

Article

Prediction of Prospecting Target Based on Selective Transfer Network

Yongjie Huang ^{1,†}, Quan Feng ^{2,3,†}, Wanting Zhang ¹, Li Zhang ⁴ and Le Gao ^{1,*}¹ Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen 529000, China² State Key Laboratory of Public Big Data, Guizhou University, Guiyang 550025, China³ College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China⁴ College of Computer Science and Technology, Nanjing Forestry University, Nanjing 210095, China

* Correspondence: le.gao@nscg-gz.cn

† These authors contributed equally to this work.

Abstract: In recent years, with the integration and development of artificial intelligence technology and geology, traditional geological prospecting has begun to change to intelligent prospecting. Intelligent prospecting mainly uses machine learning technology to predict the prospecting target area by mining the correlation between geological variables and metallogenic characteristics, which usually requires a large amount of data for training. However, there are some problems in the actual research, such as fewer geological sample data and irregular mining features, which affect the accuracy and reliability of intelligent prospecting prediction. Taking the Pangxidong study area in Guangdong Province as an example, this paper proposes a deep learning framework (SKT) for prospecting target prediction based on selective knowledge transfer and carries out intelligent prospecting target prediction research based on geochemical data in Pangxidong. The irregular features of different scales in the mining area are captured by dilation convolution, and the weight parameters of the source network are selectively transferred to different target networks for training, so as to increase the generalization performance of the model. A large number of experimental results show that this method has obvious advantages over other state-of-the-art methods in the prediction of prospecting target areas, and the prediction effect in the samples with mines is greatly improved, which can effectively alleviate the problems of a small number of geological samples and irregular features of mining areas in prospecting prediction.

Keywords: transfer learning; dilated convolution; deep learning; prospecting target prediction



Citation: Huang, Y.; Feng, Q.; Zhang, W.; Zhang, L.; Gao, L. Prediction of Prospecting Target Based on Selective Transfer Network. *Minerals* **2022**, *12*, 1112. <https://doi.org/10.3390/min12091112>

Academic Editor: Yosoon Choi

Received: 25 July 2022

Accepted: 27 August 2022

Published: 31 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Mineral resources are an essential strategic resource to ensure national stability and economic development. After years of geological exploration, surface minerals have gradually become scarce, resulting in a further increase in the difficulty of prospecting. The prospecting work started in deep surfaces and overburdened areas. In recent years, the development of machine learning technologies, especially deep learning, has achieved satisfactory success in many fields, e.g., in computer vision [1], natural language processing [2], synthetic aperture radar [3], and atmospheric prediction [4]. Meanwhile, it has also interested and been selected by geological researchers [5–8]. Therefore, the traditional geological prospecting method is gradually turning into an intelligent prospecting method based on machine learning.

An intelligent prospecting target refers to the use of machine learning technologies to mine the correlation between geological variables and metallogenic characteristics to predict the metallogenic target area [9,10]. Although quantitative prospecting target prediction based on machine learning technologies is in its infancy, some progress has been made.

These works can be divided into the following three categories according to the method of implementation:

- (1) Methods based on ensemble learning combine multiple supervised learning algorithms for prospecting target prediction. For example, the authors in [11] determined the hyperparameters of the random forest by simulating the natural evolution process, which were used to improve the accuracy of the model in predicting the prospecting target area. The authors in [12] used the isolation forest algorithm to predict outliers to determine the prospecting target area. The authors in [13] proposed the use of metric learning in the random forest to project the sample features into the feature space, separating the background and mining targets, so as to improve the prediction accuracy of the model.
- (2) Methods based on the support vector machine (SVM) divide mining samples and other samples through a hyperplane. For example, the authors in [14] separated the “mines” samples from “non-mines” samples through the optimal hyperplane and determined three prospecting target areas on both sides of the hyperplane. The authors in [15] used a genetic algorithm to optimize the hyperparameters of SVM to reduce its influence on prospecting target prediction.
- (3) Methods based on depth neural networks project the geoscience data into the same depth neural network space and extract effective features through multiple nonlinear transformations for prospecting target prediction. For example, the authors in [16] used three-layer convolution to extract the features of a Zn-element concentration distribution map to predict the prospecting target area. The authors in [17] used AlexNet to extract the features of multiple ore-forming factor maps to determine four prospecting target areas.

Although the aforementioned methods have achieved success, problems such as the small number of geological samples and the irregular features of mining areas in prospecting target prediction have not been properly solved. Fortunately, in recent years, in-depth research has been carried out on these problems. These works can also be divided into the following two categories according to the method of implementation:

- (1) Methods based on data augmentation increase the diversity of geological samples by cropping, changing the chromatic aberration and size, and distorting features. For example, the authors in [18] added random noises into geological data to predict a prospecting target by a deep convolutional network. The authors in [19] first oversampled the samples with mines and then used the random forest to determine the prospecting target areas. The authors in [20] proposed recombining pixel pairs of geological samples to assist in prospecting target prediction. These methods increase the number of geological samples but cannot cope with the irregular features of the mining areas.
- (2) Methods based on multiscale feature transformation acquire mining area features of different scales through irregular sampling for training [21,22]. For example, the authors in [23] proposed to extract irregular features of different scales using multi-group convolution or a pooling operation and fused them for prospecting target prediction. The authors in [24] used four convolution operations with different sizes to extract and fuse irregular features of geological data to improve the prediction accuracy. However, these methods do not consider the small number of samples in the prospecting target area.

To summarize, the motivation behind our method comes from two aspects: (1) Selective transfer of label information learned from large tasks to small-shot tasks can effectively improve its generalization performance; therefore, we try to selectively transfer knowledge from a well-trained large prospecting target prediction task to the target task to assist training. (2) The other is to expand the receptive field of the convolution kernel without increasing the number of parameters. In this way, we can utilize it to deal with the irregular features of the mining areas. Therefore, we combine these two motives and propose a novel

deep learning framework based on selective knowledge transfer (SKT) for prospecting target prediction to solve the problems existing in prospecting target prediction. As shown in Figure 1, we first introduce a large well-trained prospecting target prediction network as the source network and define multiple target networks with the same structure. Next, we use a soft mask to alleviate the differences in metallogenic effects that different elements may cause. Then, we use dilated convolution to capture features of different scales as inputs of the target networks. Based on this, we selectively transfer the weight parameters from the source network for the training of the different target networks. Finally, we perform top-down distillation on the large-scale and small-scale target network to mine the hidden knowledge between feature maps of different scales for small-scale target network learning. It is worth noting that the input size of $W \times H \times n$ is obtained by processing the geochemical data, where W denotes width, H denotes height, and n denotes the number of geochemical elements. Meanwhile, the output is generated by voting through all the target networks and contains two values representing the votes obtained for with and without mines, respectively. The whole SKT framework works in an end-to-end manner. Extensive experimental results show that our method is significantly competitive with state-of-the-art methods.

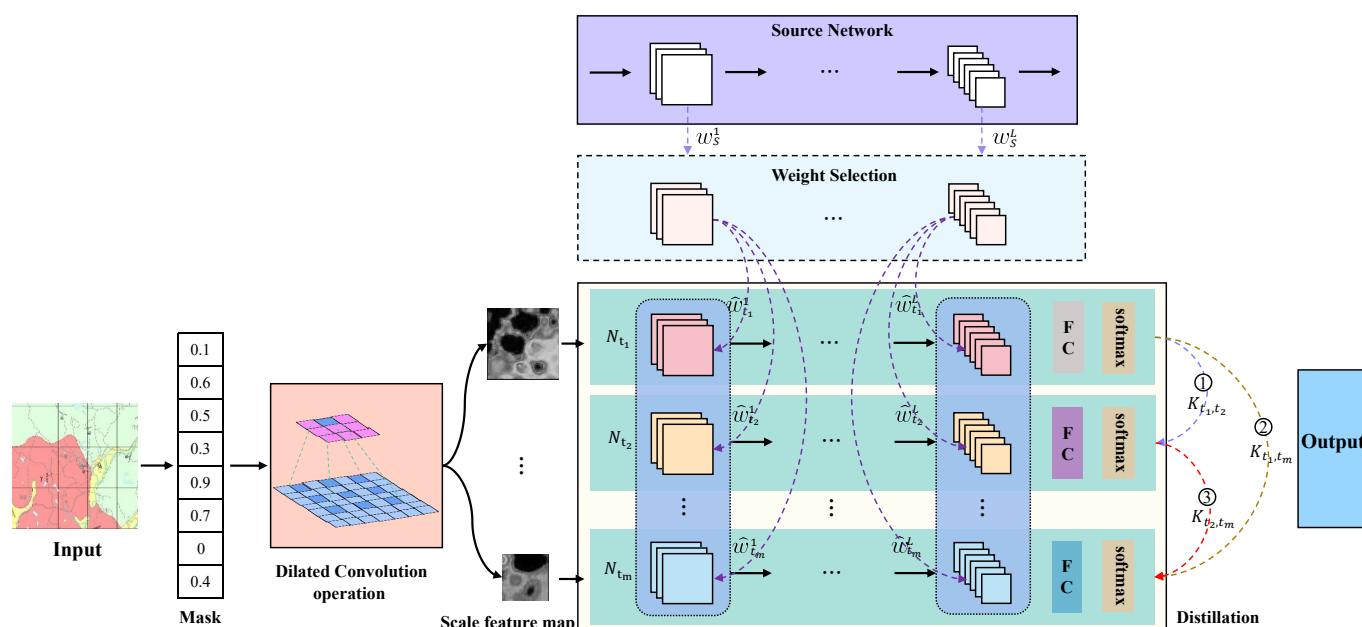


Figure 1. Schematic architecture of the SKT framework. w_s^l denotes the convolution kernels of the l -th convolution layer of the source network, N_{t_v} , $v = 1, 2, \dots, m$ denotes the target networks, $\hat{w}_{t_v}^l$ denotes the convolution kernels transferred to N_{t_v} from the source network, and K_{t_i, t_j} denotes the direction of knowledge distillation, which means that N_{t_i} guides N_{t_j} to learn. Knowledge distillation proceeds from the top down, e.g., knowledge distillation performs in the order of ①, ②, ③. Weight selection makes the source network weights sparse through the Hadamard product and then transfers to the target network.

The main contributions of this paper include the following:

1. A deep learning framework for prospecting target prediction is proposed, which provides a new way for prospecting target prediction.
2. A novel selective knowledge transfer mechanism is designed, which selectively transfers knowledge from the source network to target networks, which increases the performance of the target networks in prospecting target prediction during testing without adding additional computational cost.
3. For the first time, a soft mask strategy is proposed to maintain the consistency of related mineral elements. Its purpose is to utilize the metallogenic indicative significance of the

main mineral elements and associated mineral elements to complete the prospecting target prediction task.

The rest of this paper is organized as follows. Section 2 introduces the study area and data processing methods. Section 3 infers the structure of the SKT framework. Section 4 exhibits our experiment, including analysis of the experimental results and visualization. Section 5 summarizes the work of this paper.

2. Study Area and Data

2.1. Study Area

The study area is the Pangxidong study area in Guangdong Province, China. Figure 2 shows a simple geological map of the study area, where the red dots are the mining areas. It is a part of the Qinzhou–Hangzhou metallogenic system, and the metallogenic geological conditions are superior. Its tectonic background belongs to the eastern margin of the southern segment of the Qinzhou–Hangzhou metallogenic system. Since the late Paleozoic, it has been uplifted for a long time and is an important metallogenic area for precious metals and non-ferrous metals, with many kinds of minerals [25]. The minerals that have been explored or are being exploited include gold ore, silver ore, lead–zinc ore, tungsten and molybdenum ore, iron ore, etc. [26]. The types of deposits mainly include ductile shear zone silver–gold deposits, porphyry-type molybdenum–tungsten–copper polymetallic deposits, sedimentary reformation type, magmatic hydrothermal type, and contact metasomatic lead–zinc polymetallic deposits [27]. Minerals are mainly distributed in the northwest and southeast, along the Pangxitong fault zone and the Gucheng–Shachan fault zone, in a northeast direction.

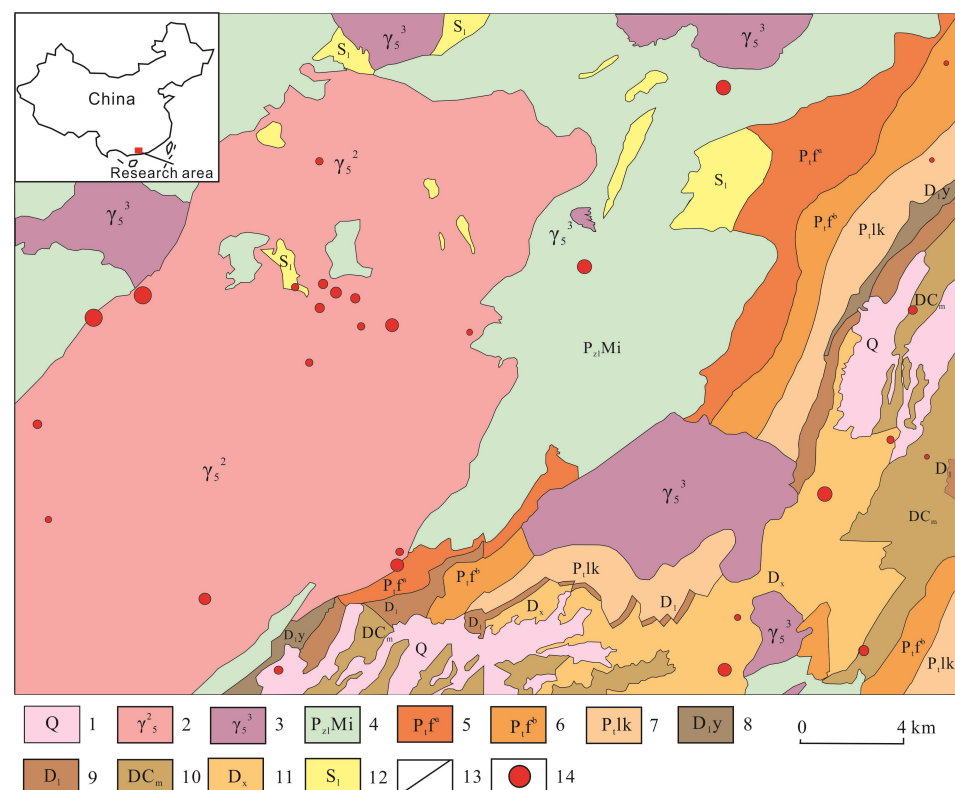


Figure 2. Simple geological map of the study area (1—Quaternary; 2—Early Yanshanian granite; 3—Late Yanshanian granite; 4—Upper Proterozoic migmatite; 5—Lower Member of Middle-Upper Proterozoic Fengdongkou Formation; 6—Upper Member of Middle-Upper Proterozoic Fengdongkou Formation; 7—Middle-Upper Proterozoic Lankeng Formation; 8—Devonian Yangxi Formation; 9—Devonian Laohutou Formation; 10—Devonian-Carboniferous Maozifeng Formation; 11—Devonian Xindu Formation; 12—Silurian Liantan Formation; 13—Faults; 14—Deposit).

2.2. Data Processing

The experimental data consist of geochemical data and were extracted from a 1:50,000 stream sediment survey. They were obtained from the Pangxidong study area in Guangdong Province. The sampling area of the stream sediment was 1694 km², and the average sampling density was 4.27 samples km². Sixteen chemical elements, including Au, B, Sn, Cu, Ag, Ba, Mn, Pb, Zn, As, Sb, Bi, Hg, Mo, W, and F, were analyzed from the stream sediment samples. Table 1 shows some of our original data, with X and Y as the coordinates of the sampling points and the rest as the values of the geochemical elemental content, totaling to 7237 points. The area under curve (AUC) is defined as the area under the receiver operating characteristic (ROC) curve enclosed by the coordinate axes [28]. Based on thesis [29], we leveraged SVM and statistical methods to screen out geochemical elements, which are indicative of mineralization. Specifically, we calculated the AUC of each element using the SVM algorithm, and then the standard deviation of the AUC was calculated by

$$S_{AUC} = \sqrt{\frac{AUC(1 - AUC) + (C_p - 1)(Q_1 - AUC^2) + (C_n - 1)(Q_2 - AUC^2)}{C_p \times C_n}} \quad (1)$$

where S_{AUC} denotes the standard deviation of the AUC; C_p and C_n denote the numbers of samples with and without mines, respectively; and $Q_1 = \frac{AUC}{2 - AUC}$ and $Q_2 = \frac{2AUC^2}{1 + AUC}$ are temporary variables. To determine whether the AUC was significantly different from 0.5, we defined the following random variable:

$$Z_{AUC} = \frac{AUC - 0.5}{S_{AUC}} \quad (2)$$

Table 1. Geochemical data of Pangxidong.

X	Y	Au	B	Sn	Cu	Ag	Ba	Mn	Pb	Zn	As	Sb	Bi	Hg	Mo	W	F
422.24	2418.80	0.9	3	8.7	4	0.025	33	147	27	26	1.17	0.31	0.23	0.04	2.67	0.79	212
421.37	2418.80	0.54	4	2.56	7	0.078	88	209	12	23	0.9	0.29	0.13	0.04	0.82	1.16	204
419.76	2418.25	0.81	3	1.52	5	0.043	1111	423	42	14	0.51	0.35	0.06	0.07	0.59	0.38	101
420.12	2418.40	0.37	2	1.65	6	0.046	941	498	38	17	0.53	0.31	0.1	0.02	0.57	0.33	111
420.55	2418.60	1.09	4	1.53	8	0.033	427	338	37	29	0.74	0.28	0.09	0.07	1.68	0.73	186
433.81	2397.92	2.31	121	2.2	4	0.075	365	239	16	18	4.31	0.96	0.43	0.066	0.77	3.01	186
424.17	2415.02	0.43	5	2.18	4	0.069	42	250	13	14	1.21	0.33	0.4	0.031	1.04	1.53	201
423.74	2415.31	0.51	5	4.85	7	0.004	30	242	47	31	0.5	0.26	0.32	0.016	1.02	3.07	210
425.14	2414.87	0.46	6	2.08	7	0.061	28	298	15	25	1.49	0.35	0.24	0.075	1.75	1.3	217
425.14	2415.15	0.47	6	1.95	7	0.055	54	420	9	18	1.07	0.37	0.14	0.042	1.08	0.85	108
424.86	2414.76	0.5	5	1.46	4	0.036	21	355	6	11	1.1	0.33	0.17	0.022	0.98	1.14	130
424.47	2414.47	0.59	6	2.6	4	0.038	29	170	6	21	1.01	0.33	0.2	0.039	1.32	2.02	192
424.82	2414.37	0.43	11	2.26	2	0.027	22	210	6	19	1.19	0.31	0.2	0.03	1.5	1.76	177
425.22	2414.46	1.05	45	4.2	3	0.065	39	125	6	25	1.94	0.39	0.54	0.046	2.17	3.04	396
424.41	2414.11	0.4	6	1.84	3	0.054	16	231	6	10	0.77	0.28	0.11	0.016	0.83	0.91	93
424.72	2413.83	0.9	7	3.87	9	0.094	135	231	48	45	2.15	0.41	0.73	0.069	1.82	2.54	327
424.35	2413.78	0.68	6	2.68	3	0.059	26	130	5	25	2.32	0.34	0.37	0.045	1.09	1.69	201
431.88	2411.24	0.81	4	3.58	15	0.053	140	143	83	34	2.76	0.36	2	0.052	2.05	8.94	241
432.90	2411.89	0.39	3	3.4	10	0.077	121	133	93	33	1.73	0.34	4.24	0.061	2.33	6.41	230
433.60	2410.63	0.42	5	2.62	1	0.042	81	152	12	21	1.55	0.31	0.3	0.054	0.77	0.94	135
433.91	2411.37	0.8	4	3.49	2	0.025	120	88	31	33	3.17	0.33	0.72	0.062	0.85	1.48	231
434.07	2410.91	0.42	6	2.99	4	0.057	109	117	12	22	2.38	0.32	0.27	0.048	0.63	2.19	210
434.73	2410.27	0.37	4	3.11	9	0.045	171	132	12	21	1.93	0.31	0.12	0.042	0.69	1.98	180
434.09	2409.69	0.36	5	2.92	6	0.044	135	123	5	21	1.76	0.29	0.09	0.046	0.71	0.6	156
432.45	2414.37	1.86	46	1.82	11	0.059	96	278	13	22	11.58	0.42	0.35	0.017	0.78	3.05	150
432.29	2414.59	3.76	65	6.04	4	0.036	41	164	67	20	9.37	0.79	0.67	0.047	2.42	7.75	486
432.60	2414.75	2.63	83	2.09	22	0.049	112	208	30	17	26.06	0.54	0.64	0.033	1.18	3.09	201

The random variable Z_{AUC} satisfies a standard normal distribution. The critical value is obtained by comparing Z_{AUC} with the standard normal distribution table. Based on Table 2, we chose a Z_{AUC} greater than the critical value of 2.58 when the significance level was 0.01. In other words, a total of eight elements, including Au, Sn, Cu, Ag, Ba, Sb, Hg, and Mo, were selected as indicative elements for prospecting target prediction.

Table 2. AUC and Z_{AUC} calculational results of 16 geochemical elements.

Element	AUC	Z_{AUC}	Element	AUC	Z_{AUC}
Au	0.6024	2.8395	B	0.5901	2.4839
Sn	0.6065	2.9595	Cu	0.6311	3.6977
Ag	0.6762	5.1563	Ba	0.6147	3.2020
Mn	0.5573	1.5617	Pb	0.5778	2.1341
Zn	0.5450	1.2232	As	0.5655	1.7893
Sb	0.5942	2.6017	Bi	0.5901	2.4839
Hg	0.6393	3.9516	Mo	0.5983	2.7203
W	0.5778	2.1341	F	0.5696	1.9037

We combined experimental results and the geological environment of the study area for analysis. The Pangxidong study area has faults, developed fold structures, and strong magmatic activity. The metallogenic geological environment is similar to many areas with rich mineral resources in the Qinzhou–Hangzhou metallogenic belt. The types of minerals involved are mainly Au, Ag, Pb, Zn, W, Mo, Fe, and Mn metal ores. Then, Pb, Zn, and Cu metal ores are visible in Devonian strata in the study area. In addition, the northeast Gucheng–Shachan and Pangxitong faults in the study area are the main ore-conducting and storage structures for Au, Ag, Sn, and other minerals [30]. In summary, most of the eight indicator elements we selected are consistent with the metallic ores in the study area. This shows the validity of our choice of indicator elements.

By referring to the processing method of the stream sediment data in [31], we used inverse distance weight interpolation to generate 3072×3072 grids (elemental content map) according to the values of the geochemical elemental content. Specifically, we calculated the distance between the nearby discrete points and the grid (x_0, y_0) by

$$D_i = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2} \quad (3)$$

where D_i denotes the distance from the i -th discrete point near (x_0, y_0) , and (x_i, y_i) denotes the coordinates of the i -th discrete point.

Based on this, we estimated the value of the grid (x_0, y_0) as follows:

$$Z(x_0, y_0) = \sum_{i=1}^N \frac{1}{(D_i)^2} Z_i / \sum_{i=1}^N \frac{1}{(D_i)^2} \quad (4)$$

where $Z(x_0, y_0)$ denotes the estimated value at the grid (x_0, y_0) , Z_i denotes the observed value at the i -th discrete point, and N denotes the number of discrete points involved in the calculation.

Figure 3 shows the elemental content maps of the eight elements. We normalized the values of the elemental content map, making its mean value 0 and its variance 1. For a value x in the elemental content map, the normalized value is as follows:

$$\hat{x} = \frac{x - \bar{x}}{\sigma} \quad (5)$$

where $\hat{x}$ denotes the normalized value, $\bar{x}$ denotes the mean of the values of the elemental content map, and σ denotes the variance of the values of the elemental content map.

We divided the elemental content map of size 3072×3072 into two parts, where the upper part (2560×3072) was used for making the training dataset, and the lower part

(512 × 3072) was used for making the test dataset. We defined two 256 × 256 windows to slide on the upper and lower elemental content maps with a step length of 128. There were 437 patches in the training dataset, where 78 patches were with mines and 359 patches were without mines. On the other hand, the test dataset contained 69 patches, including 17 patches with mines and 52 patches without mines. In addition, Gaussian noise with a mean value of 0 and a variance of 0.01 was added to augment the data [32,33]. In the end, the generated training dataset had 2169 patches, with 1092 patches with mines and 1077 patches without mines, and the generated test dataset had 275 patches, with 119 patches with mines and 156 patches without mines. We combined the elemental content maps of the eight elements, i.e., our experimental dataset consisted of 2444 patches with a size of 256 × 256 × 8. In addition, we would randomly clip the data into a 224 × 224 size before each iteration as the input of our model.

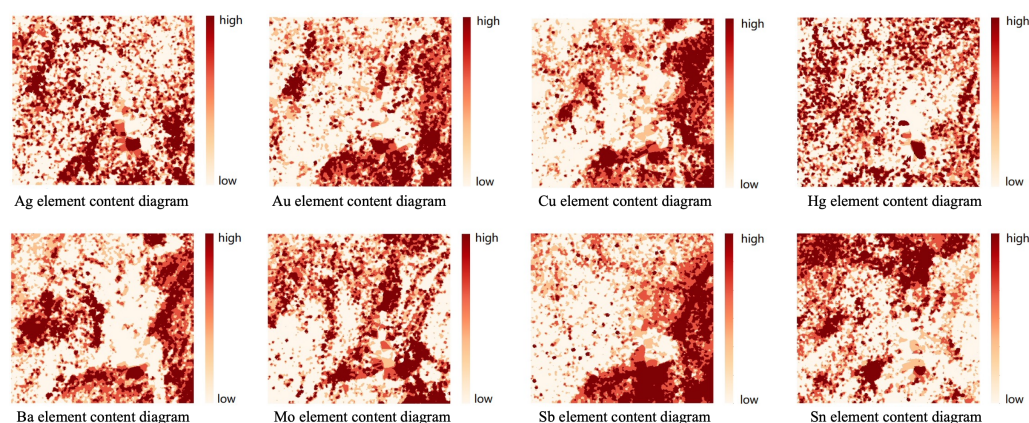


Figure 3. Elemental content diagrams.

3. Methodology

In this section, we propose a method of selective knowledge transfer for prospecting target prediction. Specifically, Section 3.1 formally addresses this issue. Section 3.2 explains how to use the soft mask to keep the weight between the associated mineral elements and the main mineral elements consistent and use dilated convolution to capture feature maps of different scales. Section 3.3 discusses selective knowledge transfer. Section 3.4 raises self-distillation to mine the hidden knowledge of target networks with different multiscale features. Section 3.5 exports the objective function of SKT used to deploy this mechanism.

3.1. Problem Formulation

In this section, we transform geochemical data into a geochemical elemental content map through inverse distance weight interpolation [31]. Then, we define a sliding window with a size of 256 × 256 by referring to [34]. At the same time, a geochemical dataset $\mathcal{D} = \{x_h^i, y_h^i\}_{h=1, i=1, 2, \dots, n}^N$ is constructed by cutting the elemental content map with a step size of 128. Here, N denotes the number of samples, i denotes the class of the geochemical element, and $x_h \in \mathbb{R}^d$ and $y_h \in \{0, 1\}$ denote the feature vector and the corresponding label, where 0 and 1 mean “without mines” and “with mines”, respectively, and d denotes the size of the sample feature space. We suppose that the SKT framework consists of a well-trained source task network $\mathcal{N}_S$ and multiple target task networks $\mathcal{N}_T = \{N_{t_1}, N_{t_2}, \dots, N_{t_m}\}$. These networks have the same structure and L convolution layers, where the numbers of input and output channels in the l -th convolution layer are M_l and M_{l+1} , respectively. The convolution kernel at the l -th layer in the source network $\mathcal{N}_S$ is defined as $W_S^l = [w_{S1}^l, w_{S2}^l, \dots, w_{SM_{l+1}}^l]$, where $W_S^l \in \mathbb{R}^{M_{l+1} \times M_l \times K \times K}$, $w_{Sj}^l \in \mathbb{R}^{M_l \times K \times K}$, and $K \times K$ is the size of the convolution kernel. For all convolution kernels at the l -th layer of $\mathcal{N}_T$, it is defined as $\mathcal{W}_T^l = [W_{t_1}^l, W_{t_2}^l, \dots, W_{t_e}^l] \in \mathbb{R}^{e \times M_{l+1} \times M_l \times K \times K}$, where the convolution kernels of

the task network t_v are $W_{t_v}^l = \{w_{t_v1}^l, w_{t_v2}^l, \dots, w_{t_vM_{l+1}}^l\}$, and the setting of the convolution kernel is consistent with that in $\mathcal{N}_S$. We propose using selective transfer knowledge for prospecting target prediction.

3.2. Congruence of Related Mineral Elements

We consider that the concentration of geochemical elements has indicative significance for mineralization, and different elements have differences in predicting mineralization. Generally, mined minerals contain associated minerals. Their main mineral elements have an important indicative significance for mineralization, while the associated mineral elements also have a certain indicative significance [35]. Therefore, we try to introduce a soft mask $\mathbf{M} = [\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_n]$ strategy based on the above consideration. Its purpose is to make the corresponding weight of the associated mineral elements as consistent as possible with that of the main mineral elements and increase the diversity of effective samples as follows:

$$\hat{x}_h^i = x_h^i \times \mathcal{M}_i \quad (6)$$

where x_h^i denotes the h -th sample in the i -th geochemical element, and $\mathcal{M}_i$ denotes the weight corresponding to the i -th geological element in $\mathbf{M}$, which is a scalar; $\hat{x}_h^i$ denotes the h -th sample in the i -th geochemical element by the mask operation. In this way, we obtain the feature maps of different geochemical elements.

To meet the challenge brought by the irregular features of the mining areas, we perform the dilated convolution operation on the features of different geochemical elements to generate feature maps of different scales. Specifically, we define a set of dilated coefficients $\rho = \{\rho_1, \rho_2, \dots, \rho_m\}$, where each coefficient corresponds to a feature map and target network at different scales. Then, we perform a dilated convolution operation on the v -th scale feature map as the input of the target network N_{t_v} , where the value of the position $(p_{\text{row}}, p_{\text{col}})$ on the feature map is calculated by

$$\mathbf{f}_{t_v}^i = \sum_{u=-r}^r \sum_{g=-r}^r \hat{x}_{t_v}^i(\rho_{t_v}u + p_{\text{row}}, \rho_{t_v}g + p_{\text{col}}) \times o(u, g) \quad (7)$$

where $\mathbf{f}_{t_v}^i$ denotes the t_v -th sample scale feature map of the i -th geochemical element through dilated convolution, ρ_{t_v} denotes the corresponding dilated coefficient to N_{t_v} , $\hat{x}_h^i(\cdot, \cdot)$ denotes the weight of the h -th sample position of the i -th geochemical element, and $o(\cdot, \cdot)$ denotes the weight of the position on the dilated convolution kernel. u, g denotes the nearest neighbor units of the position $(p_{\text{row}}, p_{\text{col}})$, and $r = S - 1$, where S is the size of the dilated convolution kernel. For example, for a dilated convolution kernel of size 3, the nearest neighbor of u and g is 1.

3.3. Selective Knowledge Transfer

Due to the small number of samples, it easily leads to problems such as underfitting or nonconvergence in the training process of the prospecting target prediction model. Thus, we selectively transfer elements in the convolution kernels from the well-trained source task network $\mathcal{N}_S$ to assist target network learning, as shown in Figure 4. Specifically, we first define a matrix $\mathbf{P}_{t_vj}^l$ with the same size as the convolution kernel in the target network N_{t_v} . Then, we calculate the Hadamard product of $W_{t_v}^l$ as follows:

$$\hat{w}_{t_vj}^l = w_{t_vj}^l \odot \mathbf{P}_{t_vj}^l \quad (8)$$

where $\hat{w}_{t_vj}^l$ denotes the selected convolution kernels, and $\odot$ denotes the Hadamard product.

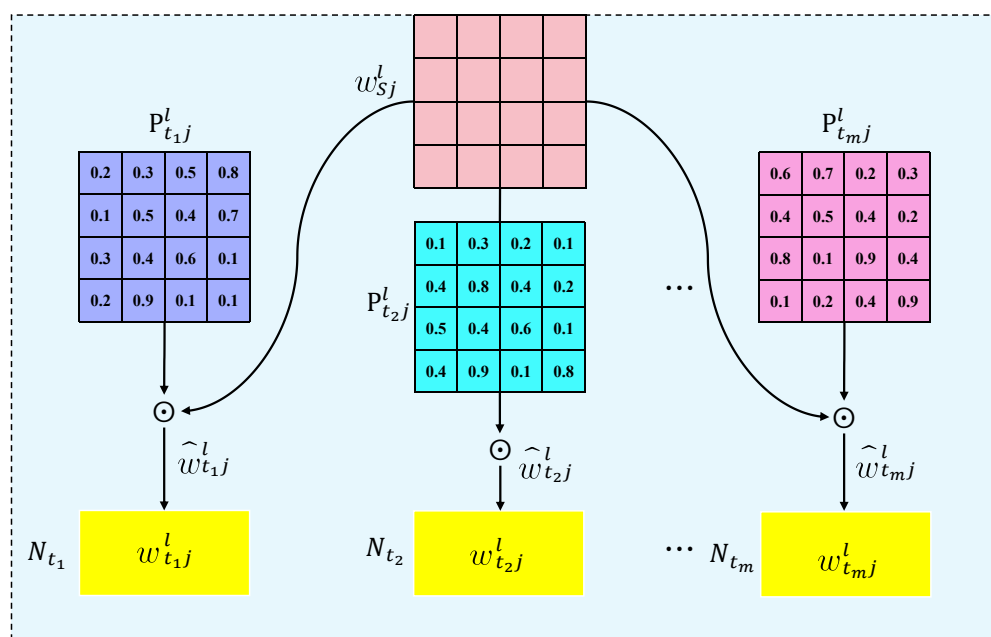


Figure 4. Description of the weight selection algorithm. w_{sj}^l denotes the convolution kernels of the l -th convolution layer of the source network, $j = 1, 2, \dots, M_{l+1}$ denotes the output channels, $P_{t_vj}^l$ denotes a matrix corresponding to the target network N_{t_v} , $v = 1, 2, \dots, m$, $\odot$ denotes the Hadamard product, $\hat{w}_{t_vj}^l$ denotes the selected convolution kernels, and $w_{t_vj}^l$ denotes the convolution kernels of N_{t_v} .

To further help target task network training, we transfer $\hat{w}_{sj}^l$ to $\mathcal{N}_T$ for training by

$$f_{t_v} = \left(\hat{w}_{sj}^l \times w_{t_vj}^l \right) * \mathbf{f}_{t_v}^i + b \quad (9)$$

where f_{t_v} denotes the convolution operation output, $*$ denotes the convolution operation, b denotes the convolution bias, and $\mathbf{f}_{t_v}^i$ denotes the t_v -th sample scale feature map of the i -th geochemical element.

3.4. Self-Distillation

To mine the hidden knowledge between the feature maps of different scales, we perform knowledge distillation from top to bottom according to the target network corresponding to the size of the feature maps. For example, there are three target networks, N_{t1} , N_{t2} , and N_{t3} , and the size of their input feature maps decreases successively. We use N_{t1} to guide N_{t2} and N_{t3} , and N_{t2} to guide N_{t3} . Specifically, we use Kullback–Leibler (KL) divergence to calculate the probability distribution of the softmax output between each pair of target networks by

$$\mathcal{L}_{t_vKD} = \sum_{j=1}^{v-1} \sum_{i=1}^n f(\mathbf{f}_{t_j}^i, \theta^{N_{t_j}}) \log \left(\frac{f(\mathbf{f}_{t_j}^i, \theta^{N_{t_j}})}{f(\mathbf{f}_{t_v}^i, \theta^{N_{t_v}})} \right) \quad (10)$$

where $\mathcal{L}_{t_vKD}$ denotes the self-distillation loss of the target network N_{t_v} ; the size of the input feature maps of N_{t_j} is larger than that of N_{t_v} ; $\theta^{N_{t_j}}$ and $\theta^{N_{t_v}}$ denote the parameters of the target networks N_{t_j} and N_{t_v} , respectively; $f(\cdot)$ denotes the softmax operation; and $\mathbf{f}_{t_j}^i$ and $\mathbf{f}_{t_v}^i$ denote the t_j -th and t_v -th sample scale feature maps of the i -th geochemical element through the dilated convolution operation.

3.5. Objective Function

The target network maps features to the corresponding label space through the full connected layer. During the training, we designed the classification loss for each target network by

$$\mathcal{L}_{t_v} = \sum_{i=1}^n L(\theta^{N_{t_v}}, \mathbf{f}_{t_v}^i, \mathbf{P}_{t_v}^{1:L}, y_h^i) \quad (11)$$

where $\mathcal{L}_{t_v}$ denotes the classification loss of the target network N_{t_v} , $L(\cdot)$ denotes cross entropy, $\theta^{N_{t_v}}$ denotes the parameter of N_{t_v} , y_h^i is the label for input x_h^i , $\mathbf{f}_{t_v}^i$ denotes the t_v -th sample scale feature map of the i -th geological element obtained through the mask operation and dilated convolution operation, and $\mathbf{P}_{t_v}^{1:L}$ denotes the matrix associated with selective knowledge transfer.

In the end, the overall optimization objective of SKT is to minimize the classification and self-distillation losses by

$$\mathcal{L}_{total} = \sum_{v=1}^m \mathcal{L}_{t_v} + \beta \sum_{v=1}^m \mathcal{L}_{t_vKD} \quad (12)$$

where $\mathcal{L}_{total}$ denotes the objective function, and β denotes the coefficient of self-distillation loss.

4. Experiments

In this section, we present extensive experiments to verify our method. Specifically, Section 4.1 describes the experimental environment and settings. Section 4.2 presents the comparison of our method with state-of-the-art methods on the geochemical dataset and the analysis of the experimental results. Section 4.3 presents the experimental results based on the relevant modules and parameters of our method. Section 4.4 visualizes the prediction result of the Pangxidong study area in Guangdong Province.

4.1. Experimental Settings

The evaluation metrics used in the experiments include Accuracy, Recall, and F1-score. All experiments are programmed and implemented with the PyTorch framework and one GeForce RTX 3090 GPU.

The SKT framework is implemented based on the ResNet-18 architecture. During training, the model uses an SGD optimizer with a momentum of 0.01, a weight decay of 1×10^{-4} , and a mini-batch size of 128. The learning rate is initially 0.1 and decreases by half every 30 epochs. We set the dilated coefficients $\rho = \{\rho_1, \rho_2, \rho_3, \rho_4, \rho_5\} = \{1, 6, 12, 18, 24\}$, referring to DeepLab [36]. The model is trained for 150 epochs on the geochemical dataset. In addition, the coefficient β of the self-distillation loss is set to 0.7.

4.2. Experimental Results and Analysis

In this section, we compare machine learning algorithms and some state-of-the-art classification methods with SKT to demonstrate that it outperforms other models in prospecting target prediction tasks. Specifically, we compare the following methods, including the traditional methods SVM [37], KNN [38], RandomForest [39], and Decisiontree [40] and the deep learning methods ResNet-18, ShufflenetV2 [41], GoogLeNet [42], MobilenetV2 [43], Mnasnet [44], SCnet [45], Efficientnet-b0 [46], T2T-vit-14 [47], and SNL [48]. It is worth noting that for fairness of comparison, SCnet and SNL are implemented based on the ResNet-18 architecture. In a machine learning algorithm, we compress each data point in the geochemical dataset into a one-dimensional tensor as the input to the algorithm.

As shown in Table 3, our model obviously performs better than ResNet-18. Specifically, the Accuracy of our model increases by 12.30%, Recall increases by 15.83%, and F1-score improves by 11.79%. Furthermore, SKT outperforms the other methods in terms of Accuracy, Recall, and F1-score. This indicates that SKT has excellent performance in prospecting target prediction and has the highest improvement in the prediction of samples with mines. Meanwhile, this also proves that our method can effectively solve problems

such as the small number of geological samples and the irregular features of mining areas in prospecting target prediction.

Table 3. Experimental results. The optimal performances are bolded.

Methods	Accuracy	Recall	F1-Score
SVM	49.51	17.64	43.73
KNN	51.45	35.29	50.09
RandomForest	59.70	25.49	54.27
Decisiontree	58.73	39.21	57.03
ResNet-18	56.79	24.50	53.21
ShufflenetV2	57.45	17.64	48.24
GoogLeNet	61.81	31.09	56.51
MobilenetV2	55.82	16.66	47.74
Mnasnet	59.22	17.64	50.61
SCnet	58.73	30.39	55.05
Efficientnet-b0	57.28	23.52	51.70
T2T-vit-14	57.76	39.21	56.19
SNL	59.70	35.29	57.07
Ours	69.09	40.33	65.00

Figure 5 is a confusion matrix, showing the number of True Negative (TN), False Positive (FP), False Negative (FN), and True Positive (TP) samples. This figure shows that (1) among these four values, TN is the highest, i.e., the number of correctly predicted the samples without mines is the largest. At the same time, FP is the lowest, i.e., the number of incorrectly predicted samples without mines is the smallest. We explain that before using Gaussian noise for data augmentation, the samples without mines are much more than the samples with mines, and the samples without mines have rich features, which are helpful to the STS framework for prediction. (2) TP denotes the number of correctly predicted samples with mines. On the contrary, FN denotes the number of incorrectly predicted samples with mines. TP is the third of the four values, lower than FN but much higher than FP. The reason for this result is the small number of samples with mines and their irregular features, which negatively affects the STS framework for predicting the samples with mines, but still has a certain performance. In conclusion, SKT can predict the samples without mines well, and it also has a certain ability to predict the samples with mines.

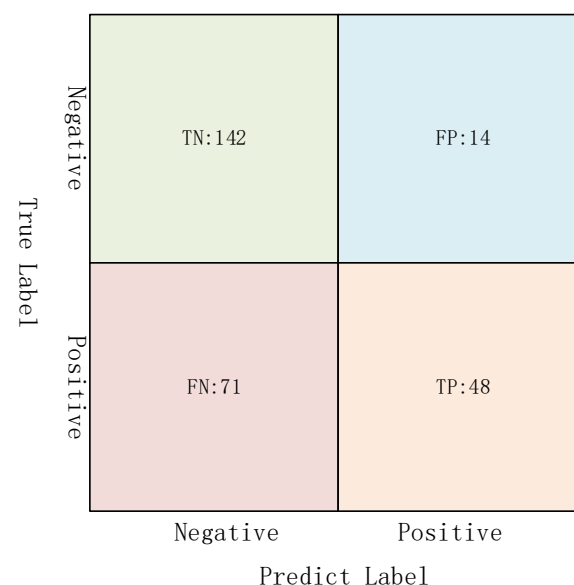


Figure 5. Confusion matrix.

In the SKT framework, the final prediction result is obtained by multiple target network voting. Table 4 shows the predicted results for the multiple target networks and voting. The following can be seen from this table: (1) As the dilated coefficient increases, the performance of the target networks first increases and then decreases. (2) When compared with the best target network, voting only has a small gap in Accuracy and F1-score, but there is a certain gap in Recall. The above result is caused by the performance of the model being affected with the increase in expansion coefficient, but at the same time, self-distillation can improve the performance of the target network, leading to first the increase and then the decrease. In addition, due to the contradiction between the dilated coefficient and self-distillation for the improvement of the model performance, we cannot determine the optimal target network. However, voting has a similar performance to the optimal target network, so the final prediction result obtained by voting is reasonable to a certain extent.

Table 4. Experimental results of the target networks and voting.

Target Network	Accuracy	Recall	F1-Score
$\rho = 1$	57.45	31.93	53.30
$\rho = 6$	65.45	41.17	62.08
$\rho = 12$	61.45	35.29	57.38
$\rho = 18$	70.18	47.05	67.34
$\rho = 24$	64.72	37.81	60.70
Voting	69.09	40.33	65.00

4.3. Correlation Analysis Experiment

We present the experimental results based on the relevant modules of SKT in Section 4.3.1. Then, in Section 4.3.2, we verify the effect of the coefficient of self-distillation loss on SKT.

4.3.1. Ablation Experiments

To evaluate the performance of the method proposed in this paper, we perform SKT ablation experiments and deploy various SKT variants. Specifically, we design the following ablation experiments: (1) remove soft mask (R-S-Mask), (2) remove dilated convolution (R-D-Convolution), (3) remove selective knowledge transfer (R-Sk-Transfer), and (4) remove self-distillation (R-S-Distillation). Based on SKT, we apply the control variable approach to soft mask, dilated convolution, and selective knowledge transfer one at a time. It is worth noting that we set the dilated coefficients $\rho = \{\rho_1, \rho_2, \rho_3, \rho_4, \rho_5\} = \{1, 1, 1, 1, 1\}$ in the second experiment. Finally, SKT is compared to the above methods.

Table 5 presents the experimental results. The following observations are made:

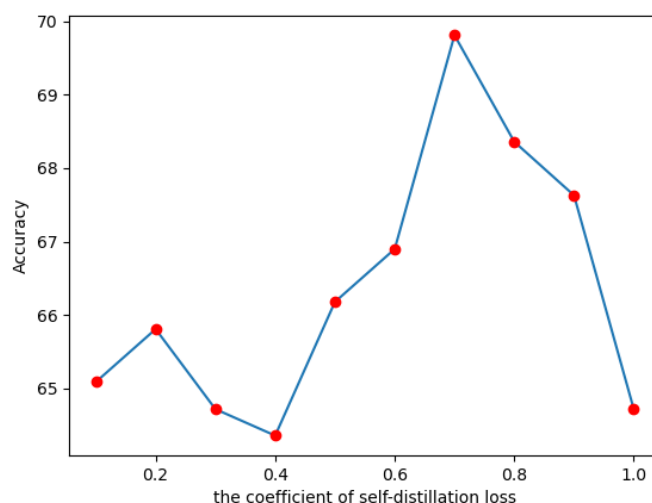
1. The soft mask makes the corresponding weight of the associated mineral elements as consistent as possible with that of the main mineral elements. Dilated convolution deals with the irregular features of the mining areas through different receptive fields. Selective knowledge transfer improves the model generalization performance to solve the problem of a small number of samples. Self-distillation mines the hidden knowledge between the feature maps of different scales. All of the aforementioned methods can improve the Accuracy, Recall, and F1-score of the prospecting prediction.
2. The contributions of these methods to SKT are different. According to the contribution from large to small, they are ranked as follows: dilated convolution, selective knowledge transfer, soft mask, and self-distillation.

Table 5. Ablation experiment results. The optimal performances are in bold.

Methods	Accuracy	Recall	F1-Score
R-S-Mask	64.72	29.41	55.43
R-D-Convolution	61.09	29.41	58.29
R-Sk-Transfer	62.54	30.25	56.83
R-S-Distillation	65.81	31.09	59.72
Ours	69.09	40.33	65.00

4.3.2. Parameter Analysis Experiments

The objective function of SKT includes self-distillation loss. To evaluate the effect of the coefficient of self-distillation loss β on SKT, we set $\beta = 0.1, 0.2, \dots, 1$ to perform a total of 10 experiments. Figure 6 shows the experimental results. From this figure, we can find the following: (1) β has a certain effect on the SKT performance. (2) The best prospecting target prediction results can be obtained when β is between 0.6 and 0.8.

**Figure 6.** The coefficient of self-distillation loss experiment results.

4.4. Visualization

In this section, we use the SKT trained in Section 4.2 to predict the prospecting target area in the Pangxidong study area and visualize the prediction results. Specifically, we first cut the elemental content map of size 3072×3072 obtained in Section 2.2 into 12×12 patches, each with a size of 256×256 . Then, we use the trained SKT to predict and visualize it. Figure 7 shows the visualization result. Based on the visualization, we come to the following conclusions: (1) The predicted prospecting target area is basically consistent with the actual mining area. (2) The prediction results of (row 7, column 4), (row 11, column 10), (row 12, column 4), and (row 12, column 11) are inconsistent with the actual mining area. In addition, we use the principal component analysis (PCA) method to reduce the dimensionality of geochemical data in the experiment to one dimension and visualize it [49]. As seen from Figure 8, we can find that (1) most of the mining areas correspond to high values, i.e., rich in these eight geochemical elements. This proves that geochemical elements have an important influence on mineralization, which is consistent with the conclusion of paper [50,51]. (2) Irregular features of geochemical-enriched regions in incorrectly predicted samples with mines (green boxes) make STS prediction difficult. In summary, we can conclude that SKT can fit the distribution of the prospecting target area in the Pangxidong study area. This again shows the effectiveness of SKT in prospecting target prediction.

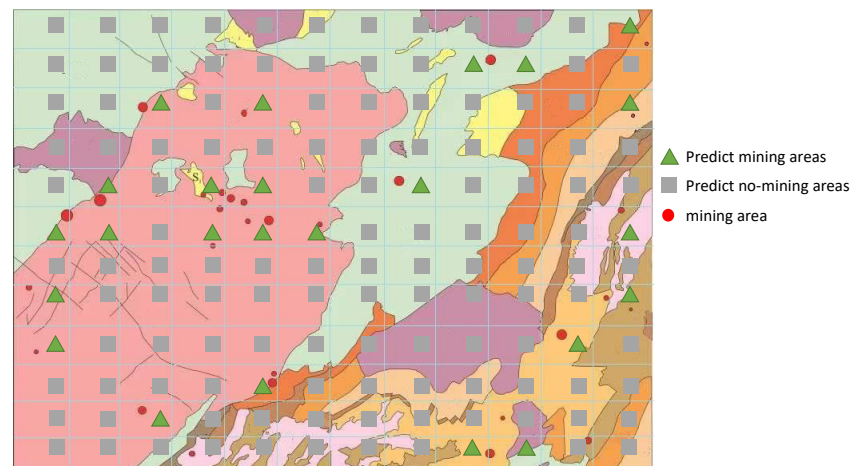


Figure 7. Visualization of experimental results in Pangxidong study area.

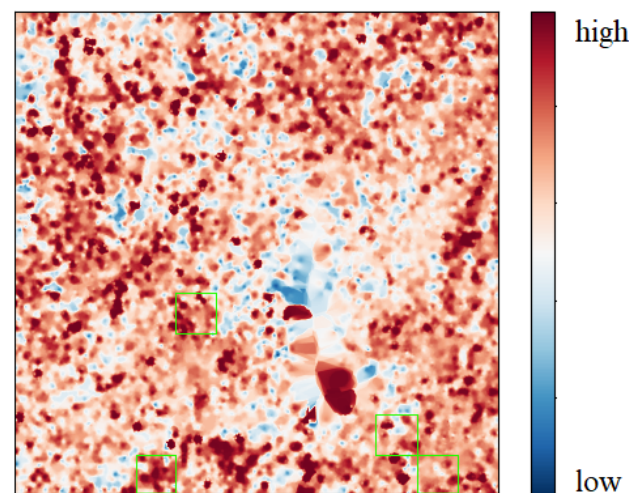


Figure 8. Visualization of geochemical data.

5. Conclusions

In this paper, a deep learning framework (SKT) for prospecting target prediction based on selective knowledge transfer is proposed to solve the problems of fewer geological sample data and irregular mining area features in the research of intelligent prospecting prediction. Taking the Pangxidong study area of Guangdong Province as an example, the prospecting target area is intelligently predicted by using the geochemical data of this area. Compared with other methods, the effectiveness of this method is proved. The main conclusions are:

- (1) In view of problems such as the small number of geological samples and the irregular features of mining areas in the research of prospecting prediction, the deep learning framework (SKT) for prospecting target prediction based on selective knowledge transfer has greatly improved the prediction of the samples with mines, which is obviously superior to other methods.
- (2) Soft mask makes the corresponding weight of associated mineral elements consistent with that of the main mineral elements as much as possible; dilation convolution enriches irregular features of the mining areas through capturing features at different scales; selective knowledge transfer improves the generalization performance of the model and solves the problem of a small number of samples; and self-distillation mines the hidden knowledge between different scale feature maps.

- (3) Parameter analysis experiments show that dilation convolution, selective knowledge transfer, soft mask, and self-distillation can improve the accuracy of SKT prediction, but their contribution to SKT gradually weakens.

Author Contributions: Conceptualization and methodology, Y.H. and Q.F.; software and validation, Y.H.; resources and data curation, L.G.; writing—original draft preparation, Y.H. and Q.F.; writing—review and editing, Y.H. and L.Z.; visualization, W.Z.; funding acquisition, L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was produced by “the Natural Science Foundation of Guangdong Province (18zxxt52), the Wuyi University Youth Team Fund (2019td10), and the Wuyi University-Hong Kong-Macau Joint Fund (2019WGALH23)”.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Scharf, T.; Kirkland, C.; Daggitt, M.; Barham, M.; Puzyrev, V. AnalyZr: A Python application for zircon grain image segmentation and shape analysis. *Comput. Geosci.* **2022**, *162*, 105057. [\[CrossRef\]](#)
2. Middy, A.I.; Nag, B.; Roy, S. Deep learning based multimodal emotion recognition using model-level fusion of audio–visual modalities. *Knowl.-Based Syst.* **2022**, *244*, 108580. [\[CrossRef\]](#)
3. Cui, S.; Ma, A.; Zhang, L.; Xu, M.; Zhong, Y. MAP-Net: SAR and Optical Image Matching via Image-Based Convolutional Network With Attention Mechanism and Spatial Pyramid Aggregated Pooling. *IEEE Trans. Geosci. Remote. Sens.* **2022**, *60*, 1–13. [\[CrossRef\]](#)
4. Elashmawy, M.; Alatawi, I. Atmospheric water harvesting from low-humid regions of Hail City in Saudi Arabia. *Nat. Resour. Res.* **2020**, *29*, 3689–3700. [\[CrossRef\]](#)
5. Eppelbaum, L.; Eppelbaum, V.; Ben-Avraham, Z. Formalization and Estimation of Integrated Geological Investigations: An Informational Approach. *Geoinformatics* **2003**, *14*, 233–240. [\[CrossRef\]](#)
6. Siebels, K.; Goita, K.; Germain, M. Estimation of Mineral Abundance From Hyperspectral Data Using a New Supervised Neighbor-Band Ratio Unmixing Approach. *IEEE Trans. Geosci. Remote. Sens.* **2020**, *58*, 6754–6766. [\[CrossRef\]](#)
7. Li, S.; Chen, J.; Liu, C. Overview on the Development of Intelligent Methods for Mineral Resource Prediction under the Background of Geological Big Data. *Minerals* **2022**, *12*, 616. [\[CrossRef\]](#)
8. Jooshaki, M.; Nad, A.; Michaux, S. A systematic review on the application of machine learning in exploiting mineralogical data in mining and mineral industry. *Minerals* **2021**, *11*, 816. [\[CrossRef\]](#)
9. Duda, R.O.; Hart, P.E. *Pattern Classification and Scene Analysis*; Wiley: New York, NY, USA, 1973; Volume 3.
10. Zekri, H.; Cohen, D.R.; Mokhtari, A.R.; Esmaili, A. Geochemical prospectivity mapping through a feature extraction–selection classification scheme. *Nat. Resour. Res.* **2019**, *28*, 849–865. [\[CrossRef\]](#)
11. Daviran, M.; Maghsoudi, A.; Ghezelbash, R.; Pradhan, B. A new strategy for spatial predictive mapping of mineral prospectivity: Automated hyperparameter tuning of random forest approach. *Comput. Geosci.* **2021**, *148*, 104688. [\[CrossRef\]](#)
12. Zhang, S.; Carranza, E.J.M.; Xiao, K.; Wei, H.; Yang, F.; Chen, Z.; Li, N.; Xiang, J. Mineral Prospectivity Mapping based on Isolation Forest and Random Forest: Implication for the Existence of Spatial Signature of Mineralization in Outliers. *Nat. Resour. Res.* **2021**, 1–19. [\[CrossRef\]](#)
13. Wang, Z.; Zuo, R.; Dong, Y. Mapping of Himalaya Leucogranites Based on ASTER and Sentinel-2A Datasets Using a Hybrid Method of Metric Learning and Random Forest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2020**, *13*, 1925–1936. [\[CrossRef\]](#)
14. Wang, Y.; Zhou, Y.; Xiao, F.; Wang, J.; Wang, K.; Yu, X. Numerical Metallogenic Modelling and Support Vector Machine Methods Applied to Predict Deep Mineralization: A Case Study from the Fankou Pb–An ore Deposit in Northern Guangdong. *Geotecton. Metallog.* **2020**, *44*, 9. [\[CrossRef\]](#)
15. Mandana, T.; Behnam, B.; Saeed, D. Intelligent geochemical exploration modeling using multiclass support vector machine and integration it with continuous genetic algorithm in Gonabad region, Khorasan Razavi, Iran. *Arab. J. Geosci.* **2021**, *14*, 1–15. [\[CrossRef\]](#)
16. Liu, Y.; Zhu, L.; Zhou, Y. Experimental Research on Big Data Mining and Intelligent Prediction of Prospecting Target Area Application of Convolutional Neural Network Model. *Geotecton. Metallog.* **2020**, *44*, 1–11. [\[CrossRef\]](#)
17. Li, S.; Chen, J.; Xiang, J. Applications of deep convolutional neural networks in prospecting prediction based on two-dimensional geological big data. *Neural Comput. Appl.* **2020**, *32*, 2037–2053. [\[CrossRef\]](#)
18. Li, T.; Zuo, R.; Xiong, Y.; Peng, Y. Random-Drop Data Augmentation of Deep Convolutional Neural Network for Mineral Prospectivity Mapping. *Nat. Resour. Res.* **2020**, *30*, 27–38. [\[CrossRef\]](#)

19. Li, T.; Xia, Q.; Zhao, M.; Gui, Z.; Leng, S. Prospectivity Mapping for Tungsten Polymetallic Mineral Resources, Nanling Metallogenic Belt, South China: Use of Random Forest Algorithm from a Perspective of Data Imbalance. *Nat. Resour. Res.* **2020**, *29*, 203–227. [\[CrossRef\]](#)
20. Zhang, C.; Zuo, R.; Xiong, Y. Detection of the multivariate geochemical anomalies associated with mineralization using a deep convolutional neural network and a pixel-pair feature method. *Appl. Geochem.* **2021**, *130*, 104994. [\[CrossRef\]](#)
21. Li, D.; Yao, A.; Chen, Q. Learning to learn parameterized classification networks for scalable input images. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 19–35.
22. Yang, T.; Zhu, S.; Chen, C.; Yan, S.; Zhang, M.; Willis, A. Mutualnet: Adaptive convnet via mutual learning from network width and resolution. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 299–315.
23. Yang, N.; Zhang, Z.; Yang, J.; Hong, Z.; Shi, J. A Convolutional Neural Network of GoogLeNet Applied in Mineral Prospectivity Prediction Based on Multi-source Geoinformation. *Nat. Resour. Res.* **2021**, *30*, 3905–3923. [\[CrossRef\]](#)
24. Wang, X. Metallogenic Pattern and Mineral Prospectivity Modeling of the Dashui Gold Concentration District. Ph.D. Thesis, China University of Geosciences, Beijing, China, 2020.
25. Xiao, F.; Wang, K.; Hou, W.; Erten, O. Identifying geochemical anomaly through spatially anisotropic singularity mapping: A case study from silver-gold deposit in Pangxidong district, SE China. *J. Geochem. Explor.* **2020**, *210*, 106453. [\[CrossRef\]](#)
26. Zhou, Y.; Li, X.; Zhen, Y.; Shen, W.; He, J.; Yu, P.; Niu, J.; Zeng, C. Geological settings and metallogenesis of Qinzhou Bay-Hangzhou Bay orogenic juncture belt, South China. *Acta Petrol. Sin.* **2017**, *33*, 667–681.
27. Lin, Z.; Zhou, Y.; Qin, Y.; Zheng, Y.; Liang, Z.; Zou, H.; Niu, J. Ore-controlling structure analysis of Panxidong-Jinshan silver-gold orefield, southern Qin-Hang belt: Implications for further exploration. *Miner. Depos.* **2017**, *36*, 866–878.
28. Lobo, J.M.; Jiménez-Valverde, A.; Real, R. AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **2008**, *17*, 145–151. [\[CrossRef\]](#)
29. Zhang, S. Deep Learning for Mineral Prospectivity Mapping of Lala-Type Copper Deposit in the Huili Region, Sichuan. Ph.D. Thesis, China University of Geosciences, Beijing, China, 2020.
30. Zhang, Y.; Zhou, Y.Z.; Wang, L.F.; Wang, Z.H.; He, J.G.; An, Y.F.; Li, H.Z.; Zeng, C.Y.; Liang, J.; Lü, W.C.; et al. Mineralization-related geochemical anomalies derived from stream sediment geochemical data using multifractal analysis in Pangxidong area of Qinzhou-Hangzhou tectonic joint belt, Guangdong Province, China. *J. Cent. South Univ.* **2013**, *20*, 184–192. [\[CrossRef\]](#)
31. Zeyu, Z.; Qingying, Z.; Shixian, L. Comparison of two machine learning algorithms for geochemical anomaly detection. *Glob. Geol.* **2018**, *37*, 1288–1294.
32. Kumar, V.; Gupta, P. Importance of statistical measures in digital image processing. *Int. J. Emerg. Technol. Adv. Eng.* **2012**, *2*, 56–62.
33. Jia, M.; Dong, M. Analysis and comparison of Gaussian noise denoising algorithms. *J. Phys. Conf. Ser.* **2021**, *1846*, 012069. [\[CrossRef\]](#)
34. Zuo, R.; Peng, Y.; Li, T.; Xiong, Y. Challenges of geological prospecting big data mining and integration using deep learning algorithms. *Earth Sci.* **2021**, *46*, 350–358.
35. Zuo, R.; Wang, J.; Xiong, Y.; WANG, Z. Progresses of researches on geochemical exploration data processing during 2011–2020. *Bull. Mineral. Petrol. Geochem.* **2021**, *40*, 81–93.
36. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Lyu, P.; He, L.; He, Z.; Liu, Y.; Deng, H.; Qu, R.; Wang, J.; Zhao, Y.; Wei, Y. Research on remote sensing prospecting technology based on multi-source data fusion in deep-cutting areas. *Ore Geol. Rev.* **2021**, *138*, 104359. [\[CrossRef\]](#)
38. Chen, Y.; Zhao, Q.; Lu, L. Combining the outputs of various k-nearest neighbor anomaly detectors to form a robust ensemble model for high-dimensional geochemical anomaly detection. *J. Geochem. Explor.* **2021**, *231*, 106875. [\[CrossRef\]](#)
39. Wang, C.; Pan, Y.; Chen, J.; Ouyang, Y.; Rao, J.; Jiang, Q. Indicator element selection and geochemical anomaly mapping using recursive feature elimination and random forest methods in the Jingdezhen region of Jiangxi Province, South China. *Appl. Geochem.* **2020**, *122*, 104760. [\[CrossRef\]](#)
40. Dai, L.M.; Chen, Y.L.; Zhou, Y.G.; Liu, B.; Lou, D. A decision tree model for mineral potential mapping. *Prog. Geophys.* **2009**, *24*, 1081–1087.
41. Ma, N.; Zhang, X.; Zheng, H.T.; Sun, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 116–131.
42. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
43. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.

44. Tan, M.; Chen, B.; Pang, R.; Vasudevan, V.; Sandler, M.; Howard, A.; Le, Q.V. Mnasnet: Platform-aware neural architecture search for mobile. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2820–2828.
45. Liu, J.J.; Hou, Q.; Cheng, M.M.; Wang, C.; Feng, J. Improving convolutional networks with self-calibrated convolutions. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10096–10105.
46. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.
47. Yuan, L.; Chen, Y.; Wang, T.; Yu, W.; Shi, Y.; Jiang, Z.H.; Tay, F.E.; Feng, J.; Yan, S. Tokens-to-token vit: Training vision transformers from scratch on imagenet. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021; pp. 558–567.
48. Zhu, L.; She, Q.; Li, D.; Lu, Y.; Kang, X.; Hu, J.; Wang, C. Unifying Nonlocal Blocks for Neural Networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 12292–12301.
49. Amaral, T.G.; Pires, V.F.; Pires, A.J. Fault detection in PV tracking systems using an image processing algorithm based on PCA. *Energies* **2021**, *14*, 7278. [[CrossRef](#)]
50. Zhou, S.G.; Zhou, K.F.; Wang, J.L. Geochemical metallogenic potential based on cluster analysis: A new method to extract valuable information for mineral exploration from geochemical data. *Appl. Geochem.* **2020**, *122*, 104748. [[CrossRef](#)]
51. Ayari, J.; Barbieri, M.; Barhoumi, A.; Belkhiria, W.; Braham, A.; Dhaha, F.; Charef, A. A regional-scale geochemical survey of stream sediment samples in Nappe zone, northern Tunisia: Implications for mineral exploration. *J. Geochem. Explor.* **2022**, *235*, 106956. [[CrossRef](#)]