



## Detection of geochemical anomalies related to mineralization using the GANomaly network

Zijing Luo <sup>a</sup>, Renguang Zuo <sup>a,\*</sup>, Yihui Xiong <sup>a</sup>, Xueqiu Wang <sup>b,c</sup>

<sup>a</sup> State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan, 430074, China

<sup>b</sup> Key Laboratory of Geochemical Exploration of Ministry of Natural Resources, Institute of Geophysical and Geochemical Exploration, Langfang, 065000, China

<sup>c</sup> UNESCO International Centre on Global-scale Geochemistry, Langfang, 065000, China

### ARTICLE INFO

Editorial handling by Prof. M. Kersten

**Keywords:**

Geochemical mapping  
Generative adversarial networks  
GANomaly  
Deep learning  
Mineral exploration

### ABSTRACT

In this study, a GANomaly network was used to detect geochemical anomalies related to mineralization in the southern part of Jiangxi Province and its adjacent areas in China. The training data used in this study belong to a typical rare-sample category of imbalanced data samples; thus, during the training phase, only non-mineralized dataset randomly selected from the study area was used for training in order to avoid the overfitting problem caused by an imbalance between positive and negative training samples. The established GANomaly network structure can effectively extract anomalous geochemical information from the exploration geochemical data. The geochemical anomalies identified by GANomaly and known tungsten polymetallic deposits show a close spatial correlation. Further, the anomalous high-value areas are located in or around the Yanshanian intrusive rock. The performance of GANomaly for the identification of multivariate geochemical anomalies was compared to that of the deep autoencoder. The comparative results indicated that the GANomaly network can learn the internal connections and characteristics between multivariate geochemical data and can effectively avoid the influence of noise in geochemical data. Therefore, the abnormal areas identified by GANomaly are determined to be significant for mineral exploration.

### 1. Introduction

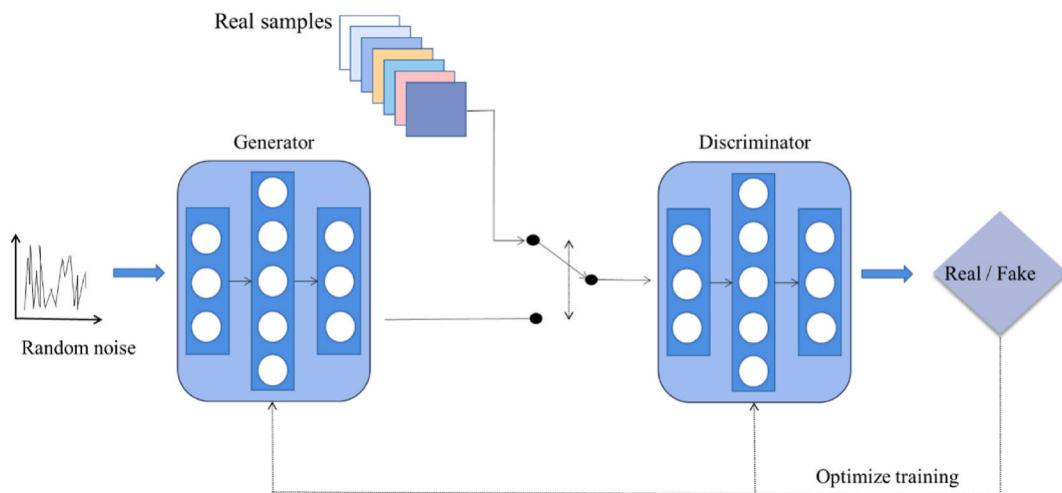
The complexity of geological systems and the multi-phase nature of ore-forming processes lead to a complex and non-single statistical distribution of geochemical data (Zuo et al., 2019). Traditional methods have limitations in processing the complex distribution of multivariate geochemical data (Xiong and Zuo, 2016; Parsa et al., 2017). Deep learning (DL) algorithms, which have been proven to be powerful tools for mining complex and nonlinear geospatial data and extracting unknown patterns related to geological processes, have been introduced in the field of exploration geochemistry (Luo et al., 2020; Chen et al., 2019; Zhang et al., 2021). Because the mineralization process is singular, mineral deposits are inherently rare (Cheng, 2007). One of the main limitations of using neural networks or other supervised machine learning algorithms is the scarcity of known mineral deposits (Li et al., 2021; Parsa, 2021). In general, the number of ore-bearing units (positive samples) is much smaller than that of ore-free units (negative samples), thus resulting in a data imbalance problem (Xiong and Zuo, 2018). Skewed class distributions will greatly underestimate the prediction

performance for minority classes (e.g., ore-bearing units) and provide an inaccurate evaluation of the classification performance (Longadge and Dongre, 2013).

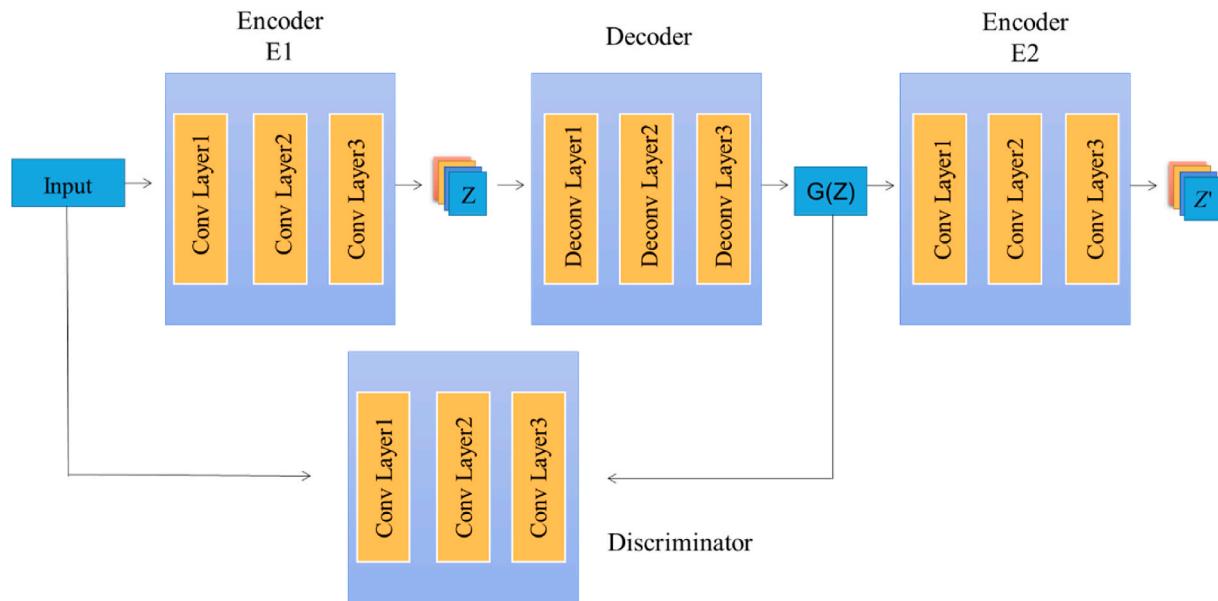
A generative adversarial network (GAN), proposed by Goodfellow et al. (2014), has become a leading method for addressing unsupervised and semi-supervised problems (Akçay et al., 2018). The basic concept of a GAN is derived from the zero-sum game in game theory, which enables discriminator  $D$  and generator  $G$  to improve the performance of the model during a mutual game (Goodfellow et al., 2014). GANs have been widely applied in the fields of image generation (e.g., Denton et al., 2015; Larsen et al., 2016), image recognition (e.g., Ledig et al., 2017), style transfer (e.g., Taigman et al., 2016; Isola et al., 2017), and anomaly detection (e.g., Schlegl et al., 2017; Zenati et al., 2018). To improve the performance of a GAN, a variety of improved algorithms, such as a conditional GAN (Mirza and Osindero, 2014), deep convolutional GAN (DCGAN) (Radford et al., 2015), and Wasserstein GAN (WGAN) (Arjovsky et al., 2017) have been proposed. In the field of geosciences, the application of a GAN is mainly focused on seismic data interpolation (Oliveira et al., 2018) and remote sensing classification (Merkle et al.,

\* Corresponding author.

E-mail addresses: [zrguang@cug.edu.cn](mailto:zrguang@cug.edu.cn) (R. Zuo), [xiongyh426@cug.edu.cn](mailto:xiongyh426@cug.edu.cn) (Y. Xiong).



**Fig. 1.** Architecture of deep generative adversarial networks (modified from Goodfellow et al., 2014).



**Fig. 2.** A diagram showing the building and calculating process of GANomaly.

2018).

The deep autoencoder and its variants, which are reconstruction-based approaches, have been widely adopted to investigate geochemical anomalies (Xiong and Zuo, 2016; Luo et al., 2020). Recently, the idea of blending the autoencoder architecture with the adversarial learning have demonstrated promise in anomaly detection problems, thus have received extensive attention. Beggel et al. (2019) adopt adversarial autoencoder architecture for anomaly detection by imposing a prior distribution on the latent representation, and then typically placing anomalies into low likelihood regions. Schlegl et al. (2017) employed the standard GAN model to train normal samples under the assumption that the latent vector of the GAN represents the true data distribution for anomaly detection. Zenati et al. (2018) proposed an Efficient-GAN-Anomaly for anomaly detection by jointly mapping from an original data space to a latent space.

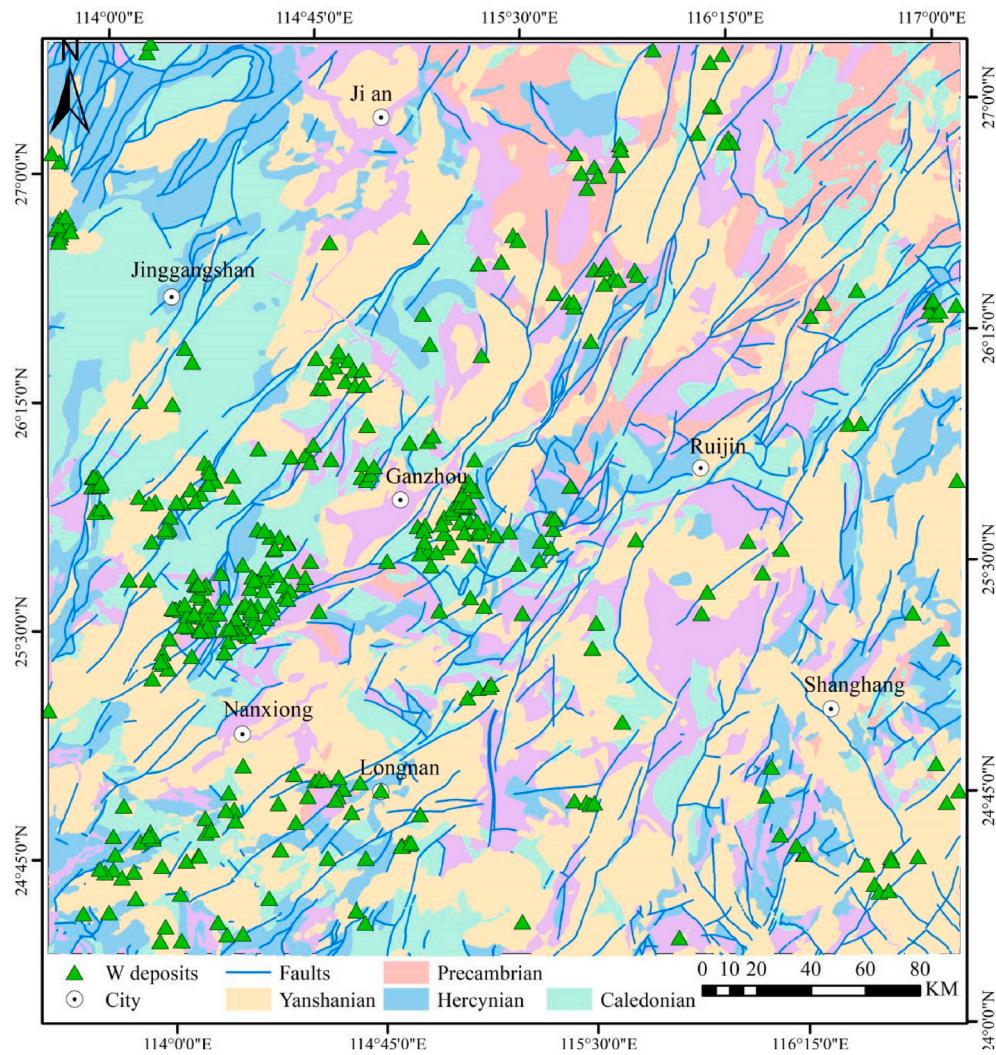
Different from the previous reconstructed-based model for anomaly detection, which calculate the reconstruction error based on the original data and their reconstructions, GANomaly model, proposed by Akcay et al. (2018), calculate the reconstruction error based on latent feature vector  $z$  and reconstructed latent feature vector  $z'$  by adding an

additional encoder structure. During the training phase, the model is aimed to learn the distribution of a normal sample to minimize the difference between the two latent feature vectors of the sample. Thus, during the testing phase, a sample with a large error between the two latent feature vectors is regarded as an abnormality (Akcay et al., 2018). The main aim of this study is to explore the efficient of GANomaly model to deal with multivariate geochemical exploration data and delineate geochemical anomalies associated with mineralization in southern Jiangxi and its adjacent areas in China.

## 2. Methods

### 2.1. GANomaly network

The GAN network consists of two parts: the generator ( $G$ ) and discriminator ( $D$ ) (Goodfellow et al., 2014) (Fig. 1). The input of generator is random noise that follows a certain distribution. The generator learns the distribution of real data from the latent space and generates new examples. The discriminator is a classic classification system structure that attempts to classify examples as either real (from



**Fig. 3.** A simplified geological map of the study area (southern Jiangxi and its adjacent areas in China).

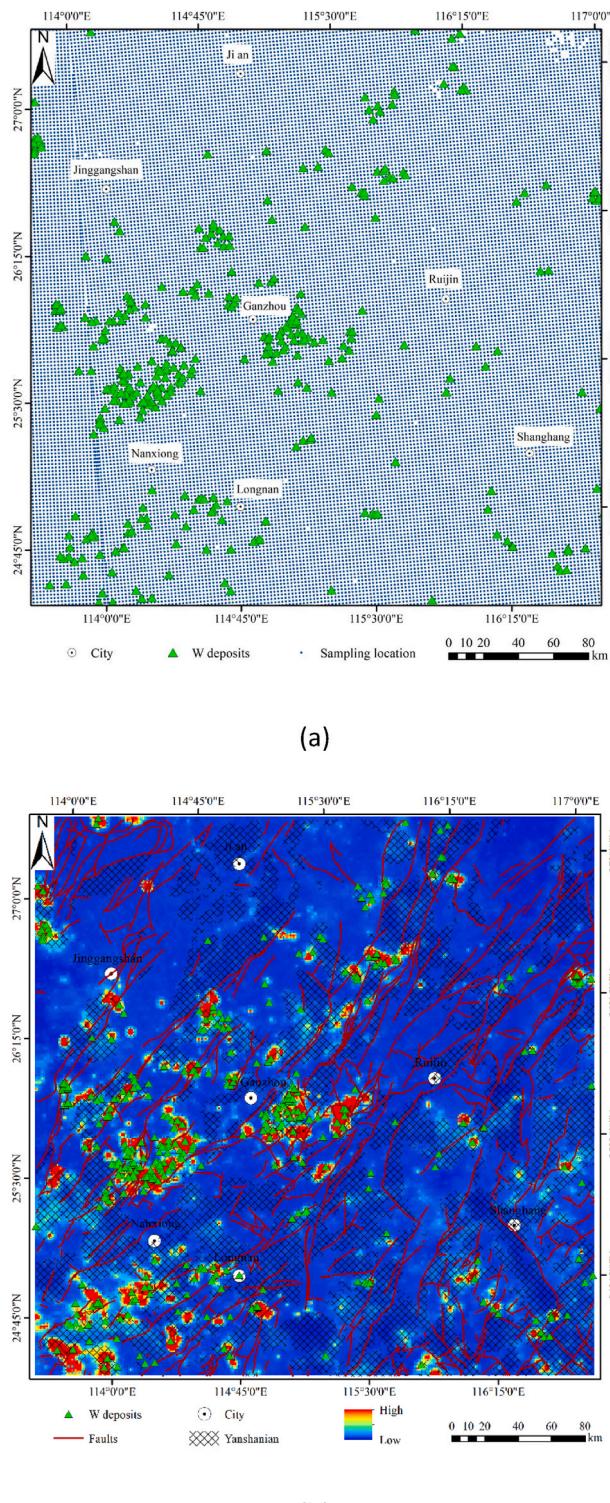
the domain) or fake (generated). The two models are trained together in a zero-sum game; that is, they are adversarial, until the discriminator cannot distinguish whether the sample is a generated or real sample, indicating the generator model can generate plausible examples.

The GANomaly model is a variant of a GAN. The architecture of GANomaly, blending the adversarial learning with an encoder-decoder-encoder pipeline, can capture the training data distribution within both original data space and latent vector space. The advantage of GANomaly for anomaly detection is that the mutual game played between the generator and discriminator in a GANomaly network can improve the ability of the decoder to recover the input data compared with the original autoencoder (AE). In addition, due to the reconstruction ability of AE and the generative capability of adversarial learning, it is possible to obtain features of the original data in deep latent space, thus can enhance the robustness for the effects of the noise interference and maintain the authenticity of the data (Chen et al., 2012). An GANomaly is composed of three sub-network parts (Akçay et al., 2018) (Fig. 2). The first sub-network is the AE network (including the encoder E1 and decoder part), which serves as the generator  $G$  part of the model. The normal sample  $x$  is input into the encoder E1 to obtain the latent feature vector  $z$ , and the latent feature vector  $z$  completes the data reconstruction using the decoder. The second sub-network is the encoder network E2, and the reconstructed data  $x'$  is input into the encoder network E2 to obtain the reconstructed latent feature vector  $z'$ . The reconstructed latent feature vector  $z'$  and latent feature vector  $z$  have the same

dimensions in order to ensure the consistency of the two encodings. The third sub-network is discriminator  $D$ , which is used for correctly distinguishing between data  $x$  and reconstructed data  $x'$ . When the discriminator has difficulty in distinguishing between these, it indicates that the generated data at this time are extremely similar to the normal sample data. When abnormal samples are input, the encoder E1 encodes the abnormal samples to obtain the latent feature vector  $z$ , but the decoder cannot reconstruct the abnormal samples correctly. This is because the training phase only uses normal samples for modeling, and its parameters are not suitable for abnormal samples. Therefore, the latent feature vector  $z'$  obtained by encoder E2 is extremely different from the latent feature vector  $z$ , and abnormal samples are identified by detecting the difference between these two vectors. The three sub-network parts of the GANomaly model in this study were constructed using the DCGAN model proposed by Radford et al. (2015). Because the convolution kernel operation can better learn the spatial characteristics of the data, it can effectively mine multiple types of geochemical information.

The loss of GANomaly network is mainly composed of three different losses which is described as follows:

- (1) Adversarial loss ( $L_{adv}$ ): To improve the ability of a decoder to restore the features and stability of the model training, the feature matching loss proposed by Salimans et al. (2016) is used for adversarial learning. This is different from the Vanilla GAN,



**Fig. 4.** Maps showing (a) sampling location and (b) W element concentration.

wherein the generator is updated according to the output of the discriminator. In the GANomaly network, the generator is updated based on the internal representation of the discriminator. Assuming a function  $f_0$ , the input  $x$  is sampled from the input data distribution  $p_x$ , the output of this function is the result of the middle layer feature of the discriminator  $D$ . Feature matching calculates the distance between the original feature

data and the feature obtained from the generated data, respectively.

$$L_{adv} = E_{x \sim p_X} \|f(x) - E_{x \sim p_X} f(G(x))\|_2 \quad (1)$$

(2) Contextual loss ( $L_{con}$ ): To optimize the similarity between the generated and real data, according to Isola et al. (2017), the GANomaly network penalizes generator  $G$  by defining the context loss ( $L_{con}$ ) to measure the L1 distance between the original data  $x$  and the generated data  $x'$ :

$$L_{con} = E_{x \sim p_X} \|x - G(x)\|_1 \quad (2)$$

(3) Encoder loss ( $L_{enc}$ ): This loss is used to minimize the distance between the latent feature vector  $z'$  generated by the encoder  $E_2$  and the latent feature vector  $z$  generated by encoder  $E_1$ , which aims to allow the generator to learn how to encode the features of the data generated by the normal sample:

$$L_{enc} = E_{x \sim p_X} \|G_E(x) - E(G(x))\|_2 \quad (3)$$

The objective function of the generator of the GANomaly network integrates the above three parts of the loss:

$$L = \omega_{adv} L_{adv} + \omega_{con} L_{con} + \omega_{enc} L_{enc} \quad (4)$$

where  $\omega_{adv}$ ,  $\omega_{con}$  and  $\omega_{enc}$  are the adjustable weight parameters of  $L_{adv}$ ,  $L_{con}$ , and  $L_{enc}$  respectively. Because noise interferes with the reconstruction error of the sample, which affects the detection ability, the GANomaly network no longer uses the reconstruction error of the sample as the basis for anomaly detection during the detection phase; instead, reconstruction errors of the deeper latent vector is used for anomaly detection. Thus, it can significantly improve the anti-interference ability of the model.

After the model training is completed, the testing sample obtains the latent feature vector  $z$  through encoder  $E_1$ , and then obtains the reconstructed latent feature vector  $z'$  through the decoder and encoder  $E_2$ . In this study, the average absolute error ( $A(x)$ ) between the two latent features  $z$  and the reconstructed latent feature vector  $z'$  was detected to identify abnormal samples in the test data.

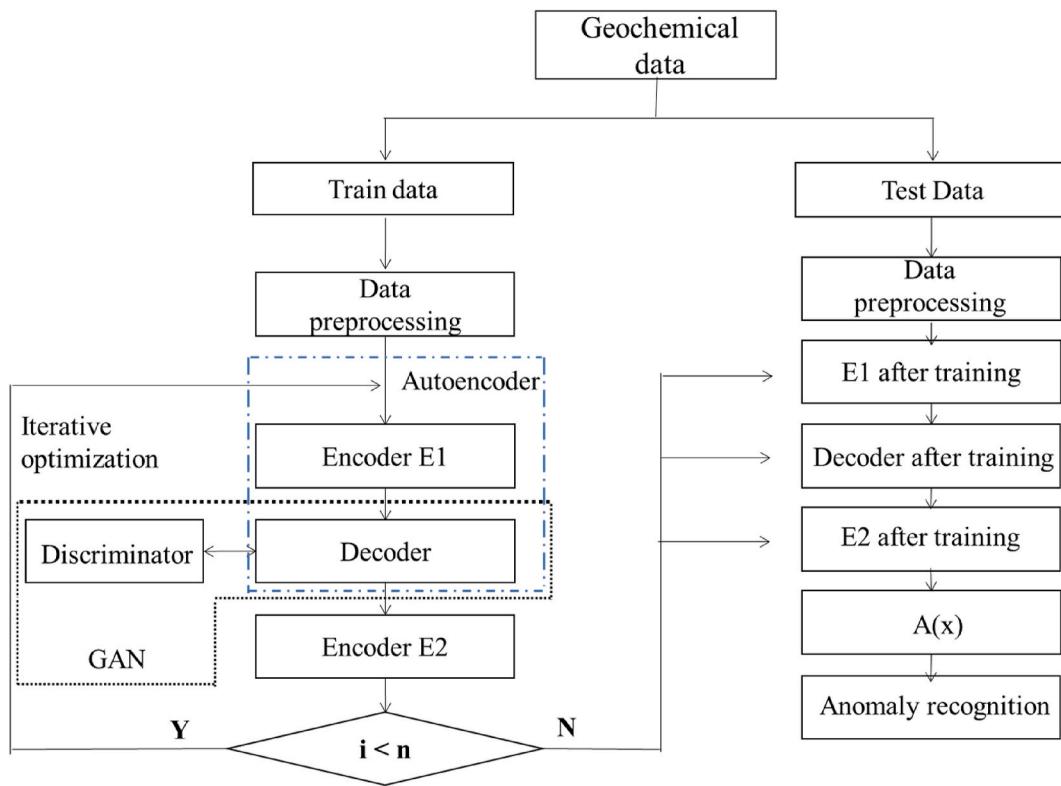
$$A(x) = \frac{1}{m} (\|G_E(x) - E(G(x))\|_1) \quad (5)$$

## 2.2. Network parameters

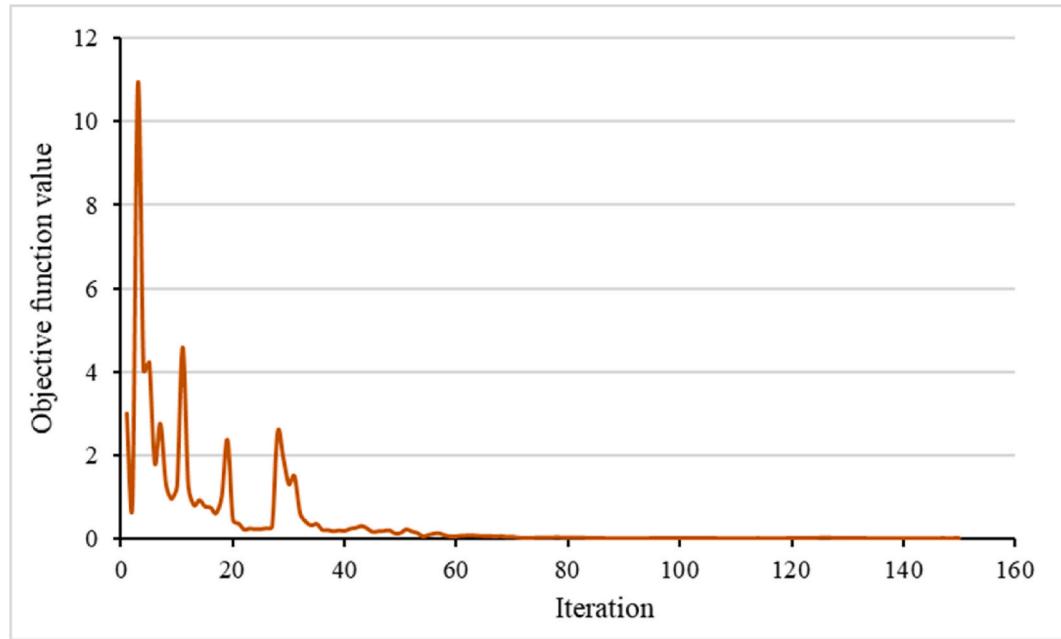
The GANomaly network structure (Fig. 2) was implemented in the Tensorflow framework using Python 3.5. Encoder  $E_1$  and the decoder are symmetric networks. Encoder  $E_1$  extracts the characteristics of the samples from a three-layer network. The number of convolution kernels in each layer of encoder  $E_1$  were set to 64, 128, and 256, respectively. Owing to the symmetrical structure of the GANomaly, the number of convolution kernels in each deconvolutional layer of the decoder were thus set to 256, 128, and 64, respectively. The convolution kernel size in both the encoder and decoder layers was set to  $4 \times 4$ . The intermediate convolution layer used the BatchNorm layer and the LeakyReLU activation function to optimize the output distribution of the intermediate layer and improve the training efficiency. The Adam optimization algorithm (Kingma and Ba, 2014) was used to optimize the network during training, and the initial learning rate  $r$  was set to 0.00001. In addition, the batch size used in the experiment is 128.

## 2.3. Performance evaluation

The receiver operating characteristic curve (ROC) (Fawcett, 2006; Parsa et al., 2018) and the success-rate curve (Agterberg and Bonham-Carter, 2005; Parsa et al., 2016) were adopted to measure the effectiveness of the geochemical anomalies extracted by the GANomaly



**Fig. 5.** A diagram showing the process of extracting geochemical anomaly information process using GANomaly (modified from [Dai et al., 2019](#)).



**Fig. 6.** A plot of variation in objective function versus the number of iterations.

model. The abscissa of the ROC curve represents the false positive rate:  $FP/(FP + TN)$ . The ordinate of the ROC curve represents the true positive rate:  $TP/(TP + FN)$ . The AUC value is the area under the ROC curve and the coordinate axis, and its range is between 0 and 1. The larger the AUC value, the better the effectiveness of the model. The success-rate curve shows the relationship between the deposit points and proportion of the study area. A geochemical anomaly map is divided into several categories according to the calculation results of the model, and

the number of grid elements of known deposits in each category is calculated to draw the success rate curve. The curve can prove the fitting degree of the model to the known deposit points. If a small area contains most of the deposit points, it is proved that the model is effective for simulating the known ore points.

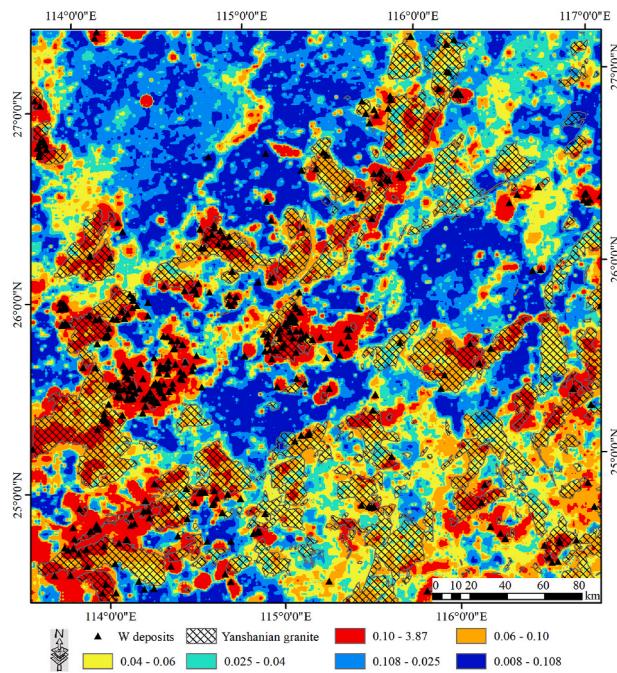


Fig. 7. A geochemical anomaly map obtained using GANomaly.

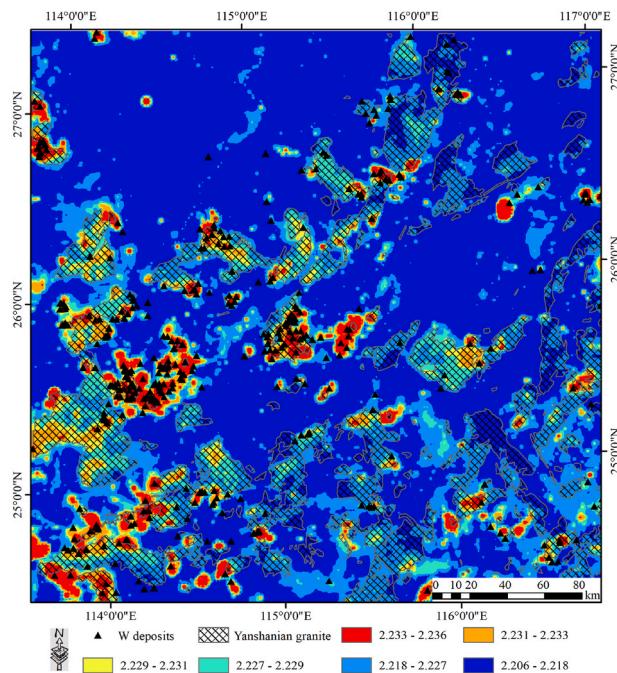


Fig. 8. A geochemical anomaly map obtained using deep AE.

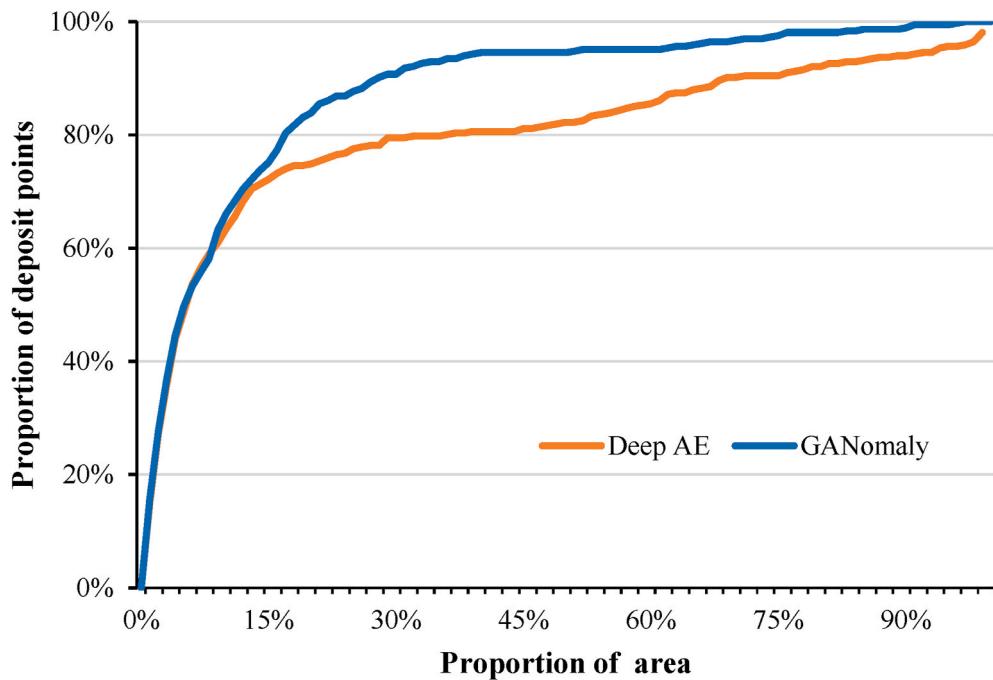
### 3. Study area and data

Jiangxi and its surrounding areas (Fig. 3) are world-class tungsten ore-concentrating areas with high mineralization intensity and high deposit density. Large-scale tungsten mineralization occurred in this area during the middle of the Yanshanian Movement. Among these areas, the southern Jiangxi region is the area with the most abundant tungsten mineral resources and has the longest mining history of wolframite in China. Geochemical concentration distributions show that large amounts of tungsten and tin accumulations are mainly distributed in the Cathaysian Block, which are closely related to the metallogenic

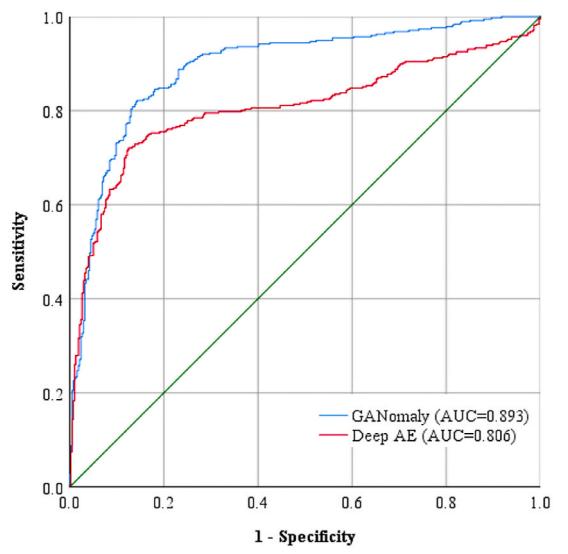
system of the Mesozoic granite province, and the anomaly center contains the world's largest tungsten-tin metallogenic province in the Nanling region (Wang et al., 2007, 2013). This region occupies a pivotal position in terms of tungsten mineral resources in China and the world (Nie and Wang, 2007).

The southern Jiangxi Province is located in the Cathaysia plate within the subduction zone of the Eurasian continental plate and the western Pacific plate and at the intersection of the Wuyishan and Nanling metallogenic belts (Feng et al., 2010) (Fig. 3). A high abundance of tungsten is found in the formations of all ages in this area (Li and Mie, 1999; Nie and Wang, 2007). Combined with the proportion of known tungsten polymetallic deposits (points) in the formation, particularly the contact relationship between tungsten deposits (body) and the formation, the favorable formations for the tungsten formation in southern Jiangxi mainly include (Zeng and He, 2017): (1) the Sinian volcanic sedimentary ferrosilicon clastic rock formation; (2) Cambrian turbidite facies containing charcoal, intercalated carbonate rocks, and sandy argillaceous formation; and (3) the Middle Devonian-Carboniferous formed by clastic rocks and carbonate rocks dominated by neritic deposits. At present, it is believed that the Indosinian and Yanshanian structures have the most obvious effect on the tungsten polymetallic mineralization in the region (Li and Mie, 1999; Xu et al., 2016). In the Indosinian, the structure is dominated by the East-West direction, and the Yanshanian is dominated by a North-Northeast direction. There are obvious differences in the structural characteristics of the different regions in southern Jiangxi. The East-West, North-Northeast, North-East, and Northwest-trending fault zones affect the tectonic pattern of Southern Jiangxi and are closely related to tungsten mineralization. These zones control the distribution and division of ore-forming bodies and related deposits in the region (Zeng and He, 2017). There is a consensus that the quartz vein type wolframite deposits in southern Jiangxi have a close genetic relationship with the Yanshanian granite bodies (Zeng and Tian, 2006). The mineralization ages are very similar to the diagenetic age of the Yanshanian granites. From a spatial perspective, the deposits are distributed around the rock mass, and there are obvious mineralization zones from the center of the rock mass. The main types of ore-bearing rock bodies are biotite, muscovite, and monzonitic granite. Each type of granite is associated with unique mineral deposits, showing a strong mineralization specificity (Nie and Wang, 2007). Superior ore-forming geological conditions have created rich tungsten resources in this area (Feng et al., 2010).

The 1:2000,000 stream sediment geochemical data in the study area (Fig. 4a) were obtained from the National Geochemical Surveying and Mapping Project of China (Xie et al., 1997; Wang et al., 2007). The original stream sediment sampling was collected with an average sampling density of approximate 1 sample/km<sup>2</sup>. Composite samples from 4 km<sup>2</sup> were used with the aim of reducing the number of samples for multi-element analysis. A single moving average method with a 4 km × 4 km or 6 km × 6 km moving window and a 2 km moving step were employed to preprocess the original data for evenly mixing the collected samples data to represent this sampling area (Xie et al., 1997). Previous studies have used principal component analysis, factor analysis, and other approaches, combined with geological background knowledge (such as the mineral deposit models), to study the geochemical evolution process of the study area, explore the combination relationship between elements, and determine the source of matter and the law of element migration (Gong et al., 2005; Wu et al., 2006; Liu et al., 2013). These studies have suggested that five elements, namely W, Sn, Mo, Bi, and Ag, are closely related to tungsten polymetallic deposits. Taking the spatial distribution map of W as an example which was obtained using the inverse distance weighted with a grid of size 1 km × 1 km, the areas linked to high W concentration have a strong spatial association with locations of tungsten polymetallic deposits (Fig. 4b). Therefore, these five geochemical elements were adopted as the input of GANomaly Network to identify multivariate geochemical anomalies related to tungsten polymetallic deposits.



**Fig. 9.** Success-rate curves of the results obtained by GANomaly and AE.



**Fig. 10.** AUC for the geochemical anomaly maps obtained by GANomaly and AE.

#### 4. Results and discussion

The original five geochemical data including W, Sn, Mo, Bi, and Ag were interpolated by the inverse distance weighting with a  $1 \text{ km} \times 1 \text{ km}$  cell size, resulting in five  $336 \times 336$  raster maps. All five raster maps were combined into a set of input feature vectors at each cell location in the set of grids. 30% of the whole data, randomly selected from the five raster maps, were adopted as the training data. The remaining data in this study area were adopted as the testing data for the recognition of geochemical anomalies. The latent feature difference of each sample calculated by Eq. (5) suggest that the data points with high latent feature difference values are considered to be geochemical abnormal points (Fig. 5). Fig. 6 shows the curve of the function value of the objective

function with the number of iterations during the training process. It can be seen that with the iterative update of the network parameters, the objective function value continues to decrease, and finally converges to 0.1, indicating that the network stay relatively stable since then.

Fig. 7 shows the results of geochemical anomalies in the study area identified using GANomaly. We can find that most of the known tungsten polymetallic deposits are located in anomalous high-value areas, which exhibit a larger difference between the potential feature  $z$  and the reconstructed potential feature  $z'$ . The results show that GANomaly can effectively identify multivariate geochemical anomalies related to tungsten polymetallic deposits. In addition, according to the spatial relationship between the geological factors and the anomaly map extracted by GANomaly, it was found that the high-anomaly areas were mostly distributed in or around the Yanshanian intrusion, which provided the heat source and some metals for the formation of tungsten polymetallic deposits. The above evidence shows that GANomaly can effectively identify the geochemical anomalies caused by the mineralization, and the extracted anomaly information can further serve the mineral exploration in this district.

The multivariate geochemical anomaly map produced by GANomaly (Fig. 7) was compared with the anomaly maps obtained by the deep AE (Fig. 8). The results of the comparison method also showed a significant correlation between the known tungsten polymetallic deposits and high-value anomaly areas. However, the results obtained by GANomaly and deep AE show that the former can enhance the ability to extract weak anomaly information compared with the latter, and also extracted some new geochemical anomalies related to tungsten polymetallic mineralization that the latter cannot. In addition, comparing the results of the GANomaly (Fig. 7) with W anomaly map (Fig. 4b), we can find that some anomalous areas which are identified by the GANomaly are not exhibited in Fig. 4b. These new anomalous areas provide important clues for the next round of mineral exploration in the study area.

Similarly, the success-rate was adopted to assess the performance of these two methods. The success-rate, depicted in Fig. 9, shows that 20% of the high-value anomaly areas extracted by GANomaly and deep AE delineated 83.88% and 74.86% of the total number of known tungsten polymetallic deposits, respectively. The ROC curve and AUC were used to further evaluate the anomaly recognition results of the three models.

The ROC curve of GANomaly to identify geochemical anomalies is shown in Fig. 10; the corresponding AUC was 0.893. The AUC values corresponding to deep AE reached 0.806. The comparison results between AE and GANomaly show that the discriminator can improve the ability of the encoder to extract the data features and the ability of the decoder to restore the features. In addition, reconstruction errors between latent features and the reconstructed latent features instead of the reconstruction errors between the original data and the reconstructed data can reduce the interference of noise or other factors in the geochemical data, making the recognition results more realistic, reflecting the objective geological conditions.

## 5. Conclusions

In this study, a method for extracting multivariate geochemical anomalies that require only non-mineralized samples during the training phase can avoid the overfitting problem caused by the imbalance between positive and negative training samples. By comparing with the deep autoencoder, the GANomaly network combining a GAN, an AE, and a CNN has the advantages of learning the spatial characteristics of the data and reducing the influence of noisy data on the extraction of abnormal information. It was found that the anomaly information extracted by the GANomaly network had a close spatial correlation with the known tungsten deposits in the study area, and also extracted many newly geochemical anomalies related to tungsten mineralization. This shows that the GANomaly network can be applied to the processing of multivariate geochemical data, providing a reference for further exploration within the study area.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We thank Dr. Mohammad Parsa and an anonymous reviewer's comments and suggestions which help us improve this study. This study was supported by the National Natural Science Foundation of China (No. 41772344).

## References

- Agterberg, F.P., Bonham-Carter, G.F., 2005. Measuring the performance of mineral potential maps. *Nat. Resour. Res.* 14 (1), 1–17.
- Akcay, S., Atapour-Abarghouei, A., Breckon, T.P., 2018. GANomaly: semi-supervised anomaly detection via adversarial training. Asian Conference on Computer Vision 622–637.
- Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein gan arXiv preprint arXiv: 1701.07875.
- Beggel, L., Pfeiffer, M., Bischl, B., 2019. Robust anomaly detection in images using adversarial autoencoders. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Cham, pp. 206–222.
- Chen, L., Guo, K., Liu, B., Liang, L., 2012. The Study of non-linear analysis method of Geochemical ore-forming anomaly. *Prog. Geophys.* 1701–1707 (In Chinese with English abstract).
- Chen, L., Guan, Q., Xiong, Y., Liang, J., Wang, Y., Xu, Y., 2019. A Spatially Constrained Multi-Autoencoder approach for multivariate geochemical anomaly recognition. *Comput. Geosci.* 125, 43–54.
- Cheng, Q., 2007. Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China. *Ore Geol. Rev.* 32, 314–324.
- Dai, J., Wang, J., Zhu, Z., Shen, C., Huang, W., 2019. Anomaly detection of mechanical systems based on generative adversarial network and auto-encoder. *Chin. J. Sci. Instrum.* 40, 16–26 (In Chinese with English abstract).
- Denton, E., Chintala, S., Szlam, A., Fergus, R., 2015. Deep generative image models using a laplacian pyramid of adversarial networks. *Adv. Neural Inf. Process. Syst.* 1486–1494 arXiv preprint arXiv:1506.05751.
- Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recogn. Lett.* 27 (8), 861–874.
- Feng, C., Zeng, Z., Qu, W., Zhang, D., Wang, S., 2010. A discussion on the chronology of tungsten mineralization and the time difference of diagenesis and mineralization in southern Jiangxi. *Miner. Deposits (S1)*, 431–432 (In Chinese with English abstract).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 2672–2680 arXiv preprint arXiv:1406.2661.
- Gong, J., Li, F., Zhang, S., Cui, F., 2005. Delineating anomalies using similarity coefficients based on element assemblage characteristics: an example of the Nanling area. *Geology and Exploration* (2), 312–322 (In Chinese with English abstract).
- Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition 1125–1134.
- Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization arXiv preprint arXiv:1412.6980.
- Larsen, A.B.L., Sonderby, S.K., Larochelle, H., Winther, O., 2016. Autoencoding beyond pixels using a learned similarity metric. In: International Conference on Machine Learning, vol. 48, pp. 1558–1566 arXiv preprint arXiv:1512.09300.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W., 2017. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4681–4690.
- Li, T., Zuo, R., Xiong, Y., Peng, Y., 2021. Random-drop data augmentation of deep convolutional neural network for mineral prospectivity mapping. *Nat. Resour. Res.* 30 (1), 27–38.
- Li, Z., Mie, Y., 1999. The distribution law of tungsten deposits in southern Jiangxi and the prediction of hidden deposits. *Jiangxi Geology* (4), 276–282 (In Chinese with English abstract).
- Liu, Y., Xia, Q., Cheng, Q., Wang, X., 2013. Application of singularity theory and logistic regression model for tungsten polymetallic potential mapping. *Nonlinear Process Geophys.* 20, 445–453.
- Luo, Z., Xiong, Y., Zuo, R., 2020. Recognition of geochemical anomalies using a deep variational autoencoder network. *Appl. Geochem.* 122, 104710.
- Longadge, R., Dongre, S., 2013. Class imbalance problem in data mining review arXiv preprint arXiv:1305.1707.
- Merkle, N., Auer, S., Müller, R., Reinartz, P., 2018. Exploring the potential of conditional adversarial networks for optical and SAR image matching. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11 (6), 1811–1820.
- Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets arXiv preprint arXiv:1411.1784.
- Nie, R., Wang, X., 2007. Research progress of tungsten deposits in southern Jiangxi. *China Tungsten Industry* 1–5 (In Chinese with English abstract).
- Oliveira, D.A., Ferreira, R.S., Silva, R., Brazil, E.V., 2018. Interpolating seismic data with conditional generative adversarial networks. *Geosci. Rem. Sens. Lett. IEEE* 15, 1952–1956.
- Parsa, M., 2021. A data augmentation approach to XGboost-based mineral potential mapping: an example of carbonate-hosted ZnPb mineral systems of Western Iran. *J. Geochem. Explor.* 106811.
- Parsa, M., Maghsoudi, A., Yousefi, M., 2018. A receiver operating characteristics-based geochemical data fusion technique for targeting undiscovered mineral deposits. *Nat. Resour. Res.* 27, 15–28.
- Parsa, M., Maghsoudi, A., Yousefi, M., Sadeghi, M., 2016. Recognition of significant multi-element geochemical signatures of porphyry Cu deposits in Noghdouz area, NW Iran. *J. Geochem. Explor.* 165, 111–124.
- Parsa, M., Maghsoudi, A., Yousefi, M., Sadeghi, M., 2017. Multifractal analysis of stream sediment geochemical data: implications for hydrothermal nickel prospection in an arid terrain, eastern Iran. *J. Geochem. Explor.* 181, 305–317.
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks arXiv preprint arXiv:1511.06434.
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., 2016. Improved techniques for training GANs. *Adv. Neural Inf. Process. Syst.* 2234–2242 arXiv preprint arXiv:1606.03498v1.
- Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G., 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. International Conference on Information Processing in Medical Imaging 146–157.
- Taigman, Y., Polyak, A., Wolf, L., 2016. Unsupervised cross-domain image generation arXiv preprint arXiv:1611.02200.
- Wang, X., Xu, S., Chi, Q., Liu, X., Wang, W., 2013. Mass accumulation and distribution of metallogenic elements in the South China block. *Geochimica* 42 (3), 229–241 (In Chinese with English abstract).
- Wang, X., Zhang, Q., Zhou, G., 2007. Nationalscale geochemical mapping projects in China. *Geostand. Geanal. Res.* 31 (4), 311–320.
- Wu, X., Gong, Q., Xiang, Y., 2006. Discussion on the method of extracting tungsten ore information from regional geochemical exploration data in Nanling area. *Computing Technology for Geophysical and Geochemical Exploration* (2), 182–186 (In Chinese with English abstract).
- Xie, X., Mu, X., Ren, T., 1997. Geochemical mapping in China. *J. Geochem. Explor.* 60, 99–113.
- Xiong, Y., Zuo, R., 2016. Recognition of geochemical anomalies using a deep autoencoder network. *Comput. Geosci.* 86, 75–82.
- Xiong, Y., Zuo, R., 2018. GIS-based rare events logistic regression for mineral prospectivity mapping. *Comput. Geosci.* 111, 18–25.
- Xu, F., Li, B., Li, B., Wang, Q., Gan, K., Deng, D., 2016. Classification and metallogenic fluid characteristics of the tungsten deposits in southern Jiangxi Province. *Geol. Resour.* (4), 339–344 (In Chinese with English abstract).

- Zenati, H., Foo, C.S., Lecouat, B., Manek, G., Chandrasekhar, V.R., 2018. Efficient gan-based anomaly detection arXiv preprint arXiv:1802.06222.
- Zeng, Z., He, G., 2017. The division of tungsten mineralization units in southern Jiangxi and their mineralization geological characteristics. In: Geological Society of Jiangxi Province. (eds.) 2016 Collection of Papers of the Geological Society of Jiangxi Province III, pp. 31–38 (In Chinese with English abstract).
- Zeng, Z., Tian, Y., 2006. A review of the history of tungsten deposit prospecting in southern Jiangxi and a new round of tungsten prospecting thinking. Resource Survey & Environment (4), 339–344 (In Chinese with English abstract).
- Zhang, C., Zuo, R., Xiong, Y., 2021. Detection of the multivariate geochemical anomalies associated with mineralization using a deep convolutional neural network and a pixel-pair feature method. Appl. Geochem. 130, 104994.
- Zuo, R., Xiong, Y., Wang, J., Carranza, E.J.M., 2019. Deep learning and its application in geochemical mapping. Earth Sci. Rev. 192, 1–14.