Neural Computing and Applications (2020) 32:2037–2053
https://doi.org/10.1007/s00521-019-04341-3

DEEP LEARNING & NEURAL COMPUTING FOR INTELLIGENT SENSING AND CONTROL

Applications of deep convolutional neural networks in prospecting prediction based on two-dimensional geological big data

Shi Li[1,2] · Jianping Chen[1,2] · Jie Xiang[3]

Received: 21 March 2019 / Accepted: 28 June 2019 / Published online: 11 July 2019

Abstract
There are many challenges in the task of predicting ore deposits from big data repositories. The data are inherently complex and of great significance to the intervenient spatial relevance of deposits. The characteristics of the data make it difficult to use machine learning algorithms for the quantitative prediction of mineral resources. There are considerable interest and value in extracting spatial distribution characteristics from two-dimensional (2-d) ore-controlling factor layers under different metallogenic conditions. In this paper we undertake such analysis using a deep convolutional neural network algorithm named AlexNet. Training on the 2-d mineral prediction and classification model is performed using data from the Songtao–Huayuan sedimentary manganese deposit. It mines the coupling correlation between the spatial distribution of chemical elements, sedimentary facies, the outcrop of Datangpo Formation, faults, water system, and the areas where manganese ore bodies are present, as well as the correlation among different ore-controlling factors by employing the AlexNet networks. By comparing the training loss, training accuracy, verification accuracy, and recall of models trained by different scales of grids and different combinations of ore-controlling factor layers, we further discuss the most appropriate scale division and the optimal combination of ore-controlling factors to make the model achieve its strongest robustness. It is found that the prediction performance of AlexNet networks reaches its peak when selecting a grid division of 200 pixels × 200 pixels (the actual distance is 10 km × 10 km) and inputting the distribution layers of 21 chemical elements maps, lithofacies–paleogeographic map, formation and tectonic map, outcrop map of Datangpo Formation, and water system map. The training loss, training accuracy, verification accuracy, and recall of the optimal model are 0.0000001, 100.00%, 86.21%, and 91.67%, respectively. The proposed method is successfully applied to the 2-d metallogenic prediction in Songtao–Huayuan study area. And five metallogenic prospective areas from A to E are delineated with large probability for potential ore bodies.

Keywords Geological big data · Prospecting prediction · Convolutional neural networks · Songtao–Huayuan Mn deposit

1 Introduction

Big data has been the fourth scientific paradigm in recent years, and numerous promising progresses in the quantitative prediction of mineral resources with the help of big data have been achieved. The academician Zhao Pengda once proposed to introduce big data concepts into geosciences and anticipated the idea that digital mineral prospecting can realize a leap from mathematical geology to digital geology and fill the gaps of traditional qualitative prospecting. Additionally, Zhao [1] summarized four geological big data-based theories of prospecting systems and carried out scientific quantitative evaluation and analysis

✉ Jianping Chen
3s@cugb.edu.cn

[1] School of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China

[2] Land Resources Information Development and Research Key Laboratory of Beijing, Beijing 100083, China

[3] MNR Key Laboratory of Metallogeny and Mineral Assessment, Institute of Mineral Resources, CAGS, Beijing 100037, China

on mineral prediction. With the help of prediction methods during the age of big data and the detailed works about the potential evaluation of important mineral resources, Xiao et al. [2] explored the basic theoretical principles for the prediction and evaluation of mineral resources and summarized the major working processes in the digital and information age. Chen et al. [3] emphatically introduced the idea and methodology for constructing "geology cloud" under the big data setting and discussed applications of big data in geosciences. Moreover, Yu et al. [4] proposed new methods driven by geological big data-based models to quantitatively predict and evaluate mineral resources, while Zheng et al. [5] advocated the idea that the national geological information service system could boost the share of geological information under the big data setting, balance the data service and information knowledge service, and serve the national construction and socioeconomic development.

In the last few decades, numerous algorithmic models for the quantitative prediction and evaluation of mineral resources have emerged. Machine learning and the prediction of mineral resources are increasingly interconnected in the geological big data age. As an important branch of artificial intelligence (AI), machine learning provides not only effective means for processing abundant evidence feature layers related to mineral prediction, analyzing big data [6], and identifying and predicting models, but also supplies technical support to construct big data-based intelligent prospecting models [7]. It also possesses the potential for identifying nonlinear relationship between known deposits and evidence layers; thus, it can endow researchers with strong prediction ability. The most extensively applied algorithms during the evaluation of mineral prediction include artificial neural networks (ANNs), regression trees (RTs), support vector machines (SVMs), and random forest (RF). In 2017, by applying deep self-encoded network to the 39 geochemical variable data obtained from the stream sediments survey in certain area of southwestern Fujian, Zuo and Xiong [8] successfully identified the geochemical anomalies related to Fe-polymetallic mineralization. Other scholars also successfully applied the big data-based machine learning to the extraction of anomaly information to further conduct quantitative prediction of mineral resources [9–16]. Obviously, under the setting of polyphyletic, volume, and heteroid geological big data, AI-based quantitative prediction of minerals that satisfies the two main lines of demand and "data link" [17] is the inexorable development trend.

However, the data pattern becomes gradually complex, and the relationship among the data gradually expands along with the growth of data volume, contributing to a greater challenge to classification and prediction, and resulting in poor performance of traditional superficial machine learning algorithms. Deep learning methods were proposed to address these challenges owing to their neural network structures that resemble more the model of human brain. Along the basic principles of traditional superficial structures, deep learning can improve the analysis performance by increasing the number of hidden layers. Neural network methods have been repeatedly applied domestically to geochemical processing. Chen [18] had successfully achieved metallogenic predictions in ore fields and their peripheries by applying evidence weight methods and resilient BP neural networks. Using neural network models to identify nonlinearities, Chen et al. [19] centralized the geological and geophysical information into a neural network model from other levels to carry out a comprehensive prediction of oil–gas reservoirs. In 2016, Yang [20] analyzed and processed data from the soil geochemical survey in certain property of Panzhihua city by integrating deep learning and principal component analysis and successfully constructed the volume model, thereby proving that improved deep learning methods can increase the accuracy and intelligentization level of geochemical analysis. In 2001, Albora et al. [21] applied cellular neural networks (CNNs) of deep learning to separate gravity anomalies from magnetic anomalies and drew the conclusion that, with the help of model trial and data analysis, neural network methods can highlight shallow anomalies. Currently, this algorithm has been successfully applied to separate gravity anomalies of ore bodies from wall rocks in chromite properties [22, 23] and provides more reasonable basis for mineral prediction. Hinton and Salakhutdinov [24] achieved classification of data using deep learning models. And when compared to existing methods, the obtained deep neural network structure accomplishes more abstract characteristics and possesses stronger classification ability and better generalization capability. In 2018, Liu et al. [25] used Zhaojikou Pb–Zn deposit of Anhui Province as a case study, constantly excavated the coupling correlation between the distribution characteristics of Pb and the underground occurrence space of ore bodies with the help of convolutional neural network (CNN) algorithms, and successfully delineated the metallogenic prospects of an unknown area. This demonstrated that convolutional neural networks (CNN) have certain advantages in the extraction of spatial features of chemical elements and prospecting prediction.

Hinton and Alex Krizhevsky et al. announced that AlexNet achieved a breakthrough in deep learning in terms of image recognition on the premise of its TOP-5 error rate of 15.3% largely lower than that of other algorithms [26]. Because the data volume of geological big data is massive, a method capable of extracting spatial distribution characteristics and simultaneously excavating the correlation among various ore-controlling factor layers under different

metallogenic conditions is still missing. On the basis of previous researches, this paper adopts the AlexNet of deep convolutional neural networks (D-CNN) model; trains the relationships between the distribution of known Mn deposits in Songtao–Huayuan and the content of 21 elements, sedimentary facies, outcrop of Datangpo Formation, faults, and water system; and finally obtains the classification model based on the D-CNN. Application of this model to the 2-d prediction in unknown areas of Songtao–Huayuan fills the gaps of previous predicting methods which merely emphasize the quantitative extraction of anomalies and neglect the spatial distribution characteristics and the correlation of different metallogenic conditions.

2 Geological background

This paper uses Songtao–Huayuan Mn deposit as a case study and carries out mineral prediction based on D-CNN. The area of study is located in the adjacent area of Hunan and Guizhou (Fig. 1), the most important accumulation area of Mn resources in China, where a great progress in Mn prospecting has been achieved recently [27]. Songtao–Huayuan belongs to the northeastern edge of the south passive marginal fold-thrust belt in Upper Yangtze paleolandmass and crosses the two fourth-level tectonic units, namely Tongren thrust belt and Fenggang detachment fold belt. The tectonic lines in the area are mainly NNE-trending and NE-trending, and a few are S–N-trending. The morphology of folds is mainly gentle and open, while the faults belong mainly to normal fault or reverse fault in conformity with the strike of strata. The outcrop consists mainly of strata from Jixian Fanjingshan Group to Lower Silurian, while the Middle and Upper Silurian, Upper Paleozoic, as well as the Mesozoic Triassic, Jurassic, Lower Cretaceous, and Cenozoic are all absent. As the main mineralization type in the area, the Late Neoproterozoic sedimentary Mn mineralization occurs essentially at the bottom of the black shale within the first Member of Nanhua System Datangpo Formation in the form of rhodochrosite.

There are many researches on the paleogeographic environment of the area and its ore-controlling role during early Datangpo period of Nanhua Period. On the one hand, some researchers studied the sedimentary facies for analyzing the lithofacies association and lateral change rule, so that they could recover the sedimentary environment and paleogeographic pattern of early Datangpo period according to the geochemical characteristics [28–34]. On the other hand, with the help of lithogeochemical methods, some scholars analyzed the oceanic paleoenvironment conditions of early Datangpo period, including the paleoclimate, paleotemperature, paleosalinity, redox conditions, bioorganic action, and submarine hydrothermal activity [35–37]. Recently, Zhou and Du et al. [38–40] held the opinion that "Datangpo-type" Mn deposits are relevant to the formation and evolution of the Nanhua rift basin derived from the breakup of Rodinia Supercontinent, as well as a series of "graben-horst"-type paleogeographic patterns formed in the adjacent area of Hunan, Guizhou, and Chongqing, and are also controlled by a series of syndepositional faults. Consequently, Mn deposits are products derived from the leakage, sedimentation, and mineralization of ancient natural gas in the series of secondary graben basins. Moreover, they pointed out that Mn

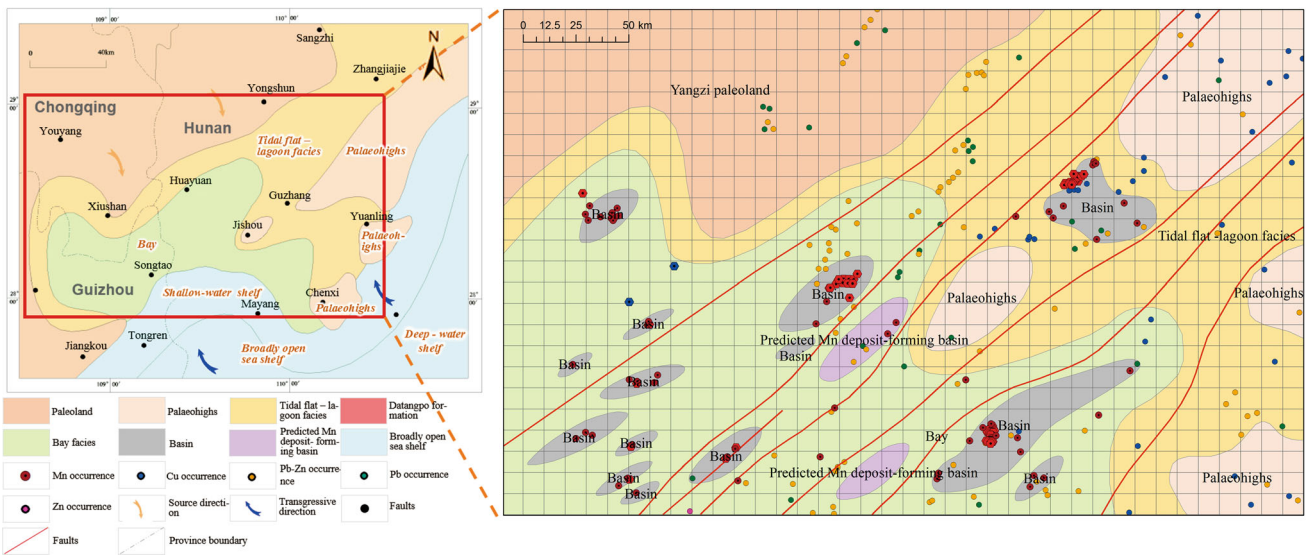


Fig. 1 Lithofacies paleogeographic map of Songtao–Huayuan during the Datangpo period of Nanhua Period

deposits are also influenced by the glacial–interglacial climate of Nanhua Period. And the syndepositional faults are not only the channels for the ascent of deep-seated manganese and ancient natural gas, but also the link connecting the internal system with the superficial system of ancient natural gas leakage-sedimentation-type Mn metallogenic system.

According to the results of previous research, we can infer the following viewpoints: (1) The migration and concentration of manganese are controlled by the paleogeographic pattern of early Datangpo period; (2) the hydrothermal activities caused by seafloor syndepositional faults and seafloor volcanic eruption may provide plentiful metallogenic materials; (3) bioorganic matter plays a significant role in speeding up the precipitation and mineralization of manganese; (4) the various physical and chemical conditions of oceanic paleoenvironment, especially the redox conditions, are important factors influencing the precipitation and precipitation forms of manganese. In this paper, we construct the basic prediction maps for the mineral prediction by collecting the lithofacies paleogeographic maps, formation and tectonic maps, distribution maps of Mn deposits, geochemical distribution maps of Mn, and other maps of Datangpo period.

3 Research methods

The whole framework of our proposed 2-d metallogenic prediction model based on the AlexNet is shown in Fig. 2, and it is mainly composed of three major parts.

The first part is data acquisition and preprocessing. The acquisition of big data is the necessary technical means to achieve informatization from digitization, as well as the premise to realize prospecting from the mining of big data. Since the geological big data are characterized by their massive data volume, wide varieties, and quick update rate, demand analysis is required to build logic structure tree and URL structure tree, and to carry out the collection of data accordingly. The collected data include text data and 2-d data, but the collection of 2-d data should consider the construction of prospecting conceptual model based on the mining of text as its premise, and the acquisition of 2-d data should be conducted under the guidance of prospecting conceptual model. Due to format disparities and coordinates inconsistencies in the collected 2-d data, preprocessing should be carried out on the original data initially. Subsequently, we impose the inverse distance weighted (IDW) interpolation on the content of the geochemical elements and generate TIFF images. Additional procedures on the data consist of georeferencing the data, unifying the coordinates, converting the format, and then segmenting the data. Finally, we take the known positive and negative samples as the training set, and the unknown, unpredicted areas as the testing set.

The second part is prospecting prediction based on AlexNet. Currently, AlexNet network structure has been successfully applied to image processing, but hardly to the processing of geological images. In this work we select 80% of the preprocessed positive and negative samples as the training set, and 20% as the validation set. Different layers are resampled to 224×224 pixels, which are inputted in the form of different bands into the network. Then we apply the spatial distribution characteristics of different factor layers to train the prediction model by using the excellent ability of the network structure to extract spatial characteristics. Therefore, the model simultaneously contains the potential relevance among different factors. Moreover, accuracy and recall are applied to check the classification quality of AlexNet model. After going through the AlexNet network, each test sample eventually get two scores, namely the ore-bearing score and non-ore score.

The third part is output of prediction results and ore-bearing probability. After normalizing the outputted scores of full-connection layer by softmax layer, they are converted into scores ranging from 0 to 1, namely the ore-bearing probability and non-ore probability. The output is then saved as a list in CSV format. Different areas in the divided grid are represented by different colors of the serial number list of known positive and negative samples and that of the prediction results, which are then outputted in the form of JPG format.

With the help of the above technical flow, the polyphyletic and volume 2-d factor layer data can be more intelligently processed, and the inherent spatial distribution characteristics and potential relevance can be used to train the classification model to achieve the purpose of prospecting prediction. This way we can accomplish the 2-d prediction process with t localization and probability determinations, as well as providing a new approach for prospecting prediction based on geological big data.

3.1 Data acquisition and preprocessing

The acquisition of data can be realized by acquiring the corpus from both LAN and public area networks, and it mainly consists of the collection of text data and layer data.

The searching and sieving of LAN data discovered by text data are realized by the secondary development of Everything software on the basis of C# platform, as well as the method of MySQL relational database. The data are acquired through P2P online transmission and FTP off-line transmission. In the face of the polyphyletic, volume, and heteroid public area network data, we propose a double iterative scheme based on keywords searching and URL.

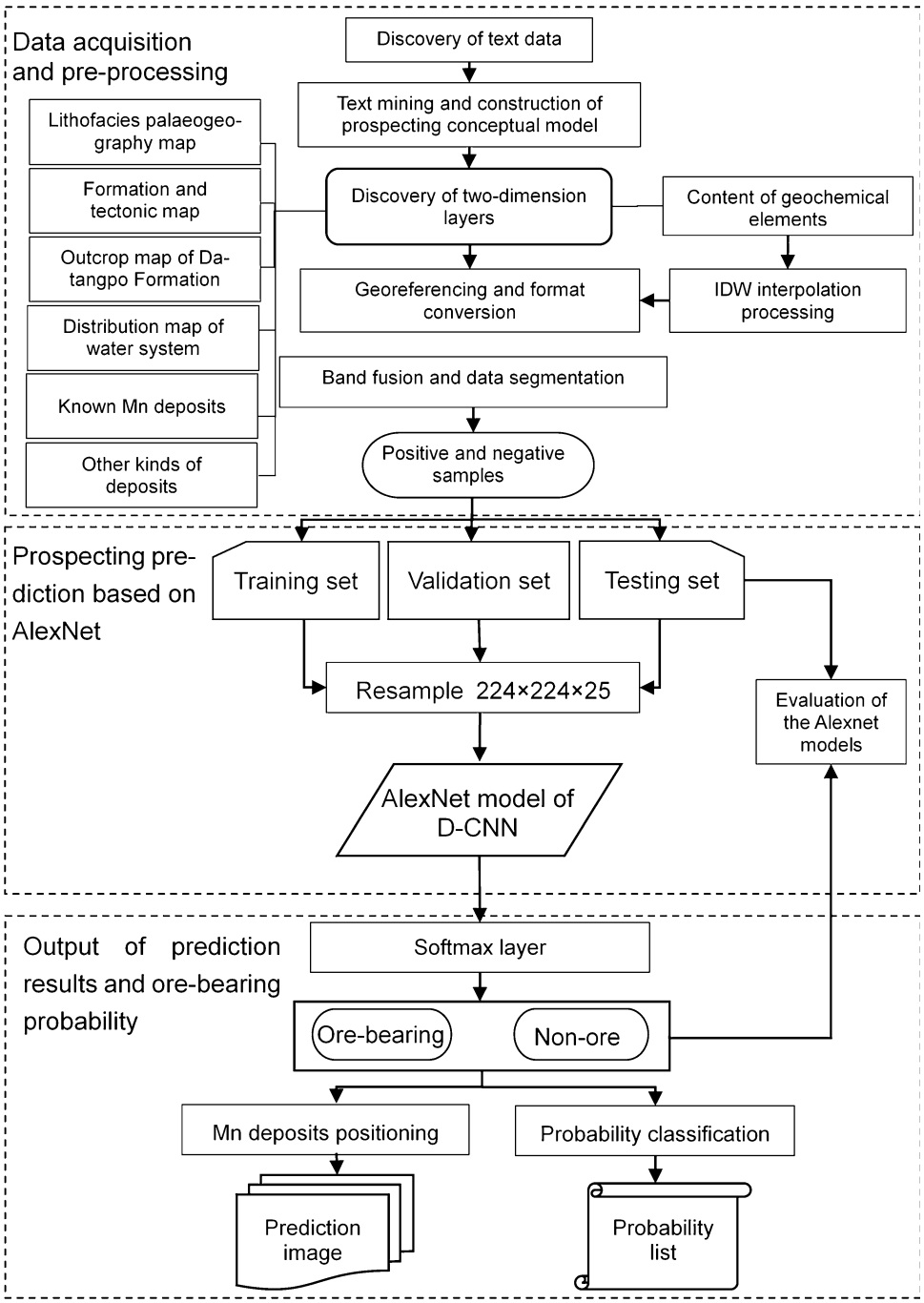


Fig. 2 Technical framework

We construct logic structure tree based on expanded structure nets of associated words in geological dictionaries and generate initial URL seeds by searching keywords with Baidu, Google, and other popular search engines; then we analyze the data content of URL seeds and extract information to give birth to new keywords for supplementing the structure tree; finally, we generate new URL by searching keywords again, namely continuously discover new URL and keywords through the URL seeds and use the machine learning to generate URL structure tree, and then repeat the mutual iteration of both processes again and again to form a comprehensive searching encirclement in both the positive and negative directions. The text data related to Songtao–Huayuan and collected through this research altogether include 543 news, 17,527 pieces of related literature, and 18 regional reports. Eventually, the prospecting conceptual model for Songtao–Huayuan Mn deposit can be obtained by matching the keywords excavated from text data with existing deposit model knowledge based on naive Bayes method. The prospecting model

for the Mn deposit can then be constructed according to the prospecting conceptual model and the discoveries of actual 2-d layer data of the area under study (Table 1).

The total collected data of prospecting factors contain the content statistics table of 39 geochemical elements, distribution map of water system, lithofacies paleogeographic map, formation and tectonic map, outcrop map of Datangpo Formation, distribution map of known Mn deposits, distribution map of known Pb–Zn deposits, distribution map of borehole data, and so on. This experiment employs the IDW interpolation to the 39 geochemical elements, then unifies and matches the coordinates for the distribution maps of 21 geochemical elements of relative importance with the distribution map of water system, lithofacies paleogeographic map of Datangpo period, distribution map of faults, and outcrop map of Datangpo Formation, and then converts these maps into TIFF format (Fig. 3a). Subsequently, it inputs the different factor layers serving as different channels for the regional image of each sample, and divides the data of Songtao–Huayuan into three groups of grid for contrast experiment. The grid sizes from west to east and from north to south successively are 20×20 km^2 (228 (19×12) grids), 10×10 km^2 (912 (38×24) grids), and 5×5 km^2 (3648 (76×48) grids). Figure 3b is an example showing the grid with the grid size of 200 pixels $\times$ 200 pixels (10×10 km^2), the label is 0-911, and the number of inputted layers is 25 (Fig. 3b).

3.2 AlexNet network structure

The AlexNet adopted in this paper was designed by the ImageNet competition champion Hinton and his student Alex Krizhevsky in 2012. The first five layers of this network model are convolutional layers, and the last three layers are full-connection layers. The advantages of AlexNet network over traditional ones are as follows:

3.2.1 Nonlinear rectified linear unit (ReLU) activation function

The AlexNet model adopts the ReLU activation function instead of the previously widely used sigmoid nonlinear activation function. The adoption of ReLU activation function can greatly reduce the computational burden of the whole process and render the training time several folds faster. In a traditional deep network, the gradient would easily disappear during the back-propagation of the sigmoid function (when the sigmoid approaches the saturation area, the derivative would be close to 0 due to the too slow transformation and the information would then be lost), and as a result, the training of deep network cannot be completed. In contrast, ReLU would assign a part of the neural cells output to 0, resulting in the sparsity of the network, interdependence reduction of parameters, and alleviation of overfitting occurrence. The ReLU activation function is given as follows:

Table 1 Prospecting model for the sedimentary Mn deposits in Songtao–Huayuan

Predicting factors	Content
Metallogenic epoch	Datangpo period of Nanhua Period
Geotectonic location	Zhangjiajie–Huayuan fold-thrust belt
Paleogeography	Extensional rift basin
Sedimentary facies	Mn-bearing shale subfacies of semi-restricted bay
Sedimentary sequence	Mainly occurs in two third-level sequences, namely the third sequence (NHS_3) and the fourth sequence (NHS_4) of Nanhua Period, and belongs to the transgressive systems tract (TST) and condensed section (CS)
Paleoclimate	Semi-restricted bay environment; the climate got warm and belonged to the interglacial warm-wet climate; correspondingly, ablative transgression emerged
Sedimentary formation	Black shale formation
Tectonics	The intersection of tectonics, deep faults, and contemporaneous faults of basement
Mn-bearing rock series	Composed of black silt-containing carbonaceous shale and banded rhodochrosite layer. When the thickness of Mn-bearing rock series is larger than 35 m, the thickness of Mn ore body can reach up to 5–7 m and the Mn grade can exceed 24% (Minle), while the thickness of Mn-bearing rock series is less than 10 m and Mn ore body with a certain scale can be hardly formed (Shanmuzhai)
Indicator of strata	Thick moraine conglomerate
Geochemistry	The anomaly area with the value of Mn larger than 1200 ppm has good register relationship with mineralization characteristics

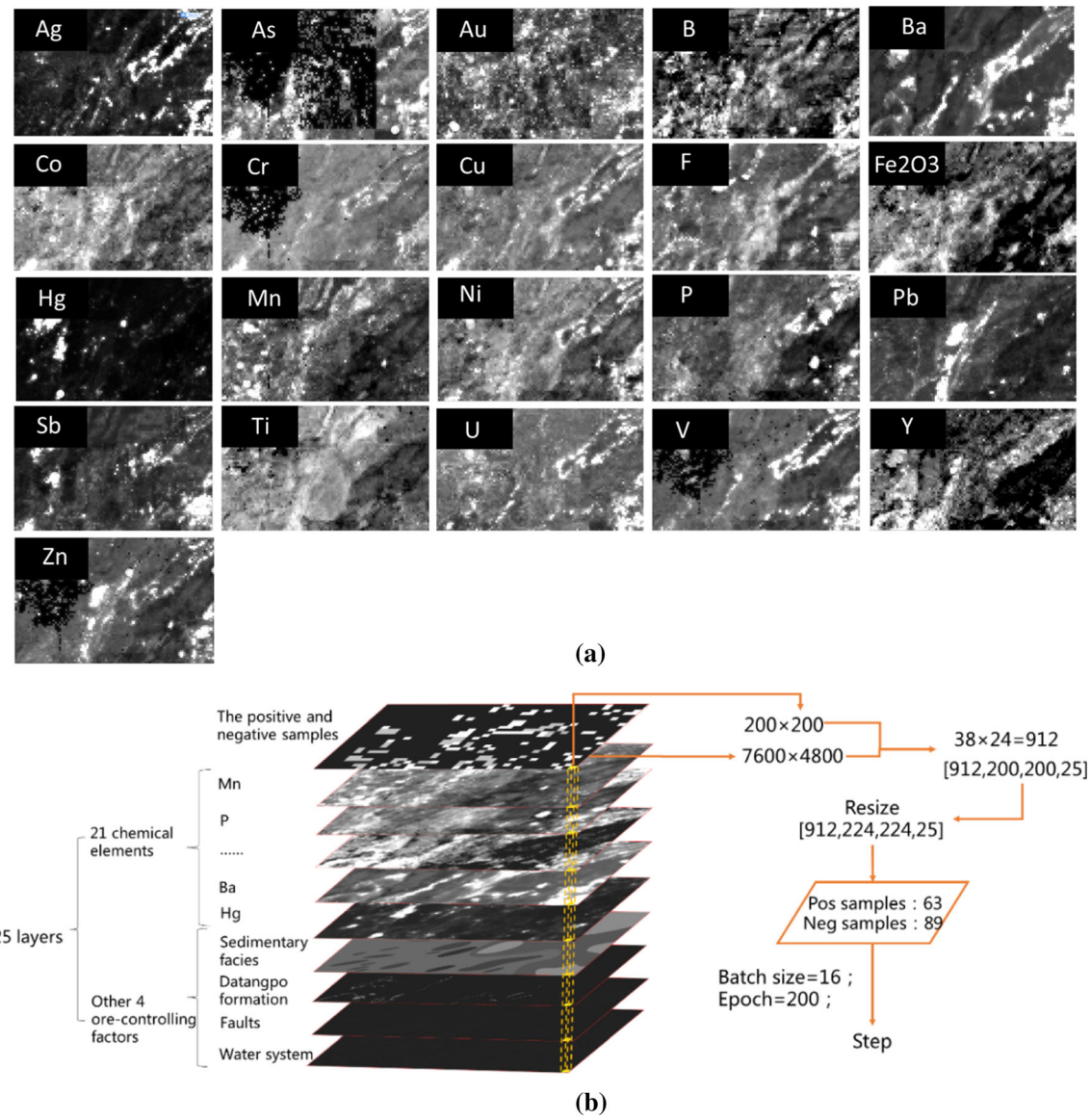


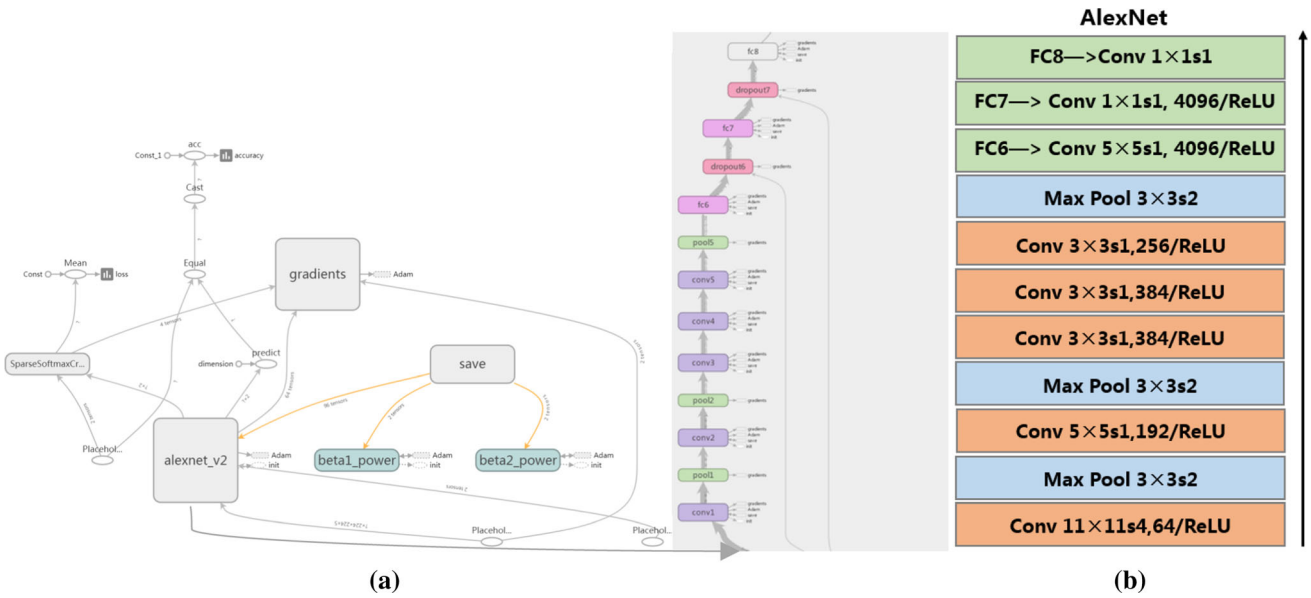
Fig. 3 Preprocessing of 2-d data. **a** Format conversion of the data; the layers include the distribution maps of 21 interpolated geochemical elements, lithofacies paleogeographic map of Datangpo period, outcrop map of Datangpo Formation, distribution map of water system, and distribution map of faults. In the display map of positive and negative samples, the 63 grids of the positive samples correspond to the grid areas (gray ones) where the known Mn deposits are located, the 89 grids of the negative samples are the grid areas (white ones) where other minerals or negative boreholes are located, and the black grids represent the unknown, unpredicted areas. **b** Segmentation of image (the image is 7600 × 4800 pixels, and the horizontal resolution and vertical resolution are both 96 dpi. The image is divided into 912 grids by the grid size of 200 × 200 pixels; therein, 80% of the grids are training samples. Randomly select 16 grids as the training sample data each time and circulate the process for 200 times)

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0. \end{cases}$$

3.2.2 Overfitting prevention of dropout layer

AlexNet can control overfitting by introducing the dropout method with various combinations of weights. During the training process, the activation state of hidden layer neurons can be controlled by threshold values with specific norm and the neurons exceeding the threshold values would be restrained during the forward propagation and back-propagation. During training, a probability of 0.5 is used to randomly inactivate the neurons. As a result, dropout layer reduces the occurrence of overfitting.

Before entering the AlexNet network, the sample sets have been resampled to 224 × 224, which is suitable for the self-defined AlexNet network. The hyper-parameter settings of the AlexNet model adopted in this paper are shown in Fig. 4b:



(a) (b)

Fig. 4 **a** Network structure of prediction model (input layer — > AlexNet network — > softmax layer) and **b** hyper-parameter setting of AlexNet model (where Conv refers to the convolution layer, Max pool refers to the pooling layer, S is the stride, and ReLU is the activation function)

(1) The first layer (convolution layer): convolution kernel = [11, 11], stride = 4, output channel = 64, padding = valid (not extending edge).
(2) The second layer (pooling layer): pool_size = [3, 3], stride = 2, output channel = 64(default), padding = valid.
(3) The third layer (convolution layer): convolution kernel = [5, 5], stride = 1, output channel = 192, padding = same (expand the edge to make the input and output the same size).
(4) The fourth layer (pooling layer): pool_size = [3, 3], stride = 2, output channel = 192, padding = valid.
(5) The fifth layer (convolutional layer): pool_size = [3, 3], stride = 1, output channel = 384, padding = same.
(6) The sixth layer (convolution layer): convolution kernel = [3, 3], stride = 1, output channel = 384, padding = same.
(7) The seventh layer (convolution layer): convolution kernel = [3, 3], stride = 1, output channel = 256, padding = same.
(8) The eighth layer (pooling layer): pool_size = [3, 3], stride = 2, output channel = 256, padding = valid.

Each convolution layer contains the ReLU activation function and, subsequently, the max pooling. As to the activation function of CNN, the effect of ReLU exceeds the sigmoid in deeper networks, and it successfully solves the gradient dispersion problem of sigmoid encountered in deeper networks. Similarly, dropout layers have been imposed on the three fully connected layers (FC) (except the FC8) to prevent overfitting. This research selects three convolution layers to replace FC layers, so as to increase the computational efficiency of GPU. Finally, it outputs the ore-bearing probability and non-ore probability of each sample area through softmax layer (Fig. 4a).

3.3 Localization and probability determination

In this work we endow the predicted ore-bearing areas and the areas of known deposits with different colors after obtaining the coordinates from the prediction results and then output and save them in JPG format; meanwhile, we determine the ore-bearing probability through the softmax activation function given by

$$P_k = \frac{\exp(S_K)}{\sum_{k=0}^{K} \exp(S_K)},$$

where S_K refers to the input of softmax; we commonly consider S_K as the model score (also known as the non-normalized score) of the kth class; P_k refers to the probability.

The softmax activation function is only used for the neurons with more than one output so as to guarantee that the sum of all the outputted neurons equals 1.0. Thus, all the outputted probabilities are less than or equal to 1. And it is highly intuitive to compare the output values. If we consider P_k as the "probability" of containing ore or not containing ore, for example, then with an "ore-bearing" output of type "A" with probability equal to 0.8, we can access that the ore-bearing probability of the area delineated by the prediction model is 80%.

4 Experiments

4.1 Determination of optimal parameters

By comparing the grid division at different scales, as well as the training loss, training accuracy, verification accuracy, and recall (Table 2) of models trained by different factor layer combinations, we analyze the optimal sample division scale and the optimum input number of ore-controlling factor layers to make the model achieve its strongest robustness by controlling the scale and factor combination. The contrast experiments are divided into five groups, in which the first three belong to the contrast experiments with different grid sizes on the premise of, respectively, inputting five ore-controlling layers (Mn layer and other four ore-controlling factor layers).

By comparing the accuracy and loss curves for each group of the trained classification model (Fig. 5) we found that the metallogenic prediction performance with a grid size of 200 × 200 pixels (namely the actual distance of 10 km × 10 km) is optimal. Moreover, after the contrast experiment of different grid sizes with replacing the Mn layer with the distribution layers of 21 chemical elements, we found that the smaller the grid size, the more the number of selected prospecting factor layers, and the higher the verification accuracy of the prediction model. Moreover, the smaller the grid size, the larger the computational burden, and the lower the computational efficiency. In other words, we found that when we employ a grid size of 200 × 200 pixels, as well as the interpolated layers of 21 chemical elements, lithofacies paleogeographic map, formation and tectonic map, outcrop map of Datangpo Formation, and distribution map of water system in our experiments, the trained AlexNet convolutional network model reaches its best performance for the training loss, training accuracy, verification accuracy, and recall of 0.0000001, 100.00%, 86.21%, and 91.67%, respectively. We applied this model to predict Songtao–Huayuan, and the predicted areas with the highest ore-bearing probability according to the model possess high probability of containing undiscovered ore bodies.

4.2 Experimental results and discussion

In this paper we took the Songtao–Huayuan sedimentary Mn deposit as a case study and mine the coupling correlation between the spatial distribution of Mn, sedimentary facies, the outcrop of Datangpo Formation, faults, water system, and the occurrence space of Mn deposits, as well as the correlation among different ore-controlling factors by employing the AlexNet network of D-CNN model. We then segment (200 × 200 pixels) different factor layer data and label them from 0 to 911; subsequently, we train the 2-d mineral prediction and classification model with known positive and negative samples (Fig. 6). The training accuracy, verification accuracy, training loss, and recall of the trained model are 100.00%, 86.21%, 0.0000001, and 91.67%, respectively, indicating the reliability of the model. By applying this model to the 2-d metallogenic prediction in unknown areas, the ore-bearing areas eventually determined by the model score S_k and the ore-bearing probability computed by the softmax function can be obtained, as shown in Table 3.

The ore-bearing grid areas of Songtao–Huayuan predicted by AlexNet model in Experiments 1 to 5 are shown in Fig. 7. It can be seen from the overall distribution of the prediction results that if the grid size is too large, the prediction results would be insufficiently accurate, and if the grid size is too small, the prediction results would be too sparse and loose to achieve the purpose of delineating metallogenic prospects. That means, the more the layers participating in training, the more the predicted ore-bearing areas. Although the model contains the potential relevance

Table 2 Comparison on model indexes of multiple groups of experiments

Serial number	Number of layers	Size of grid/pixel	Positive sample/piece	Negative sample/piece	Step/time	Training loss	Training accuracy (%)	Verification accuracy (%)	Recall (%)
Experiment 1	5	400*400	32	62	990	0.000002	100.00	61.11	66.67
Experiment 2	5	200*200	63	89	1590	0.000000	100.00	75.86	58.33
Experiment 3	5	100*100	81	120	2190	0.000003	100.00	82.05	81.25
Experiment 4	25	400*400	32	62	990	0.000002	100.00	66.67	33.33
Experiment 5	25	200*200	63	89	1590	0.0000001	100.00	86.21	91.67

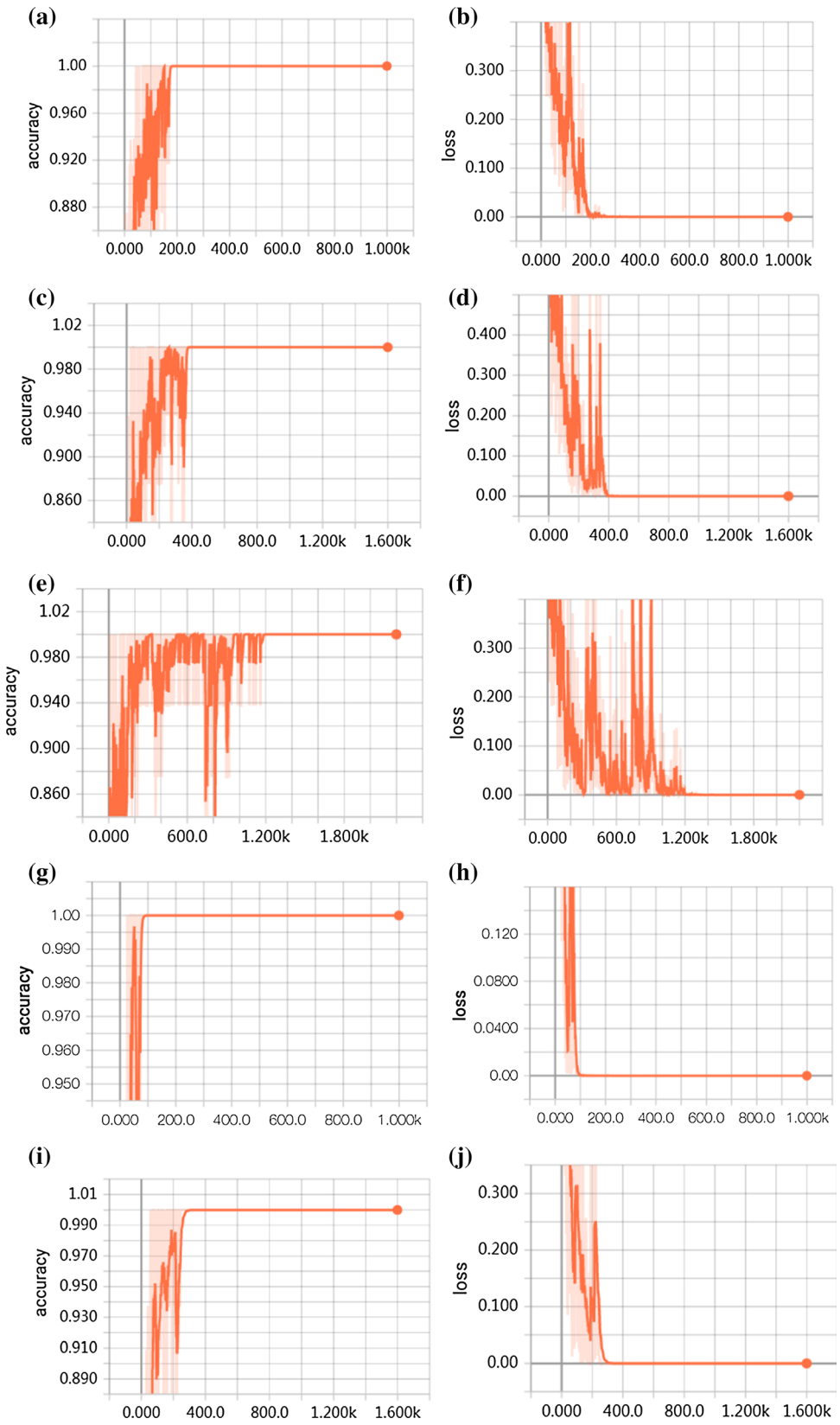


Fig. 5 Accuracy and loss curves of prediction model trained by samples with different grid sizes from Experiments 1 to 5; herein, figures **a**–**f** show the curves with 5 ore-controlling layers and figures **g**–**j** show the curves with 25 ore-controlling layers. The grid size of figures **a**, **b**, **g**, and **h** is 400 × 400 pixels; the grid size of figures **c**, **d**, **i**, and **j** is 200 × 200 pixels; and that of figures **e** and **f** is 100 × 100 pixels

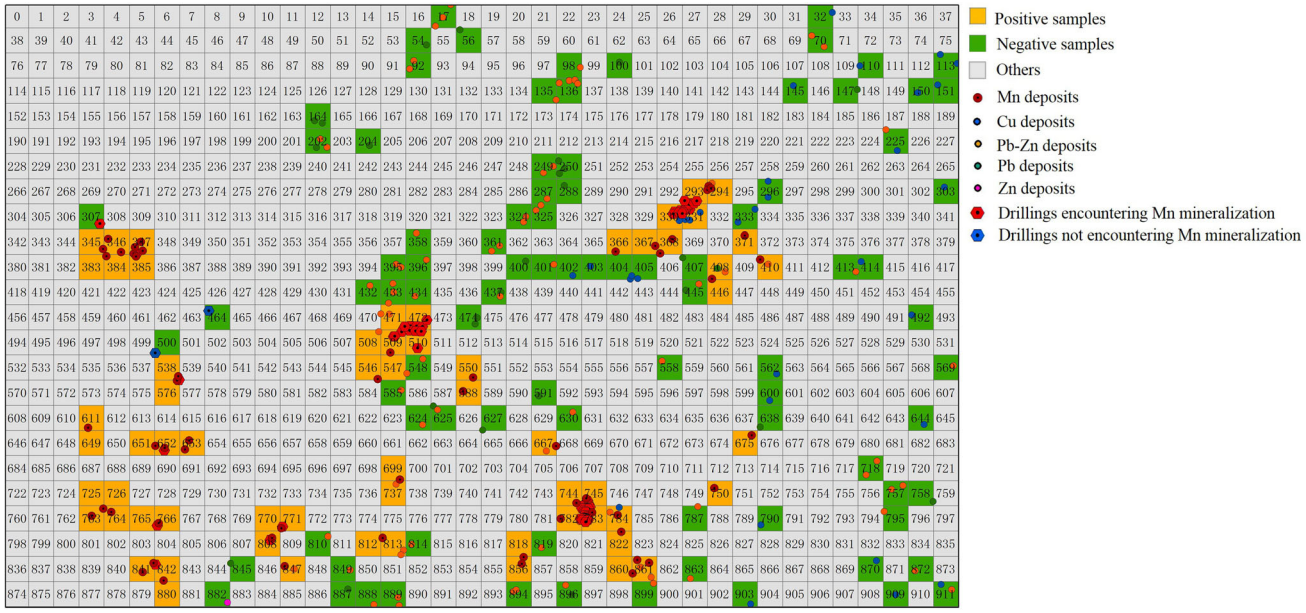


Fig. 6 Regional code and distribution of positive and negative samples

which can improve the reliability to some extent, it still faces uncertainty and interference from factors that have no or negative impact. Therefore, the prediction results should be further analyzed in combination with the geological background of the study area.

As to sedimentary minerals, sedimentary environment is the most direct ore-controlling factor; therefore, a certain sedimentary mineral of a certain geological period must occur in the specific horizon of a certain sedimentary system; moreover, the morphology, scale, and quality of some ore bodies are even related to sedimentary system. As a result, this research considers sedimentary facies as one of the ore-controlling factor layers to participate in training, and the finally superimposed results show that the predicted ore-bearing areas are mainly located in extensional rift basins or their surrounding regions. In particular, when the grid size is sufficiently small, both known and predicted deposits are mostly restrained in basins, while in paleouplift and Yangtze paleo-landmass, predicted deposit is absent. Moreover, the smaller the grid size, the higher the registration degree of prediction area and outcrop location of Datangpo Formation, and the more obvious the spatial distribution characteristics. This illustrates that the D-CNN model applied in this research has inherently embraced the sedimentary facies, spatial morphology, and distribution characteristics of strata outcrop, as well as their influence on Mn mineralization; in turn, these factors can serve as important basis for prospecting prediction. Consequently, the reliability of the model has been further verified. It can be inferred from the comparison on model evaluation indexes of multiple groups of experiments shown in Table 2 that the prediction effect of the model is best when the grid size is 200 × 200 pixels (Experiment 5) and the input number of factors is 25. Therefore, the analysis of prediction results is accomplished by mainly combining with the results of Experiment 5 and superimposing with various ore-controlling factor layers.

Since localization, quantification, and probability determination are the three problems that should be solved during the quantitative prediction of mineral resource, we decide to impose normalization on the model scores for softmax function and output the ore-bearing and non-ore probabilities of all coded areas (Fig. 8). The grid areas are then divided into five levels according the magnitude of the predicted ore-bearing probability (0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0), and the red ones represent the grid areas with a predicted ore-bearing probability in the range 0.8–1.0. Excluding the areas containing known ore bodies, the remaining ones are the predicted ore-bearing areas and possess a great probability of containing undiscovered deposits. Most of the areas predicted by the model are located near the known Mn deposits, which are in the known rift basins or their vicinity. Therefore, as the figure shows, the red areas with high probability cover all the known Mn deposits. During the verification process of the model, 20% of known samples served as validation set; as a result, the verification accuracy and recall of the model can reach up to 86.21% and 91.67%, respectively. The orange grid areas with probabilities in the range 0.6–0.8 represent potential ore-bearing areas. The figure shows that the metallogenic prediction areas with probabilities in the range 0.8–1.0 are distributed along the strike of faults characterized by NE45° distribution, and the predicted ore-bearing areas are not only located in rift basins, but also in

Table 3 Statistical table for the predicted ore-bearing probability of Songtao–Huayuan (in which the ore-bearing probability is more than zero)

No.	Ore	Non	No.	Ore	Non	No.	Ore	Non	No.	Ore	Non
28	0.01	0.99	366	1.00	0.00	480	0.07	0.93	586	0.01	0.99
30	1.00	0.00	367	1.00	0.00	481	1.00	0.00	588	1.00	0.00
31	0.01	0.99	368	1.00	0.00	482	0.99	0.01	596	0.02	0.98
32	1.00	0.00	369	1.00	0.00	484	1.00	0.00	597	1.00	0.00
66	1.00	0.00	370	1.00	0.00	485	0.72	0.28	598	0.12	0.88
67	1.00	0.00	371	1.00	0.00	499	0.17	0.83	602	0.99	0.01
68	1.00	0.00	372	1.00	0.00	501	0.92	0.08	603	0.99	0.01
101	0.19	0.81	373	1.00	0.00	502	1.00	0.00	610	1.00	0.00
104	1.00	0.00	374	1.00	0.00	506	1.00	0.00	611	1.00	0.00
105	0.01	0.99	383	1.00	0.00	508	1.00	0.00	612	0.96	0.04
138	0.04	0.96	384	1.00	0.00	509	1.00	0.00	613	1.00	0.00
141	1.00	0.00	385	1.00	0.00	510	1.00	0.00	614	0.62	0.38
142	1.00	0.00	386	1.00	0.00	512	1.00	0.00	615	1.00	0.00
218	1.00	0.00	393	0.99	0.01	516	0.01	0.99	617	0.04	0.96
251	0.44	0.56	397	1.00	0.00	517	1.00	0.00	623	0.98	0.02
255	0.34	0.66	398	1.00	0.00	518	0.02	0.98	634	0.03	0.97
256	1.00	0.00	406	1.00	0.00	519	1.00	0.00	635	0.01	0.99
257	1.00	0.00	408	1.00	0.00	520	0.03	0.97	641	1.00	0.00
292	1.00	0.00	409	1.00	0.00	521	0.31	0.69	642	1.00	0.00
293	1.00	0.00	410	1.00	0.00	522	0.99	0.01	649	1.00	0.00
294	1.00	0.00	411	0.99	0.01	524	0.39	0.61	650	1.00	0.00
302	1.00	0.00	425	1.00	0.00	537	1.00	0.00	651	1.00	0.00
304	0.87	0.13	437	0.56	0.44	538	1.00	0.00	652	1.00	0.00
305	1.00	0.00	439	1.00	0.00	539	1.00	0.00	653	1.00	0.00
308	1.00	0.00	440	1.00	0.00	540	1.00	0.00	654	1.00	0.00
310	0.27	0.73	441	0.02	0.98	545	0.99	0.01	661	1.00	0.00
311	1.00	0.00	442	0.36	0.64	546	1.00	0.00	662	1.00	0.00
326	1.00	0.00	443	1.00	0.00	547	1.00	0.00	663	1.00	0.00
328	0.28	0.72	444	0.97	0.03	549	1.00	0.00	664	1.00	0.00
330	1.00	0.00	446	1.00	0.00	550	1.00	0.00	666	0.92	0.08
331	1.00	0.00	447	1.00	0.00	551	1.00	0.00	667	1.00	0.00
332	1.00	0.00	451	0.01	0.99	556	0.01	0.99	668	1.00	0.00
334	1.00	0.00	456	0.03	0.97	559	1.00	0.00	673	0.04	0.96
335	1.00	0.00	457	1.00	0.00	560	0.03	0.97	674	0.97	0.03
342	1.00	0.00	458	0.04	0.96	572	0.08	0.92	675	1.00	0.00
345	1.00	0.00	460	0.26	0.74	573	0.36	0.64	677	0.13	0.87
346	1.00	0.00	462	1.00	0.00	575	1.00	0.00	679	1.00	0.00
347	1.00	0.00	471	1.00	0.00	576	1.00	0.00	680	1.00	0.00
350	0.02	0.98	472	1.00	0.00	577	1.00	0.00	684	1.00	0.00
356	1.00	0.00	477	0.01	0.99	582	0.38	0.62	685	1.00	0.00
No.	Ore	Non	No.	Ore	Non	No.	Ore	Non	No.	Ore	Non
686	1.00	0.00	722	1.00	0.00	774	1.00	0.00	843	1.00	0.00
687	0.78	0.22	723	1.00	0.00	775	1.00	0.00	846	1.00	0.00
688	1.00	0.00	724	1.00	0.00	777	0.99	0.01	847	1.00	0.00
689	1.00	0.00	725	1.00	0.00	780	1.00	0.00	852	1.00	0.00
690	1.00	0.00	726	1.00	0.00	782	1.00	0.00	853	1.00	0.00
697	1.00	0.00	727	1.00	0.00	783	1.00	0.00	855	0.99	0.01
700	1.00	0.00	728	1.00	0.00	784	1.00	0.00	856	1.00	0.00

Table 3 (continued)

No.	Ore	Non	No.	Ore	Non	No.	Ore	Non	No.	Ore	Non
701	1.00	0.00	729	1.00	0.00	785	0.98	0.02	857	1.00	0.00
702	1.00	0.00	730	1.00	0.00	787	1.00	0.00	859	1.00	0.00
704	1.00	0.00	732	0.99	0.01	788	1.00	0.00	860	1.00	0.00
705	1.00	0.00	734	1.00	0.00	791	0.43	0.57	861	1.00	0.00
706	0.73	0.27	736	0.99	0.01	792	0.99	0.01	862	1.00	0.00
707	0.03	0.97	737	1.00	0.00	793	0.95	0.05	864	0.02	0.98
708	0.80	0.20	738	0.03	0.97	794	1.00	0.00	865	0.16	0.84
710	1.00	0.00	741	1.00	0.00	799	1.00	0.00	866	1.00	0.00
711	0.71	0.29	742	0.04	0.96	800	1.00	0.00	871	0.93	0.07
712	1.00	0.00	743	0.99	0.01	801	1.00	0.00	873	1.00	0.00
713	0.02	0.98	744	1.00	0.00	802	1.00	0.00	874	1.00	0.00
715	1.00	0.00	745	1.00	0.00	803	1.00	0.00	875	1.00	0.00
716	1.00	0.00	746	1.00	0.00	804	1.00	0.00	876	1.00	0.00
722	1.00	0.00	747	1.00	0.00	805	1.00	0.00	877	1.00	0.00
723	1.00	0.00	749	1.00	0.00	807	1.00	0.00	878	1.00	0.00
724	1.00	0.00	750	1.00	0.00	808	1.00	0.00	879	1.00	0.00
725	1.00	0.00	751	1.00	0.00	809	1.00	0.00	880	1.00	0.00
726	1.00	0.00	752	1.00	0.00	811	0.99	0.01	881	0.01	0.99
727	1.00	0.00	753	1.00	0.00	812	1.00	0.00	884	1.00	0.00
728	1.00	0.00	754	0.98	0.02	813	1.00	0.00	897	1.00	0.00
729	1.00	0.00	755	1.00	0.00	815	1.00	0.00	898	1.00	0.00
730	1.00	0.00	760	1.00	0.00	816	1.00	0.00	900	0.01	0.99
732	0.99	0.01	761	1.00	0.00	817	1.00	0.00	902	0.09	0.91
734	1.00	0.00	762	1.00	0.00	818	1.00	0.00	906	0.55	0.45
736	0.99	0.01	763	1.00	0.00	821	1.00	0.00	908	0.01	0.99
737	1.00	0.00	764	1.00	0.00	822	1.00	0.00			
738	0.03	0.97	765	1.00	0.00	823	1.00	0.00			
741	1.00	0.00	766	1.00	0.00	836	0.62	0.38			
742	0.04	0.96	767	1.00	0.00	838	1.00	0.00			
743	0.99	0.01	769	1.00	0.00	839	1.00	0.00			
744	1.00	0.00	770	1.00	0.00	840	1.00	0.00			
745	1.00	0.00	771	1.00	0.00	841	1.00	0.00			
746	1.00	0.00	772	1.00	0.00	842	1.00	0.00			

the outcrop of Datangpo Formation, and this is in agreement with the metallogenic type of Songtao–Huayuan Datangpo-type sedimentary Mn deposit.

It can be seen from the overlap between the prediction results of AlexNet network model and the distribution of the 21 elements that the Mn, P, and Y anomalies correspond well with deposits (Fig. 9); the P deposits in the study area are characterized by their bedded occurrence and phosphorous host rocks of Doushantuo Formation; thus, they belong to sedimentary phosphorite deposits and possess close relationship with the occurrence horizon position of Mn deposits; additionally, the Y anomalies could reflect the spatial position of occurrence horizon (Nh_1d) for Mn deposits. Therefore, the comprehensive anomalies of Mn–P–Y are of great significance to the prospecting in the study area. It can be observed from Fig. 9 that, to some extent, the predicted ore-bearing areas are influenced by high content of Mn. Nevertheless, the Mn content in several parts of the predicted areas is relatively low. It indicates that the Mn is not the only principal determinant factor and the final prediction results are co-determined by the spatial distribution characteristics of various factor layers and the relevance among these factors.

Additionally, the spatial distribution of Y, Ba, and Hg is also related to the occurrence horizon (Datangpo Formation) of predicted metallogenic areas, and the anomalies of

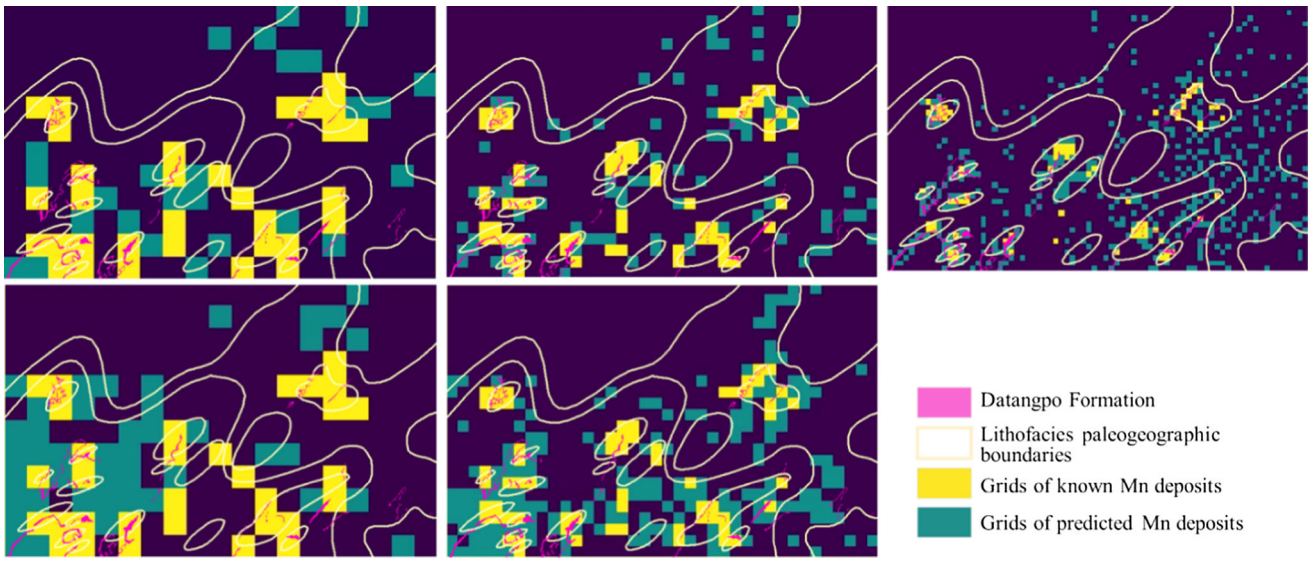


Fig. 7 Prediction results of Mn deposits in Songtao–Huayuan

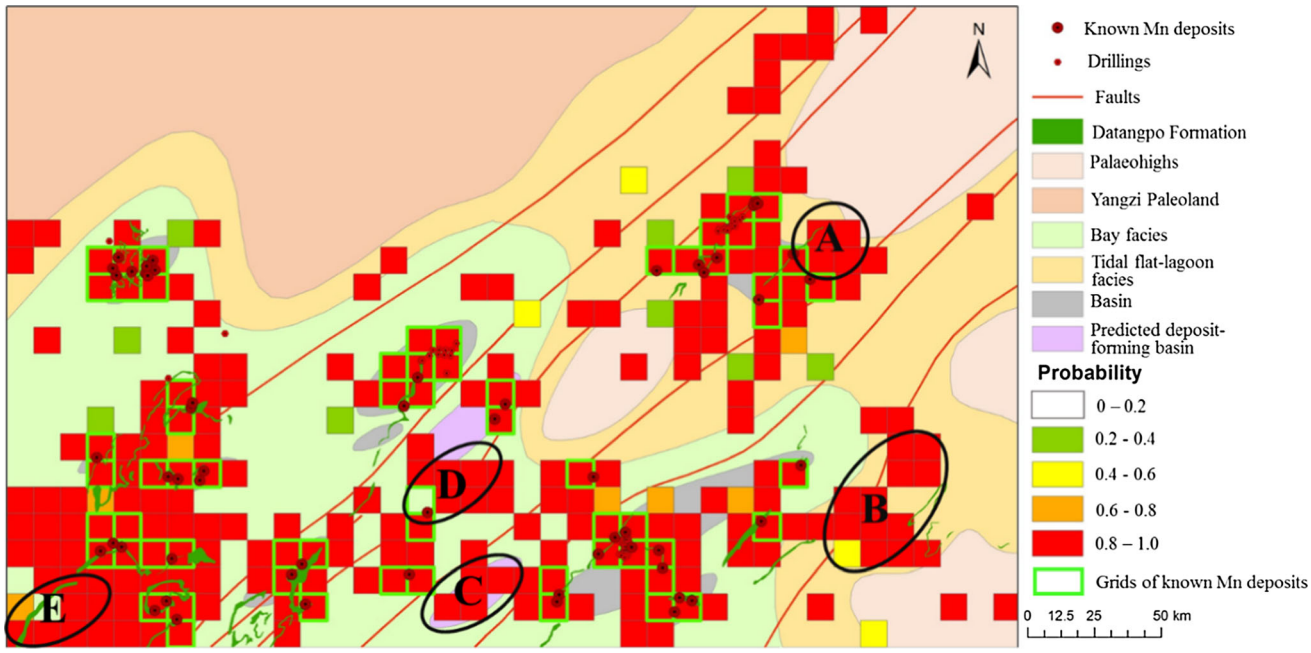


Fig. 8 Distribution of predicted ore-bearing probability and delineation of metallogenic prospects in Songtao–Huayuan

Mn–Fe_2O_3–Ti have better register relationship with Mn-bearing areas. Thus, the spatial distribution rule of siderophile elements is closely related to Mn deposits. Moreover, F, B, Ba, U, and Y are all concentrated in the surrounding of Neoproterozoic group complex and controlled by Cambrian, Sinian, and Nanhua strata, respectively. Thereinto, the enrichment of B is related to its high content in sedimentary rocks, especially in argillaceous sedimentary rocks, and the content of B in metamorphic rocks is generally much less than that in sedimentary rocks. The contents of Ba and U are closely related to black shale, and that of Hg in carbonaceous–argillaceous shale is frequently slightly higher, while Y can clearly reflect the spatial position of occurrence horizon. Therefore, the successive enrichment of F–B–Ba (U)–Hg–Y may be able to provide a basis for the prediction of regional Mn deposits or deep prediction of the property.

In combination with the geological background of Songtao–Huayuan, five metallogenic prospects numbered from A to E as shown in Fig. 8 can be eventually delineated. Prospects A, C, D, and E all are located near the known rift basins, while prospects A, B, and E all develop

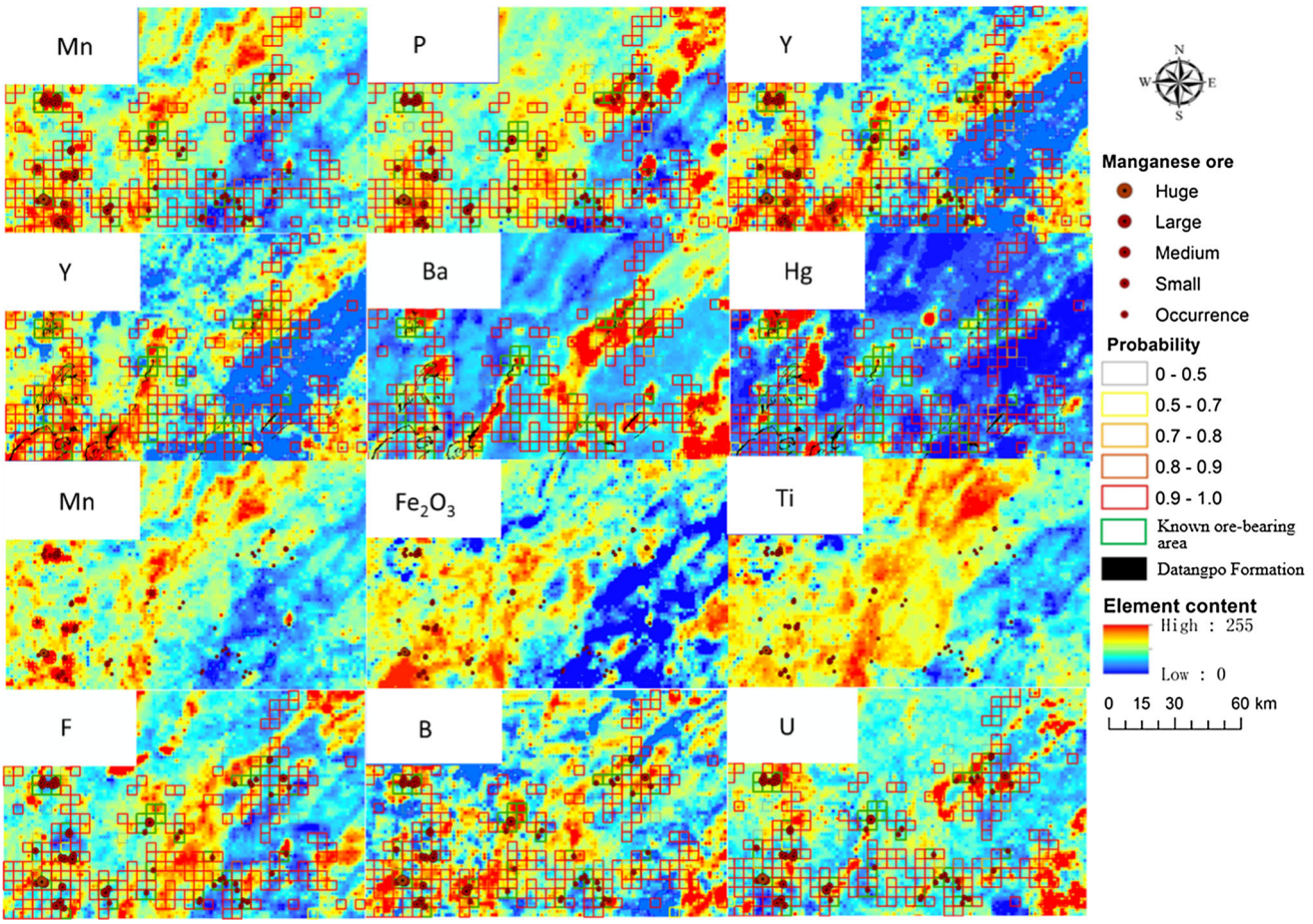


Fig. 9 Superposed map of prediction results of CNN model and the content distribution of 21 elements

the outcrop of Datangpo Formation. Additionally, prospects C and D are located in and near the predicted manganese-forming basin, respectively. It can be concluded from our comprehensive analysis that these five areas have the highest probability of containing undiscovered sedimentary Mn deposits.

5 Conclusions

Previous mineral prediction research focused only on the quantitative extraction of anomalies and commonly neglected the importance of spatial distribution characteristics (e.g., coupling correlation between the distribution characteristics of geochemical elements, sedimentary facies, tectonics, and strata outcrop and the spatial distribution of deposits). Moreover, the relevance among different metallogenic conditions is rarely taken into consideration when carrying out traditional evidence weight methods. It is not rigorous to merely analyze what are deemed as important ore-controlling factors and neglect the ones deemed irrelevant, for potential metallogenic conditions exhibiting "butterfly effects" are likely to be neglected. Therefore, we need a set of quantitative prediction methods to simultaneously make geological background and big data "speak." Accordingly, with the help of the prediction approach of mining big data, and the AlexNet algorithm of D-CNN for depicting complicated and nonlinear geoscience spatial modes, this paper proposes a kind of 2-d prospecting prediction method not only capable of extracting spatial distribution characteristics, but also capable of mining out the relevance of various ore-controlling factor layers under different ore-controlling factors. Using Songtao–Huayuan as the study area, several conclusions can be drawn:

1. With the help of the AlexNet networks, a classification model based on D-CNN is obtained by training the relationship between the distribution of known Mn deposits in Songtao–Huayuan and 21 chemical elements including Mn, as well as sedimentary facies, outcrop of Datangpo Formation, faults, and water system. The model can effectively mine the relationship between the surficial distribution characteristics of ore-controlling factors and distribution of deposits.

2. Through multiple groups of contrast experiments, this paper discusses the optimal sample division scale and the optimum input number of ore-controlling factor layers. The performance of the trained model for the whole study area is best when the grid size is 10×10 km^2 and the layer number of ore-controlling factors is 25. With these optimal parameters, the verification accuracy, recall, and training loss of the AlexNet model are 86.21%, 91.67%, and 0.0000001, respectively. Thus, the accuracy of the model could be considered as reliable.
3. By applying this model to the 2-d metallogenic prediction in unknown areas, the predicted ore-bearing probability is divided into five levels, namely 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0. In combination with the geological background of Songtao–Huayuan, a total of five metallogenic prospects numbered from A to E are delineated. Our analysis shows that these prospects have great probability of containing undiscovered ore bodies; thus, boreholes are recommended to be implemented in selected sites for verification.

However, the method still faces a few problems to be implemented in actual applications of prospecting prediction:

1. The scale and grid size determination need to be further studied. If the study area is too large, the samples would not be accurate enough; reversely, if the study area is too small, the data scope of samples would be circumscribed, and the requirement for the prediction accuracy would be extremely high. Whether the surficial characteristics are capable of reflecting the underground metallogenic position is still unknown. Additionally, grid size can indirectly influence the effect of metallogenic predictions. If the grid size is too small, although the limited samples can be maximally utilized, the computational burden would greatly increase. If the grid size is too large, effective use of the existing sample points becomes challenging and thus the prediction accuracy would decrease. Therefore, an appropriate scale should be selected to guarantee the sample capacity and the increase in the prediction accuracy as high as possible.
2. As to the problem of selecting samples, this research has maximally collected all the known manganiferous or non-manganiferous samples. The known Mn deposits served as positive samples, while negative boreholes and other kinds of deposits served as negative samples, but the sample capacity is far from sufficient for deep learning. Currently, the commonly used methods of expanding original images by geometric transformation of the images [41] or by radial transformation in polar coordinates space [42] are not suitable for geological samples. Therefore, an appropriate method is greatly needed to solve the problem of sample capacity so as to improve the effect of the trained models.
3. Under the era of big data, we should have a model that is broadly applicable to a large number of factor layers. If the sample capacity is large enough, all the layers corresponding to the ore-controlling factors produced after their match with prospecting concept model library can be used to train the AlexNet model. Consequently, more potential ore-controlling factors can be mined to really make big data "speak" and to gradually realize "intelligent" prospecting accordingly.

Acknowledgements This research was financially supported by the Chinese MOST project "Methods and Models for Quantitative Prediction of Deep Metallogenic Geological Anomalies" (2017YFC0601502).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Zhao PD (2015) Digital mineral exploration and quantitative evaluation in the big data age. Geol Bull China 34(7):1255–1259
2. Xiao KY, Sun L, Li N et al (2015) Mineral resources assessment under the thought of big data. Geol Bull China 34(7):1266–1272
3. Chen JP, Li J, Cui N et al (2015) The construction and application of geological cloud under the big data background. Geol Bull China 34(7):1260–1265
4. Yu PP, Chen JP, Chai FS et al (2015) Research on model-driven quantitative prediction and evaluation of mineral resources based on geological big data concept. Geol Bull China 34(7):1333–1343
5. Zheng X, Li JC, Wang X et al (2015) Construction of the national geological information service system in the age of big data. Geol Bull China 34(7):1316–1322
6. Li S, Chen JP, Xiang J (2018) Prospecting information extraction by text mining based on convolutional neural networks—a case study of the Lala copper deposit, China. IEEE Access 6:52286–52297
7. Yong LW, Jian ZJ, Chen JP et al (2017) Construction and prediction of prospecting model based on big data intelligence. China Min Mag 26(9):79–84
8. Zuo RG, Xiong YH (2018) Big data analytics of identifying geochemical anomalies supported by machine learning methods. Nat Resour Res 27(1):5–13
9. Twarakavi NKC, Misra D, Bandopadhyay S (2006) Prediction of arsenic in bedrock derived stream sediments at a gold mine site under conditions of sparse data. Nat Resour Res 15(1):15–26
10. Chen M, Mao S, Liu Y (2014) Big data: a survey. Mob Netw Appl 19(2):171–209
11. O'Brien JJ, Spry PG, Nettleton D et al (2015) Using random forests to distinguish gahnite compositions as an exploration

guide to Broken Hill-type Pb–Zn–Ag deposits in the Broken Hill domain, Australia. J Geochem Explor 149:74–86
12. Gonbadi AM, Tabatabaei SH, Carranza EJM (2015) Supervised geochemical anomaly detection by pattern recognition. J Geochem Explor 157:81–91
13. Kirkwood C, Cave M, Beamish D et al (2016) A machine learning approach to geochemical mapping. J Geochem Explor 167:49–61
14. Zhao JN, Chen SY, Zuo RG (2016) Identifying geochemical anomalies associated with Au–Cu mineralization using multifractal and artificial neural network models in the Ningqiang district, Shaanxi, China. J Geochem Explor 164:54–64
15. Xiong YH, Zuo RG (2016) Recognition of geochemical anomalies using a deep autoencoder network. Comput Geosci 86:75–82
16. Chen YL, Wu W (2017) Application of one-class support vector machine to quickly identify multivariate anomalies from geochemical exploration data. Geochem Explor Environ Anal 17(3):231–238
17. Chen JP, Li J, Xie S et al (2017) China geological big data research status. J Geol 41(3):353–366
18. Chen SM (2012) Research of multiple geoscience information prospecting prediction in Xikuangshan antimony ore field. China University of Geosciences, Beijing
19. Chen JP (2008) The research and application of neural network pattern recognition technique for oil and gas geochemistry base on MATLAB. China University of Geosciences, Beijing
20. Yang H (2016) Research and application of the combination of deep learning and principal component analysis. Chengdu University of Technology, Chengdu
21. Albora AM, Ucan ON, Ozmen A et al (2001) Separation of Bouguer anomaly map using cellular neural network. J Appl Geophys 46(2):129–142
22. Liu Z, Liu MC, Wei W et al (2010) Gravity anomaly separation based on cellular neural network. J China Univ Pet (Ed Nat Sci) 34(4):57–61
23. Li C, Jiang YL, Hu MK et al (2015) Study and application of gravity anomaly separation by cellular neural networks. Comput Tech Geophys Geochem Explor 37(1):16–21
24. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507
25. Liu YP, Zhu LX, Zhou YZ (2018) Application of convolutional neural network in prospecting prediction of ore deposits: taking the Zhaojikou Pb–Zn ore deposit in Anhui Province as a case. Acta Petrol Sin 34(11):3217–3224
26. Dang Y, Zhang JX, Deng KZ et al (2017) Study on the evaluation of land cover classification using remote sensing images based on AlexNet. J Geo-Inf Sci 19(11):1530–1537
27. Zhou Q, Du YS, Yuan LJ et al (2017) Exploration models of ancient natural gas seep sedimentary-type manganese ore deposit: a case study of the Nanhua Period "Datangpo" type manganese ore in the conjunction area of Guizhou, Hunan and Chongqing. Acta Geol Sin 91(10):2285–2298
28. Liu XF, Wang QS, Gao XJ (1989) Manganese ore geology in Guizhou. Guizhou People's Publishing House, Guiyang
29. Wang YG, Wang LX, Zhu SC et al (1985) Sedimentary environment and manganese formation of the Datangpo formation in eastern Guizhou. Guizhou People's Publishing House, Guiyang
30. He MH (1993) Petrographical facies and paleogeography and manganese: forming conditions of Datangpo epoch of early Sinian period in Songtao and adjacent areas. Guizhou Geol 10(1):62–67
31. He MH (2001) Sedimentary facies and palaeogeography during the early Sinian and potential of the rhodochrosite deposits in northeastern Guizhou and its adjacent areas. Sediment Geol Tethyan Geol 21(3):39–47
32. Xiang WQ, Xiao YK (2013) Geological characteristics and metallogenic regularity for Nanhuan Datangpo manganese deposits from Tongren–Songtao region. J Southwest Univ Sci Technol 28(4):31–38
33. Mou J, Wang AH, Huang DG (2014) Sedimentary micro-facies features and prospecting prediction of manganese-bearing rock series in Songtao–Yinjiang area of Guizhou. Guizhou Geol 31(2):99–104
34. Kuang WL, Li XY, Yang SX (2014) The mineralization geological characteristics and genesis of Minle type manganese deposits in northwestern of Hunan Province. Chin J Geol 49(1):305–323
35. Wang YG (1990) Old hot brine manganese deposits in a shallow-sea rift basin: an example from the Sinian manganese deposits in the Wuling mountain area. Sediment Geol Tethyan Geol 1:38–45
36. Zhao DX (1990) Intraclastic structures and gravity flow sedimentation of rhodochrosite ore in Sinian Datangpo formation. Sci Geol Sin 2:149–157
37. Liu BJ, Yu GM, Chen CS (1993) Ophiolite submarine fan of the Shigaze Group in the Yarlung Zangbo suture zone and its plate tectonic significance. Sediment Geol Tethyan Geol 13(2):16–27
38. Zhou Q, Du YS (2012) Ancient natural gas leakage and manganese mineralization: a case study of the Nanhua Datangpo manganese deposits in eastern Guizhou. Geological Publishing House, Beijing, pp 2–10
39. Zhou Q, Du YS, Qin Y (2013) Ancient natural gas seepage sedimentary-type manganese metallogenic system and ore-forming model: a case study of "Datangpo type" manganese deposits formed in rift basin of Nanhua Period along Guizhou–Hunan–Chongqing border area. Miner Depos 32(3):457–466
40. Du YS, Zhou Q, Yu WC et al (2015) Linking the Cryo-genian manganese matallogenic process in the south-east margin of Yangtze block to break-up of Rodinia supercontinent and Sturtian glaciation. Geol Sci Technol Inf 34(6):1–7
41. Hobert JP (2011) The data augmentation algorithm: theory and methodology. In: Brooks S, Gelman A, Jones G et al (eds) Handbook of Markov chain Monte Carlo. CRC Press, London, pp 253–293
42. Salehinejad H, Valaee S, Dowdell T et al (2018) Image augmentation using radial transform for training deep neural networks. In: IEEE international conference on acoustics, speech and signal processing. IEEE Press, New York, pp 3016–3020