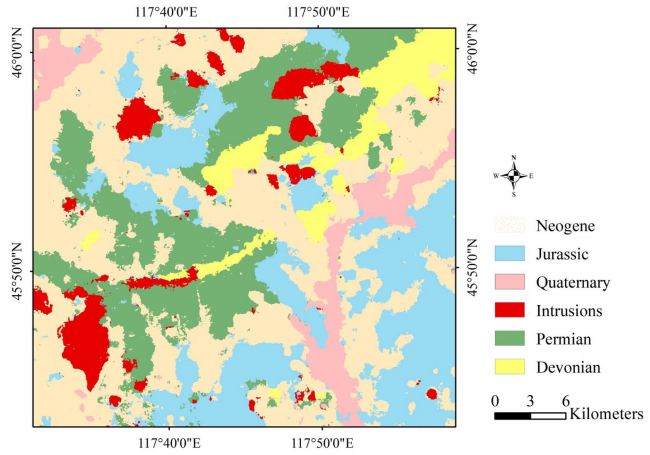

Fig. 11. Classification map obtained by CNN and fused data.

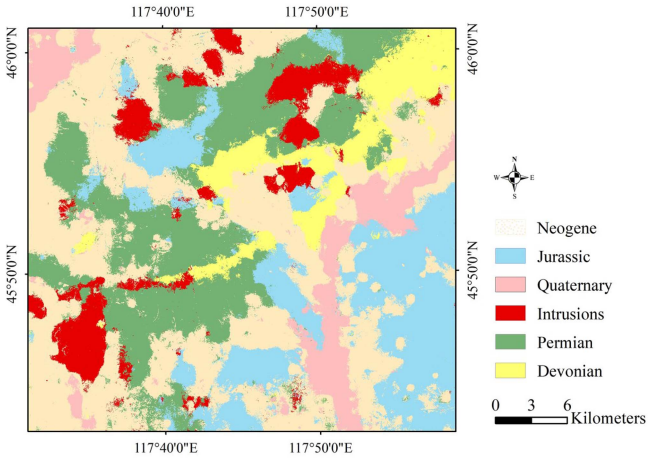

Fig. 13. Classification map obtained by RF model and fused data.

agreement with the geological map (see Fig. 3) with an overall classification accuracy of 83%.

RF was used to map the lithological units and was compared with the CNN model. The parameters need to be set for the RF were the number of decision trees (T) and the characteristic variables selected at each node (m). The out-of-bag error was used to evaluate the ability of the RF to fit the training data and was employed to estimate the classification accuracy of each lithological unit under the optimum parameters. m is the arithmetic square root of the number of geochemical elements, with a value of 6. Fig. 12 shows the out-of-bag error rate obtained by the classification model with different values of T. When $T = 300$, the out-of-bag error rate is the lowest, and the model achieves the best classification effect on the test set. A total of 2500 samples (pixels) were randomly selected as training samples from the center of each lithologic unit. Fig. 13 presents the classification map obtained by RF, in which some misclassifications can be found on the boundaries among various lithologic units. For example, a part of the Jurassic strata is wrongly classified as Permian and Neogene strata, resulting in a relatively low classification accuracy of 78%.

D. Comparative Experiments and Analysis

Comparing the two classification maps obtained by the CNN and RF, CNN had a better performance than RF and further improved the overall classification accuracy by approximately 5%. RF is a pixelwise classification model that ignores the spatial characteristics of the neighboring data, which may lead to the salt-and-pepper phenomenon in the obtained classification map. CNN can solve such problems because it considers the spatial dependence of nearby regions. Moreover, the boundaries of the different lithologic units identified by the CNN were clearer than those identified by the RF.

To further validate the improvements in lithologic mapping using multisource data fusion and a CNN, several comparative experiments were conducted using 1) CNN and fused multisource data (RG_f + CNN), 2) RF and fused multi-source data (RG_f + RF), 3) CNN and fused remote sensing images (ASTER + Sentinel-2A + CNN), and 4) RF and fused remote sensing images (ASTER + Sentinel-2A + RF). The abovementioned tests were implemented using the same parameters and computer environment.

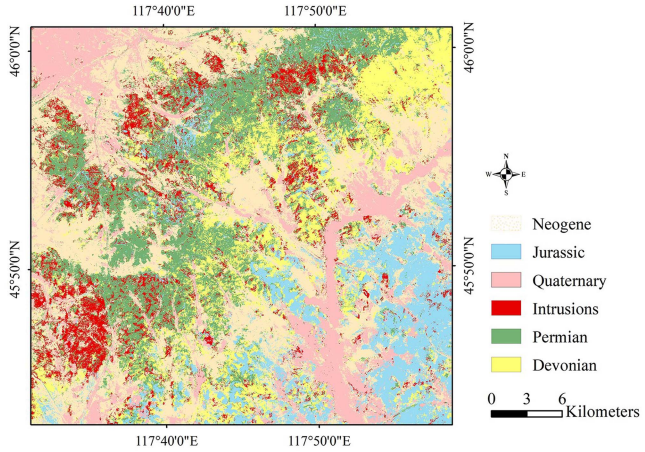

Fig. 14. Classification map obtained by CNN based on ASTER and Sentinel-2A fusion images.

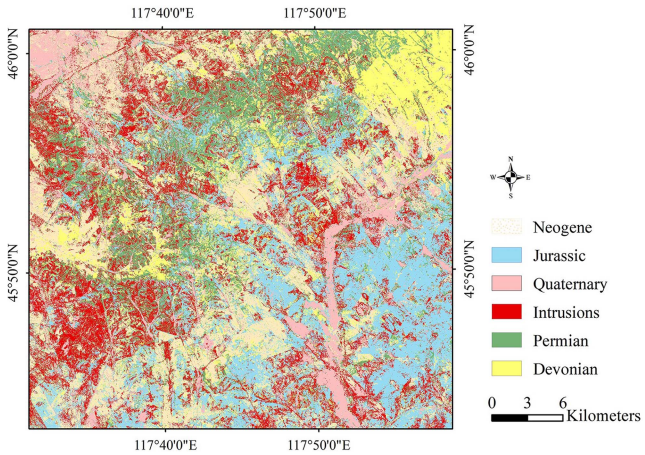

Fig. 15. Classification map obtained by RF based on ASTER and Sentinel-2A fusion images.

Figs. 14 and 15 show the classification map based on CNN and RF using ASTER + Sentinel-2A. From a visual perspective, it is difficult to distinguish the boundaries of lithologic units, implied that the fusion of remote sensing images (ASTER + Sentinel-2A) performed worse than that of multisource data (ASTER + Sentinel-2A + geochemical data). The abovementioned observations demonstrate excellent performance in geological mapping when using a combination of remote sensing and geochemical data.

From the perspective of quantitative index evaluation, this article counted the overall accuracy, classification accuracy of each category, and kappa coefficient obtained from four groups of experiments (see Table III). The kappa coefficient is a quantitative measure for predicting consistency between categories and reference types [57]. A larger kappa coefficient indicates a better classification performance of the model on the whole or individual. The resulting classification of RG_f + CNN was superior to other results and had the highest overall classification accuracy. All four experiments provided a satisfactory classification of the

TABLE III
CLASSIFICATION PERFORMANCE STATISTICS FOR EACH LITHOLOGICAL UNIT

Lithological categories	RG_f+ CNN	RG_f+ RF	ASTER+ Sentinel-2A+CNN	ASTER+ Sentinel-2A+RF
Neogene	88%	85%	73%	67%
Jurassic	76%	79%	69%	45%
Quaternary	77%	63%	31%	29%
Intrusions	76%	68%	29%	13%
Permian	85%	80%	59%	45%
Devonian	80%	65%	25%	22%
Overall accuracy	83%	78%	52%	39%
Kappa	0.79	0.74	0.48	0.36

Neogene strata, which covered a large study area. However, compared with the other methods, data fusion with a CNN promoted the distinguishment of the Quaternary strata, Intrusions, and Devonian strata, which are developed in a small area. The intrusions identified using ASTER + Sentinel-2A + RF were mixed with Jurassic and Permian strata, respectively, resulting in a low classification accuracy of 13%. After fusing geochemical survey data, the classification accuracy of the intrusions increased to 68%. The accuracy in classifying the Jurassic and Permian strata also increased by 34% and 35%, respectively. Meanwhile, for the Devonian strata identified using ASTER + Sentinel-2A + CNN, the classification accuracy was improved from 25% to 80% with the support of the fused data, and the classification results were significantly improved with clearer geological boundaries that are more coincident with the geological map. Compared with RG_f + CNN and RG_f + RF, a CNN could improve the overall accuracy by 5% and the kappa coefficient by 0.05. For the Neogene, Jurassic, and Permian strata, which comprise a large proportion of the exposed area, the classification accuracy was not greatly improved because this type of lithologic unit was sufficiently represented in the training samples to construct and optimize the RF and CNN model. However, a CNN enhances the classification accuracy for lithologic units comprising a small proportion of the exposed area, providing limited training samples, such as Quaternary strata. These findings indicate the potential of CNNs for lithological mapping.

In summary, with the help of the fused data and a CNN, a high classification accuracy of 83% was achieved for the vegetation coverage areas. Fused data provided abundant information on lithologic units. Furthermore, a CNN can extract deep-level spatial characteristics from fused data that contribute to mapping lithologic units. However, some lithologic units are still misclassified by others, with an unclear contact zone and boundary, especially for Jurassic strata. This may be due to the strong metamorphism and weathering, sampling errors, similarities in the spectral characteristics, and the uncertainty of surveying and mapping of the original geological map.

V. Conclusion

Geological mapping in vegetation coverage areas is challenging. In this article, a geological mapping framework combining multi-data fusion and a CNN was applied to identify lithologic units in Jilinbaolige, Inner Mongolia, China. The following conclusions can be drawn: 1) multisource data fusion provides abundant information from different perspectives. The combination of remote sensing images and geochemical survey data is beneficial for realizing the complementary advantages of different kinds of data and affords abundant diagnostic information for geological mapping; 2) A CNN model performs better and improves the overall accuracy in classifying the lithologic units by 5% than RF model. This excellent performance indicated that the proposed framework is effective for geological mapping. How to improve the accuracy of geological mapping is a continuous work. In the follow-up study, more classification features such as structure, rock mass, and aeromagnetic suitable for the study area will be added to further optimize the classification results. In recent years, a wide variety of deep learning models have emerged. In the future, newly developed models and technologies will be used to improve the model and explore the neural network structure suitable for geological mapping.

References

[1] Z. Wang, R. Zuo, and L. Jing, "Fusion of geochemical and remote-sensing data for lithological mapping using random forest metric learning," *Math. Geosci.*, vol. 53, pp. 1125–1145, 2021.

[2] G. R. Hunt, *Spectroscopic Properties of Rocks and Minerals*. Boca Raton, FL, USA: CRC Press, 2017, pp. 295–386.

[3] W. Wang and Q. Cheng, "Mapping mineral potential by combining multiscale and multi-source geoinformation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2008, pp. II-1321–II-1324.

[4] H. He, X. Yang, Y. Li, Y. Li, and L. Wang, "Multi-source data fusion technique and its application in geological and mineral survey," *J. Earth Sci. Environ.*, vol. 32, no. 1, pp. 44–47, 2010.

[5] W. P. Loughlin, "Principal component analysis for alteration mapping," *Photogrammetric Eng. Remote Sens.*, vol. 57, no. 9, pp. 1163–1169, 1991.

[6] S. Kuhn, M. J. Cracknell, and A. M. Reading, "Lithologic mapping using random forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia," *Geophysics*, vol. 83, no. 4, pp. 183–119, 2018.

[7] J. R. Harris and E. C. Grunsky, "Predictive lithological mapping of Canada's North using random forest classification applied to geophysical and geochemical data," *Comput. Geosci.*, vol. 80, pp. 9–25, 2015.

[8] B. Demir and S. Ertürk, "Empirical mode decomposition of hyperspectral images for support vector machine classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4071–4084, Nov. 2010.

[9] A. Gasmi, C. Gomez, H. Zouari, A. Masse, and D. Ducrot, "PCA and SVM as geo-computational methods for geological mapping in the Southern of Tunisia, using ASTER remote sensing data set," *Arabian J. Geosci.*, vol. 9, pp. 1–12, 2016.

[10] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc., Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[11] L. Yu, A. Porwal, E. Holden, and M. C. Dentith, "Towards automatic lithological classification from remote sensing data using support vector machines," *Comput. Geosci.*, vol. 45, pp. 229–239, 2012.

[12] N. Li, M. Frei, and W. Altermann, "Textural and knowledge-based lithological classification of remote sensing data in Southwestern Prieska sub-basin, Transvaal Supergroup, South Africa," *J. Afr. Earth Sci.*, vol. 60, no. 4, pp. 237–246, 2011.

[13] Y. He et al., "A unified network of information considering superimposed landslide factors sequence and pixel spatial neighbourhood for landslide susceptibility mapping," *Int. J. Appl. Earth Observ. Geo-Inf.*, vol. 104, 2021, Art. no. 102508.

[14] Y. LeCun, Y. Bengio, and G. E. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[15] H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, "Exploring strategies for training deep neural networks," *J. Mach. Learn. Res.*, vol. 1, no. 10, pp. 1–40, 2009.

[16] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6,011,875, 2000.

[17] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, 2000.

[18] C. Li, L. Liu, J. Wang, and R. Wang, "Comparison of two methods of fusing remote sensing images with fidelity of spectral information," *J. Image Graph.*, vol. 9, no. 11, pp. 1376–1385, 2004.

[19] H. Yu, B. Yan, F. Gan, W. Chi, and F. Wu, "Hyperspectral image fusion by an enhanced gram schmidt spectral transformation," *Geography Geo-Inf. Sci.*, vol. 23, no. 5, 2007, Art. no. 3042.

[20] D. Huang, "Research of pixel level remote sensing image fusion method," Ph.D. dissertation, Central South Univ., Changsha, China, 2011.

[21] T. Zhang, J. Liu, K. Yang, W. Luo, and Y. Zhang, "Fusion algorithm for hyperspectral remote sensing image combined with harmonic analysis and Gram–Schmidt transform," *Acta Geodaetica et Cartographica Sinica*, vol. 44, no. 9, pp. 1042–1047, 2015.

[22] L. Jing and Q. Cheng, "A technique based on non-linear transform and multivariate analysis to merge thermal infrared data and higher-resolution multispectral data," *Int. J. Remote Sens.*, vol. 31, no. 24, pp. 6459–6471, 2010.

[23] H. Ding et al., "A method and system for improving the resolution of geochemical layers," Chinese patent no. 201811275285.4, 2018.

[24] E. H. Adelson, C. H. Anderson, J. R. Bergen, and P. J. Burt, "Pyramid methods in image processing," *RCA Engineer*, vol. 29, no. 6, pp. 33–41, 1984.

[25] B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, and I. Pippi, "Multispectral fusion of multisensor image data by the generalized Laplacian pyramid," in *Proc. Int. Geosci. Remote Sens. Symp.*, 1999, pp. 1183–1185.

[26] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[27] J. Huang, J. Li, and Y. Gong, "An analysis of convolutional neural networks for speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2015, pp. 4989–4993.

[28] Q. Tian and M. Wang, "Research progress on deep learning algorithms," *Comput. Eng. Appl.*, vol. 55, no. 22, pp. 25–33, 2019.

[29] J. He, J. R. Harris, M. Sawada, and P. Behnia, "A comparison of classification algorithms using Landsat-7 and Landsat-8 data for mapping lithology in Canada's Arctic," *Int. J. Remote Sens.*, vol. 36, no. 8, pp. 2252–2276, 2015.

[30] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 807–814.

[31] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.

[32] J. Cai, J. Cai, X. Liao, H. Huang, and Q. Ding, "Preliminary study on hand gesture recognition based on convolutional neural network," *Comput. Syst. Appl.*, vol. 24, no. 4, pp. 113–117, 2015.

[33] Z. Jiang, H. Lin, H. Zhang, and F. Jiang, "Remote sensing image of GF-1 classification using object-oriented method and convolutional neural network," *J. Central South Univ. Forestry Technol.*, vol. 41, no. 8, pp. 45–55, 2021.

[34] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[35] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.

[36] L. Breiman, "Probability and stochastic processes: With a view toward applications," *Course Technol.*, 1986.

[37] B. H. Menze et al., "A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data," *Biomed. Central Bioinf.*, vol. 10, no. 1, pp. 1–16, 2009.

[38] J. Wang, "Metallogeny of Dongwu banner copper and silver ploymetallic ore zone of Inner Mongolia," *Mineral Resour. Geol.*, vol. 17, pp. 132–135, 2003.

[39] Z. Huang et al., "Mineralization regularity and prospecting direction of polymetallic metallogenic belt in Dongwuqi, Inner Mongolia," *Geol. Surv. Res.*, vol. 36, pp. 205–212, 2013.

[40] J. Wang, R. Zuo, and J. Caers, "Discovering geochemical patterns by factor-based cluster analysis," *J. Geo-Chem. Exploration*, vol. 181, pp. 106–115, 2017.

[41] J. Wang and R. Zuo, "A MATLAB-based program for processing geochemical data using fractal/multifractal modeling," *Earth Sci. Inform.*, vol. 8, pp. 937–947, 2015.
[42] H. Liu, Q. Chi, W. Wang, X. Wang, and J. Zhou, "Study on geochemical exploration methods in mountainous and hilly grassland coverage areas in central and Eastern Inner Mongolia," *Geo-Phys. Geo-Chem. Exploration*, vol. 37, no. 3, pp. 382–388, 2013.
[43] J. Wang and R. Zuo, "An extended local gap statistic for identifying geochemical anomalies," *J. Geo-Chem. Exploration*, vol. 164, pp. 86–93, 2016.
[44] H. Liu, "Geochemical methods for semi–arid grasslands: A case study in Dong Ujimqinqi, Inner Mongolia," Ph.D. dissertation, China Univ. of Geosciences, Changsha, China, 2011.
[45] K. Yang, "Ore-forming characters and prospecting orientation of metal deposit in the Dong Ujimqin, Inner Mongolia," Ph.D. dissertation, China Univ. of Geosciences, Changsha, China, 2013.
[46] R. Zuo, Q. Xia, and H. Wang, "Compositional data analysis in the study of integrated geochemical anomalies associated with mineralization," *Appl. Geo-Chem.*, vol. 28, pp. 202–211, 2013.
[47] P. Filzmoser, K. Hron, and C. Reimann, "Principal component analysis for compositional data with outliers," *Environmetrics, Official J. Int. Environmetrics Soc.*, vol. 20, no. 6, pp. 621–632, 2009.
[48] J. Aitchison, "The statistical analysis of compositional data," *J. Roy. Stat. Soc., Ser. B (Methodological)*, vol. 44, no. 2, pp. 139–160, 1982.
[49] J. Aitchison, C. Barceló-Vidal, J. A. Martín-Fernández, and V. Pawlowsky-Glahn, "Logratio analysis and compositional distance," *Math. Geol.*, vol. 32, no. 3, pp. 271–275, 2000.
[50] M. Comas-Cufí and S. Henestrosa, "Codapack 2.0: A stand-alone, multiplatform compositional software," in *Proc. 4th Int. Workshop Compositional Data Anal.*, 2011.
[51] M. Drusch et al., "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, 2012.
[52] F. D. Van der Meer, H. Van der Werff, and F. Van Ruitenbeek, "Potential of ESA's Sentinel-2 for geological applications," *Remote Sens. Environ.*, vol. 148, pp. 124–133, 2014.
[53] T. Cooley et al., "FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation," in *Proc. Int. Geosci. Remote Sens. Symp.*, 2002, pp. 1414–1418.
[54] M. Main-Knorn, B. Pflug, J. Louis, V. Debaecker, U. Müller-Wilm, and F. Gascon, "Sen2Cor for Sentinel-2," in *Proc. Image Signal Process. Remote Sens. XXIII*, 2017, vol. 10427, pp. 37–48.
[55] M. Abrams, "The advanced spaceborne thermal emission and reflection radiometer (ASTER): Data products for the high spatial resolution imager on NASA's Terra platform," *Int. J. Remote Sens.*, vol. 21, no. 5, pp. 847–859, 2000.
[56] T. Tieleman and G. Hinton, "Lecture 6.5-RMSprop: Divide the gradient by a running average of its recent magnitude," *COURSERA, Neural Netw. Mach. Learn.*, vol. 4, no. 2, pp. 26–31, 2012.
[57] A. Viera and J. Garrett, "Understanding interobserver agreement: The kappa statistic," *Fam. Med.*, vol. 37, no. 5, pp. 360–363, 2005.

Ting Pan is currently working toward the M.S. degree in geo-detection and information technology, China University of Geosciences, Wuhan, China.

Her research interests include remote sensing geology and geological mapping.

Renguang Zuo received the B.S. and Ph.D. degrees in mineral prospecting and exploration from the China University of Geosciences (CUG), Wuhan, China, in 2004 and 2009, respectively.

He is currently a Full Professor with the State Key Laboratory of Geological Processes and Mineral Resources, CUG. He has authored and coauthored more than 100 international peer-reviewed papers on various journals. He is an Associate Editor for *Computers and Geosciences*, *Journal of Geochemical Exploration*, *Natural Resources Research*, and *Ore Geology Reviews*. His current research interests include mathematical geosciences and mineral exploration.

Ziye Wang received the B.S. degree in mining engineering from Henan University of Technology, Jiaozuo, China, in 2014 and the M.S. degree in geo-detection and information technology from the China University of Geosciences, Wuhan, China, in 2017.

He is currently an Associate Professor with the School of Earth Resources, China University of Geosciences. His research interests include remote sensing geology and geochemical anomaly detection.