https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

Recognizing geochemical spatial patterns using deformable convolutional networks guided with geological knowledge

Xinyu Zhang, Yihui Xiong, Zhiyi Chen

State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China

Corresponding author. Email: Yihui Xiong (xiongyh426@cug.edu.cn)

Abstract. This study tackles the limited quantification of irregular spatial geochemical patterns and weak interpretability in deep learning models in geochemical anomaly recognition. We propose a hybrid approach that that integrates geological knowledge (GK) into deformable convolutional networks (DCN), creating a model termed GK_DCN, with the aim of enhancing both the performance and transparency of geochemical anomaly recognition. This model introduces learnable parameters that allow the convolutional kernels to adaptively adjust their sampling locations, enabling them to more accurately capture complex, irregular geochemical anomaly patterns caused by mineralization. To enhance geological consistency, ore-controlling fault are incorporated as geological knowledge constraints, guiding the network to prioritize spatial correlations between deposits and faults. Experimental results in southern Tianshan Au-Cu polymetallic ore district demonstrate that the GK_DCN significantly enhances the accuracy and reliability of geochemical anomaly recognition verified across multiple evaluation metrics, producing more distinct spatial anomalous patterns and higher consistency with known mineral deposits by adaptively adjusting the receptive field. Visualization of the kernel offsets revealed the model's superior adaptive spatial sampling mechanism. Furthermore, using Grad-CAM to generate feature significance heatmaps highlighted the key features the model focused on during geochemical anomaly recognition, significantly improving interpretability and proving effectiveness in capturing complex geochemical patterns. This work provides an effective intelligent method for geochemical pattern recognition and offers a reference for interpretable deep learning in geochemical exploration through multi-angle visualization.

1. Introduction

Geo-anomalies, detected through various observational datasets such as geological, geochemical, geophysical, and remote sensing methods, play a vital role in identifying mineralization-related geological processes. Their significance lies in the fact that these anomalies reveal underlying causative features or events that are not directly observable (Cheng and Zhao, 2011). Hydrothermal mineralization is a systematic yet complex geological phenomenon, involving the movement of orebearing hydrothermal fluids, interactions between fluids and host rocks, mineral precipitation, and the eventual concentration of ore materials (Pirajno, 2008). Since these processes result from the interplay of multiple geological factors operating across different spatial and temporal scales, the associated geo-anomalies display considerable complexity (Cheng, 2012). Analyzing their spatial distribution in detail can significantly refine and advance geological understanding of numerous scientific questions. Geochemical anomalies associated with mineralization represent one of the most significant types of geo-anomalies for mineral exploration (Zuo et al., 2021). These anomalies often exhibit anisotropic spatial distributions that are controlled by ore-forming geological structures-such as strata, faults, folds, and magmatic intrusions-which provide essential space, heat, fluid, and material conditions required for mineralization (Pirajno, 2008). For example, hydrothermal mineralization frequently presents as linear-trending geochemical anomalies along fault zones, where fault systems act as pathways for the transport and deposition of ore-forming materials (Wang et al., 2013). Consequently, recognizing the spatial anisotropy of geochemical patterns is crucial for accurately identifying significant anomalies, thus can greatly enhance the success of mineral exploration (Cheng, 2012; Zuo, 2017; Xiao et al., 2018).

Long-term research and practice have demonstrated that integrating spatial structure through geostatistics, spatial autocorrelation analysis, spatial decomposition, moving window statistics, and spatially aware machine learning offers a more geologically realistic and robust framework for recognizing geochemical anomalies. Geostatistical techniques, such as kriging, allow for the estimation of values at unsampled locations, generating a spatially continuous model of the geochemical background. Anomalies are identified where measured values significantly exceed the kriging predictions, indicated by large prediction errors (Jimenez-Espinosa et al., 1993). Additionally, incorporating directional variograms into kriging methods, such as anisotropic ordinary kriging, enables explicit accounting for directional trends and local heterogeneity in anomalies (Reis et al., 2003). Spatial autocorrelation analysis, using methods like local Moran's I, helps detect statistically significant spatial clusters-such as high-high clusters (indicating potential anomalies surrounded by other high values) and high-low clusters (representing isolated high values) (Yin et al., 2021). Spatial decomposition primarily employs two categories of methods: trend surface analysis and multifractal filtering. Trend surface analysis involves fitting polynomial surfaces, using either global or local regression, to model regional trends. The residuals derived from this surface represent local deviations, which can serve to highlight anomalies against the broader regional pattern (Wang and Zuo, 2015). Multifractal filtering methods mainly include the Concentration-Area (C-A) model and the Spectrum-Area (S-A) model (Cheng et al., 1994, 2000). These methods plot element concentration against area and identify breaks in the observed power-law (scaling) behavior. These breakpoints are used to separate background populations from anomalous ones. This approach explicitly models the scale-dependent heterogeneity of spatial patterns

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

and establishes thresholds based on deviations from fractal behavior across different scales (Cheng, 2012). Moving window statistics methods, such as local singularity analysis and the local gap statistic, calculate local statistics (e.g., mean, median, standard deviation) within a defined spatial window (Cheng, 2007; Wang and Zuo, 2016). Values that significantly exceed the local background within their respective window are identified as geochemical anomalies. This technique effectively captures local spatial context and non-stationarity, although the choice of window size is a critical and subjective step. To account for anisotropy, these methods are often adapted by incorporating elliptical or directionally weighted windows (Xiao et al., 2018, 2020; Wang et al., 2018). The final category for analyzing geochemical spatial patterns is spatially aware machine learning. This approach primarily includes two types: models that integrate spatial features or components (Cheng et al., 2011; Wang et al., 2015), and models with inherent capabilities to capture spatial structures (LeCun and Bengio, 1998). In the first type, spatial characteristics are incorporated into traditional statistical methods-often through distance-based kernels or spatial weighting schemes-to address geographic heterogeneity and non-stationarity, where variable relationships vary across space. Commonly applied spatially weighted machine learning techniques for identifying geochemical anomaly patterns include geographically weighted regression (GWR) (Wang et al., 2015; Tian et al., 2018), spatially weighted principal component analysis (SWPCA) (Cheng et al., 2011; Xiao et al., 2012), density-based spatial clustering of applications with noise (DBSCAN) (Zhang et al., 2019; Hajihosseinlou et al., 2024), and geographical random forest (GRF) (Soltani et al., 2024). The second type involves machine learning architectures specifically designed to handle spatial data, such as convolutional neural networks (CNN) (LeCun and Bengio, 1998) and graph neural networks (GNN) (Scarselli et al., 2008). CNN learns local spatial features and mineralization-related patterns through convolutional and pooling operations. However, a key limitation is their reliance on fixed, regular convolution kernels (e.g., 3×3 grids), which restricts their ability to adequately model the anisotropic nature of geochemical distributions (Dai et al., 2017). In contrast, GNN directly represents non-Euclidean spatial relationships using nodes (e.g., sample points with geochemical attributes) and edges (encoding spatial proximity or geological links), allowing anomaly detection based on complex neighborhood interactions (Xu et al., 2023, 2024, 2025; Chen et al., 2025). Nevertheless, GNN requires high-quality data and substantial domain knowledge to define meaningful graph structures. In particular, defining appropriate edges-based on spatial distance or geological similarity-is crucial yet challenging. Improper edge definitions may introduce noise, mask genuine anomalies, and ultimately impair model performance (Gong and Cheng, 2019; Zhou et al., 2020).

Deformable Convolutional Networks (DCN) address a fundamental constraint of traditional CNN: the fixed geometric structure of their convolution kernels (Dai et al., 2017). By introducing learnable spatial offsets for each sampling point in the kernel, DCN adaptively adjust the sampling locations, effectively warping the kernel's receptive field to align with

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

irregular and complex patterns (Dai et al., 2017; Zhu et al., 2018). This flexibility allows the kernel to conform to non-rigid and deformed structures, enabling more precise feature extraction from key regions of irregular shapes (Dai et al., 2017; Zhu et al., 2019). As a result, DCNs exhibit greater robustness to geometric variations such as changes in orientation, scale, or deformation, maintaining consistent feature representation across diverse pattern states. These capabilities make DCN especially valuable in tasks involving irregular spatial structures, where they significantly improve recognition and quantification performance. By offering essential spatial adaptability, DCN provide a powerful tool for analyzing the complex and often messy geometries encountered in real-world data across various domains, which include irregular seismic data interpolation (Zhao et al., 2023; Luo et al., 2024; Sun et al., 2024), earthquake crack detection (Yu et al., 2022), flood boundary detection (Yu et al., 2023), surface wave suppression (Gao et al., 2024), underwater image enhancement (Tian et al., 2023), atmospheric forecasting (Nielsen et al., 2022), precipitation forecasting (Xu et al., 2024), morphological characteristics of clouds modelling (Liu et al., 2021), images denoising (Guan et al., 2022; Liu et al., 2024), hyperspectral image classification (Zhu et al., 2018; Zhao et al., 2021), identification of anomalous deformation areas (Zhang et al., 2022), hyperspectral anomaly detection (Wu et al., 2023), soil moisture monitoring (Na et al., 2025). By capturing nuanced spatial deformations, DCN offer a transformative approach for extracting meaningful metrics from the inherent irregularity of geoscientific data.

In this study, we utilize a DCN as the foundational model for recognizing and extracting complex anisotropic geochemical spatial patterns. Just as purely data-driven deep learning methods such as CNN face interpretability issues, so too does the DCN, whose function is regarded as complex "black boxes". While they achieve high prediction accuracy, understanding why they make a specific prediction, which features in the input data were decisive, or how their learned representations map to established geological concepts is extremely difficult (Rudin, 2019; Gilpin et al., 2018). For instance, especially for DCN, are the learned offsets geologically meaningful, or are they exploiting subtle. Current approaches to enhance the interpretability of deep learning models primarily operate at three levels: model input, model construction, and model output (Zuo et al., 2024). At the model input level, interpretability is enhanced through metallogenic models, feature engineering, and geologically constrained data augmentation methods (Zuo et al., 2024). At the model construction stage, key ore-controlling factors are integrated into the hidden layers, while the spatial coupling relationship between known mineral deposit locations and these factors is incorporated into the loss function (Xiong et al., 2022; Luo et al., 2023; Zuo et al., 2025). At the model output stage, visualization techniques are employed to examine the outputs of each hidden layer, providing insight into the extraction and integration processes of prospecting information. Meanwhile, attribution techniques are applied to assess the importance of input variables, helping to quantify their contributions to the formation of mineral deposits (Luo et al., 2023; Xu et al., 2025). In this study, we enhance the

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

interpretability of the DCN at both the model construction and output levels. During model construction, a governing equation representing the spatial correlation between known mineral deposits and ore-controlling factors is embedded into the loss function (Xiong et al., 2022; Zuo et al., 2024). This approach introduces conceptual models and expert knowledge into the training process, ensuring that the model's outputs are consistent with established geological principles (Zuo et al., 2024). At the model output stage, we utilize class activation mapping (CAM) (Jung and Oh, 2021) and its variant, Grad-CAM (Selvaraju et al., 2016), to visualize the regions within the input data that most influence the model's predictions. CAM visually identifies the most discriminative regions in an input image responsible for a specific class prediction of CNN and its variants (e.g., DCN). It leverages the weights of the final fully connected layer to compute a weighted sum of the activation maps from the last convolutional layer, thus can transform CNN and its variants from a "black box" into a more transparent model by generating a heatmap (class activation map). Besides, the learned offsets are also visualized to reveal how DCN dynamically adapts sampling locations, enhancing understanding of model behavior for spatial pattern quantification. Ultimately, the constructed model was applied to the study area of the southern Tianshan Au-Cu polymetallic ore district to verify its effectiveness and interpretability in identifying geochemical anomalies.

2. Geological setting and Datasets

2.1. Geological setting

The South Tianshan Metallogenic Belt, extending across Central Asia from Uzbekistan through Tajikistan, Kyrgyzstan, and into western China (Xinjiang), is one of the world's most significant gold and copper provinces (Fig. 1). Its formation is intrinsically linked to the protracted and complex tectonic history of the Central Asian Orogenic Belt, specifically the final closure of the Paleo-Asian Ocean (Gao et al., 2009; Han et al., 2011). The regional geology is dominated by the collage of multiple terranes, including Precambrian continental blocks, early Paleozoic oceanic crust fragments, and island arcs, which were accreted and subsequently deformed during the Late Paleozoic collision between the Tarim Craton to the south and the Kazakhstan-Yili Block to the north (Gao et al., 2009). This continental collision, culminating in the Late Carboniferous to Early Permian, created a major suture zone characterized by extensive thrusting, folding, and large-scale strike-slip fault systems. These structures provided crucial conduits for subsequent fluid migration and mineralization.

The primary mineralization events are temporally and genetically associated with this collisional orogeny and the post-

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

collisional extensional phase. Two major mineralization styles prevail: (1) Orogenic gold deposits, often hosted in shear zones within Neoproterozoic to Paleozoic metamorphic rocks (e.g., the giant Muruntau deposit in Uzbekistan). These deposits formed from metamorphic fluids released during devolatilization of subducted slabs or thickened crust. (2) Copper-gold skarn and porphyry-style mineralization, frequently associated with Late Carboniferous to Permian post-collisional I-type granitoids intruding carbonate-rich sequences. These intrusions provided the heat and magmatic fluids responsible for widespread hydrothermal alteration and metal deposition. The conjunction of fertile source rocks (often black shales), ideal structural traps (fault jogs, shear zones, lithological contacts), and the timing of magmatism relative to tectonic stress changes created the perfect conditions for the formation of world-class gold and copper deposits. The Chinese segment of the South Tianshan, such as the Sawayaerdun gold belt, continues this metallogenic trend, hosting numerous deposits with similar genetic models (Chen et al., 2012; Goldfarb et al., 2014; Seltmann et al., 2014).

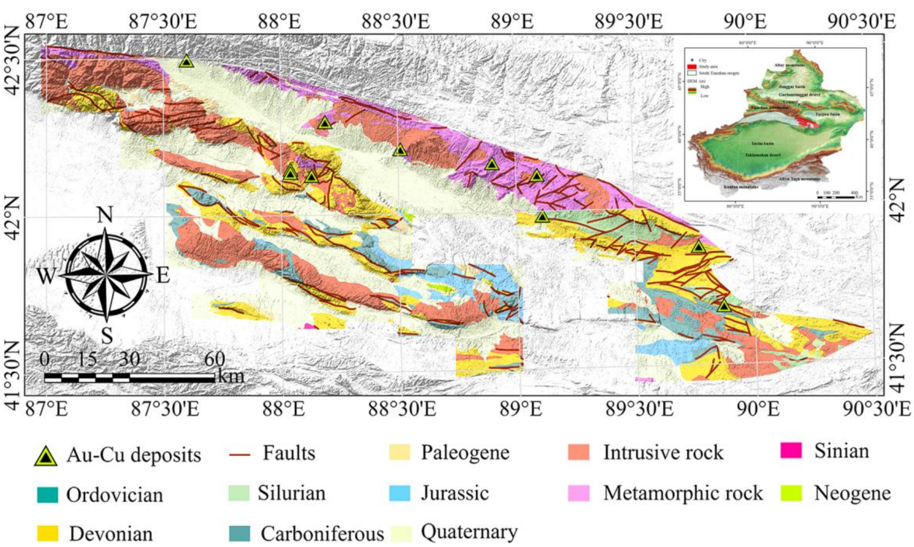


Figure 1: Simplified geological map of the Southern Tianshan showing the main tectonic units and Au-Cu deposits (modified from Xue et al. (2014); Zhao et al. (2020)).

2.2. Datasets

The 1:200,000 scale geochemical samples in this study area were sourced from the Chinese national geochemical mapping project (Xie et al., 1997). The standard sampling density was 1-2 samples per square kilometer, with every 4 km² constituting one analytical unit. Sampling density was appropriately reduced in areas where fieldwork was difficult to conduct (1 sample per 20–50 km2). Multiple sub-samples were collected within a certain range (20-50 m) around the

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

sampling point and combined into a single composite sample. The sample was sieved through a 60-mesh stainless steel screen, with the final sample weight exceeding 200g. A total of 32 elements and 7 oxides were analyzed: Bi, Cu, P, La, Li, Ag, Sn, Au, Mo, Th, U, Y, W, Sb, Hg, Mn, Cr, Sr, Nb, Pb, Ni, Ti, Cd, Co, Ba, Be, V, Zn, B, As, Zr, F, as well as Fe_2O_3, K_2O, CaO, MgO, Na_2O, Al_2O_3, and SiO_2. The detection limits and analytical methods for each element are listed in Table 1.

Table 1 Elements, analytical methods, and detection limits from the Chinese national geochemical mapping project

Elements	Unit	Detection Limit	Analytical Method	Elements	Unit	Detection Limit	Analytical Method
Ag	ng/g	0.02	ES	**Pb**	μg/g	2	ICP–MS
As	μg/g	1	HG-AFS	**Sb**	μg/g	0.1	HG-AFS
Au	ng/g	0.0003	GF-AAS	**Sn**	μg/g	1	ES
B	μg/g	5	ES	**Sr**	μg/g	5	ICP-AES
Ba	μg/g	50	ICP-AES	**Th**	μg/g	4	ICP–MS
Be	μg/g	0.5	ICP-AES	**Ti**	μg/g	100	XRF
Bi	μg/g	0.1	ICP–MS	**U**	μg/g	0.5	ICP–MS
Cd	ng/g	0.05	ICP–MS				
Co	μg/g	1	ICP–MS	**V**	μg/g	20	ICP-AES
Cr	μg/g	15	XRF	**W**	μg/g	0.5	ICP–MS
Cu	μg/g	1	ICP–MS	**Y**	μg/g	5	XRF
F	μg/g	100	ISE	**Zn**	μg/g	10	ICP-AES
Hg	μg/g	0.0005	CV-AFS	**Zr**	μg/g	10	XRF
La	μg/g	30	ICP–MS	**Al_2O_3**	%	0.05	XRF
Li	μg/g	5	ICP-AES	**CaO**	%	0.05	ICP-AES
Mn	μg/g	30	ICP-AES	**Fe_2O_3**	%	0.05	XRF
Mo	μg/g	0.4	ICP–MS	**K_2O**	%	0.05	XRF
Nb	μg/g	5	ICP–MS	**MgO**	%	0.05	ICP-AES
Ni	μg/g	2	ICP-AES	**Na_2O**	%	0.05	ICP-AES
P	μg/g	100	XRF	**SiO_2**	%	0.1	XRF

Note: XRF: X-ray fluorescence spectrometry; ICP-AES: Inductively coupled plasma-atomic emission spectrometry; ICP–MS: Inductively coupled plasma–mass spectrometry; ES: Emission spectrometry; HG-AFS: Hydride generation atomic fluorescence spectrometry; GF-AAS: Graphite furnace atomic absorption spectrometry; CV-AFS: Cold vapor atomic fluorescence spectroscopy; ISE: Ion selective electrode.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

3. Methods

3.1. Deformable convolutional networks (DCN)

Deformable convolution (Dai et al., 2017; Zhu et al., 2019), enables adaptive adjustment of the receptive field positions by incorporating learnable offset parameters for each sampling point within the convolutional kernel. Figure 2 illustrates the distinction between the sampling points of standard convolution and those of deformable convolution. This approach overcomes the limitations imposed by a fixed grid structure, thereby facilitating more flexible and precise extraction of image features exhibiting complex geometric deformations.

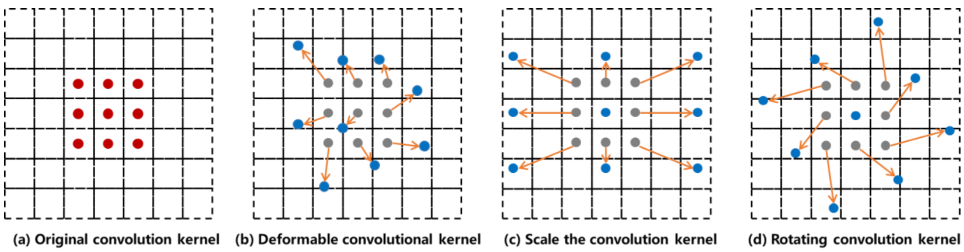


Figure 2: Illustration of the sampling locations in 3 × 3 standard and deformable convolutions. (a) Regular sampling grid of standard convolution; (b–d) deformed sampling locations of deformable convolution with augmented offsets. The red areas are the sampling locations in 3 × 3 standard convolution. The grey areas and the blue areas are the initial sampling locations and final sampling locations of the deformable convolution, respectively. The yellow arrow points from the initial sampling location to the corresponding final sampling location.

The computation involved in deformable convolution remains a form of two-dimensional convolution, with an emphasis on spatial interactions across all channels. The fundamental aspect of this method lies in learning the offsets of sampling points via a parallel branch network, allowing the convolutional kernel to dynamically adjust its sampling locations based on the content of the input feature map. This mechanism directs convolutional operations to concentrate on regions of interest, substantially enhancing the network's capacity to represent features associated with geometric transformations. In this study, a standard 3×3 two-dimensional convolutional kernel, denoted as R, is employed as an illustrative example.

$R = (-1,-1), (-1,0), \cdots, (0,1), (1,1), (1)$

In conventional convolutional kernels, the weight matrix is denoted by w, the input feature map by x, and p_n represents any pixel within the convolutional window R. For each output position p_0 in the feature map, the convolution operation can be mathematically expressed as follows:

$$y(p_0) = \sum_{p_n \in R} w\,(p_n) \mathrm{x}(p_0 + p_n), (2)$$

In the context of deformable convolution, the introduction of an offset $\Delta p_n \{\Delta p_n \mid n = 1, \cdots, N\} N = |R|$ modifies the original formulation, transforming it into:

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

$$y(p_0) = \sum_{p_{\mathrm{n}} \in R} w\,(p_{\mathrm{n}}) \times (p_0 + p_{\mathrm{n}} + \Delta p_n), (3)$$

This adjustment results in sampling points that are spatially shifted, with the offset positions denoted as $p_{\mathrm{n}} + \Delta p_n$. Since the offset Δp_n generally assumes non-integer values, the computation of the convolution must be performed using bilinear interpolation, as described by:

$$x(p) = \sum G\,(q, p) x(q), (4)$$

The value at any position p is thus a function of $p = p_0 + p_{\mathrm{n}} + \Delta p_n$ and is computed over all spatial locations q within the input feature map x by employing the bilinear interpolation kernel $G(\cdot,\cdot)$. Notably, the two-dimensional interpolation kernel is separable and can be decomposed into the product of two one-dimensional kernels, which serves to optimize computational efficiency:

$$G(q, p) = g(q_x, p_x) g\big(q_y, p_y\big), (5)$$

Where $g(a, b) = max(\,0{,}1 - |a - b|)$.

Figure 3 delineates the detailed implementation procedure of deformable convolutional layers. Initially, the learned offset vectors are applied to the fixed sampling grid of the input feature map, enabling adaptive adjustment of each sampling point's position. Subsequently, bilinear interpolation is utilized to estimate feature values at the offset, non-integer coordinate locations, thereby ensuring that the sampled feature distribution effectively concentrates on the target region. Figure 4 provides a comparative visualization between standard convolution and deformable convolution with respect to their receptive fields for geochemical pattern recognition. By incorporating offsets, the receptive field in deformable convolution transcends the constraints imposed by the fixed, regular grid of standard convolution. This flexibility allows the receptive field to adaptively assume irregular spatial configurations that better correspond to the actual geometric structure of the target object, thereby substantially enhancing the accuracy of feature extraction.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

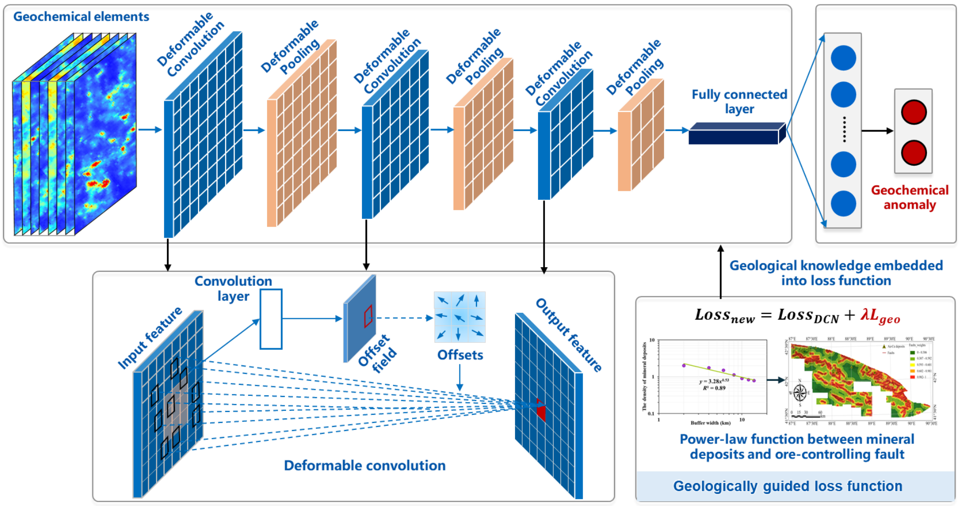


Figure 3: The framework of proposed geological knowledge guided deformable convolution networks.

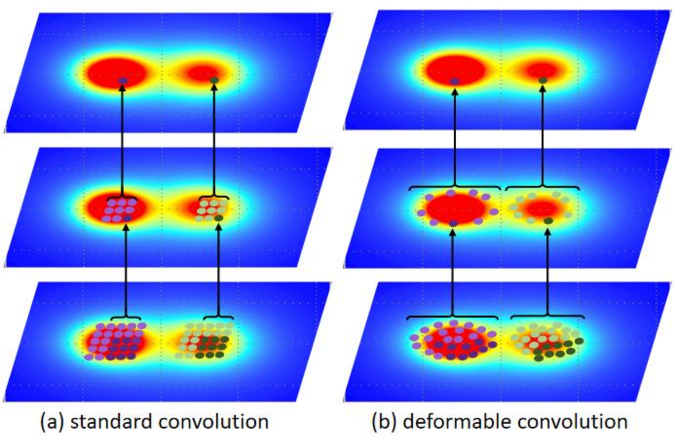


Figure 4: Illustration of the sampling locations for (a) normal convolution and (b) deformable convolution. Maps showing irregular geochemical patterns. It is observed that deformable convolutions can adaptively extracts the features of the input by adjusting its shape according to the actual patterns by shifting the convolutional kernel, but normal convolutions only describe the fixed receptive field.

3.2. Geologically-constrained DCN

This study introduces soft constraints on deformable convolutional networks to enhance geochemical anomaly detection by incorporating geological prior knowledge. In typical geochemical anomaly recognition tasks, deformable convolutional neural networks optimize their parameters by minimizing the cross-entropy loss, which measures the divergence between predicted and true label distributions. To improve this optimization process, the present work

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

augments the loss function with an additional penalty term derived from established geological principles, thereby guiding the model to learn feature representations that better conform to geological laws (Semenov et al., 2019).

The conventional loss function L_{DCN}, is defined as follows:

$$L_{DCN}(p,\hat{p}) = -\sum_{x} p\,(x) \log \hat{p}\,(x), (6)$$

Where$p(x)$ and the predicted distribution $\hat{p}(x)$ denote the true and predicted distributions, respectively. Building upon this, a novel penalty term grounded in geological knowledge is formulated and integrated into the loss function. Following the approach proposed by Zuo (2016) for penalty term construction, the relationship between the distance control factor and the spatial distribution of mineral deposits is modeled by a power-law function w, expressed as:

$$w = \frac{m}{m_{\max}} = \frac{Nd^k}{m_{\max}}, (7)$$

Here, N is a constant, d represents the distance between the control factor and the mineral deposit, m denotes the density of mineral points at distance d, and k corresponds to the line fitting parameters relating log^m and log^d. Here, w serves as a control equation embedding prior geological knowledge to characterize the spatial coupling between known mineral occurrences and their controlling factors. As the DCN progressively learns the spatial distribution patterns between mineral deposits and their surrounding grid units, it becomes essential to extract spatial structural features encapsulated by the weight function w and incorporate them into the training process. Consequently, a geology-informed penalty term L_{geology} is constructed, formulated as:

$$L_{geology} = ||f_{softmax}(\hat{p}(x)) - f_{sigmoid}(\sum_{\mathrm{i=1}}^{\mathrm{n}} a\, w_i + b)||_2, (8)$$

In this expression, a and b are trainable parameters within the combined kernel used for feature aggregation, where a represents weights and b denotes bias terms; n is the number of feature maps. The aggregated features undergo normalization via the function f_{sigmod}, and the network output is subsequently transformed into mineral potential prediction values through the mapping function $f_{softmax}$.

Finally, a total loss function L_{total} was constructed in the variable convolution that integrates prior geological knowledge, and its expression is as follows:

Ultimately, a comprehensive loss function L_{total} is developed for the deformable convolutional network, integrating prior geological knowledge, and is expressed as:

$$L_{total} = L_{DCN} + \lambda L_{geology}, (9)$$

This formulation effectively constrains the model to produce predictions that are consistent with both data-driven learning and established geological understanding (Fig. 3).

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

3.3. Gradient-weighted Class Activation Mapping

To improve the interpretability and discriminative localization capabilities of deformable convolutional neural networks, the integration of class activation mapping (CAM) techniques can be employed. The conventional CAM approach leverages the weights from the global average pooling (GAP) layer and the final classification layer to visualize the discriminative regions utilized by the CNN during classification. By projecting the output layer's weights back onto the convolutional feature maps, the relative importance of different image regions can be identified.

Initially, it is necessary to remove all fully connected layers following the last convolutional block, as CAM requires a fully convolutional architecture to maintain spatial information up to the final layer. A GAP layer is introduced subsequent to the last deformable convolutional layer to substitute the fully connected layers (Jung and Oh, 2021). The function of this GAP layer is to compute the spatial average value F^k of each feature map in the final convolutional layer, which can be mathematically expressed as:

$$F_k = \frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} f_k\,(x, y), (10)$$

where $f_k(x, y)$ denotes the activation at spatial location (x, y) in the k-th channel of the feature map output by the last deformable convolutional layer, and X and Y represent the width and height of the feature map, respectively.

Following the GAP layer, a single fully connected layer with a softmax activation function is appended. For a given class c, this layer assigns a weight w_k^c to each averaged feature map value $f_k(x, y)$. The linear classification logit score S_c for class c is then computed as:

$$S_c = \sum_k w_k^c\, F_k = \frac{1}{X \cdot Y} \sum_{x,y} \sum_k w_k^c\, f_k(x, y), (11)$$

Here, S_c is a scalar representing the classification score. To generate the class activation map, the weights w_k^c are multiplied element-wise with the corresponding feature maps F^k and summed across all channels:

$$M_c(x, y) = \sum_k w_k^c\, f_k(x, y), (12)$$

This operation preserves spatial information along the width and height dimensions. Subsequently, bilinear interpolation is applied to upsample the matrix M_c to the original input image size, thereby producing the complete CAM visualization. In summary, each feature map channel corresponds to a specific class of visual features extracted by a convolutional kernel from the input image. The weights w_k^c implicitly indicate the significance of these features for the classification of category c, reflecting the degree of attention that the model allocates to each feature with respect to that class.

However, CAM technique necessitates the substitution of the fully connected layer with a GAP layer and is limited to analyzing only the final convolutional layer. To overcome these constraints, we adopted the Gradient-weighted CAM

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

(Grad-CAM) approach, which derives the requisite weights indirectly through gradient computations rather than depending on the GAP layer and softmax activation (Selvaraju et al., 2016). This method can be applied to a wide range of contemporary models incorporating deformable convolutional layers without modifying the existing network architecture or requiring retraining. Consequently, it enables the generation of class-specific activation heatmaps for convolutional layers situated at various depths within the network.

The Grad-CAM algorithm involves computing the gradient of the target score—typically corresponding to the class of interest—with respect to the feature maps of a selected convolutional layer. From these gradients, the importance weight α_k^c for each channel k is obtained, as expressed by the following equation:

$$\alpha_k^c = \frac{1}{Z}\sum_{i=1}^{u}\sum_{j=1}^{v}\frac{\partial y^c}{\partial A_{ij}^k}, (13)$$

Here, c denotes the target class, α_k^c represents the weight of the k-th channel for class c, and y^c is the linear classification logit score for class c. The partial derivative $\frac{\partial y^c}{\partial A_{ij}^k}$ corresponds to the sensitivity of the output score y^c with respect to the activation at spatial location (i, j) in the k-th feature map, where u and v indicate the width and height of the feature map, respectively.

Mathematically, the weight α_k^c serves a role analogous to the weight w_k^c in the original CAM formulation. By linearly combining these weights with the corresponding feature maps, the class activation map M_c can be computed as follows:

$$L_{Grad-CAM}^c = ReLU\left(\sum_k \alpha_k^c A^k\right), (14)$$

The application of the ReLU function ensures that only features exerting a positive influence on the class c are retained. Finally, the resulting heatmap "M_c" is upsampled to match the input image dimensions using bilinear interpolation, thereby facilitating effective visualization of the class-discriminative regions (Fig. 5).

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

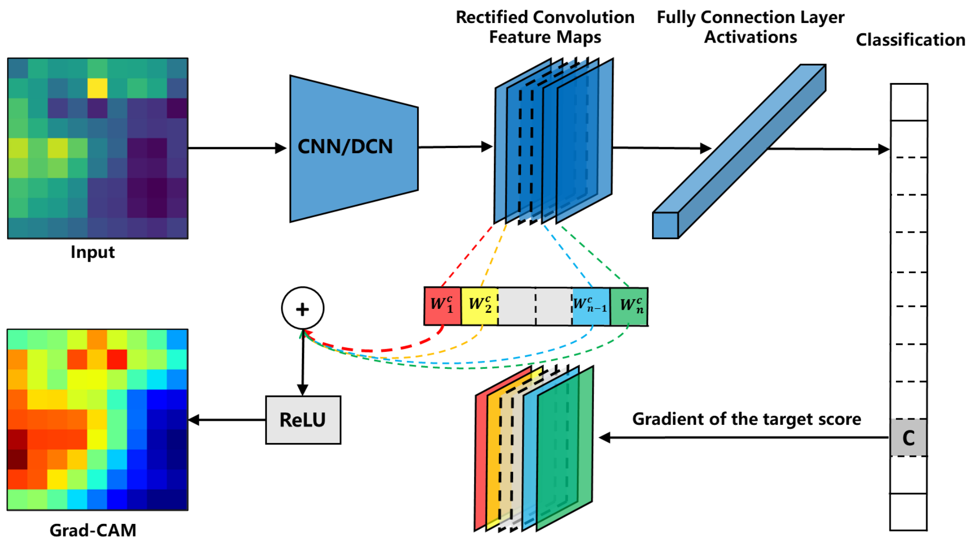


Figure 5: The workflow diagram for obtaining Grad-CAM within convolution neural network and deformable convolution networks.

310 **4. Results and Discussions**

The process begins by preprocessing the geochemical data: each of the 39 elements is interpolated onto a 1 km × 1 km grid using inverse distance weighting. Small cubes are then cropped from this 3D grid and fed into a GK_DCN for feature extraction and anomaly recognition. As a supervised algorithm, the GK_DCN requires a dataset labeled with known anomalies (positive samples) and background (negative samples). A critical aspect of the model is its ability to leverage the varying discriminative power of different spatial positions within the data cubes, which significantly boosts its learning capacity.

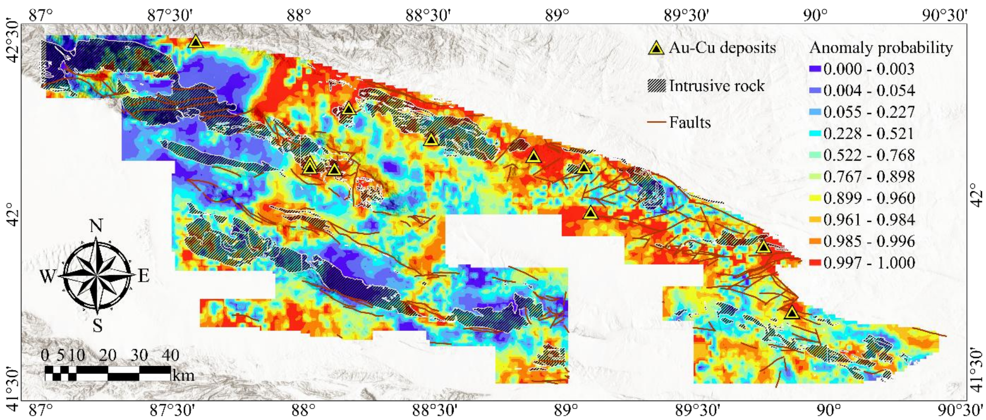
Geochemical anomalies that deviate from regional patterns are key indicators of mineral deposits (Cheng, 2012). To model these anomalies, favorable areas were defined as 3x3 grid blocks centered on known Au-Cu deposits. From each central grid, a 9 x 9 cell patch was extracted, generating 84 positive samples representing mineralized areas. An equal number of negative samples with known deposits were randomly selected from barren regions, following Nykänen et al. (2015). The similar strategy was used for negative sample augmentation generating 84 patches. The dataset was split 8:2 for training and validation, resulting in a final input data cube of dimensions 134 × 9 × 9 × 39 (67 patches per class).

4.1. Recognizing geochemical anomalies by GK_DCN

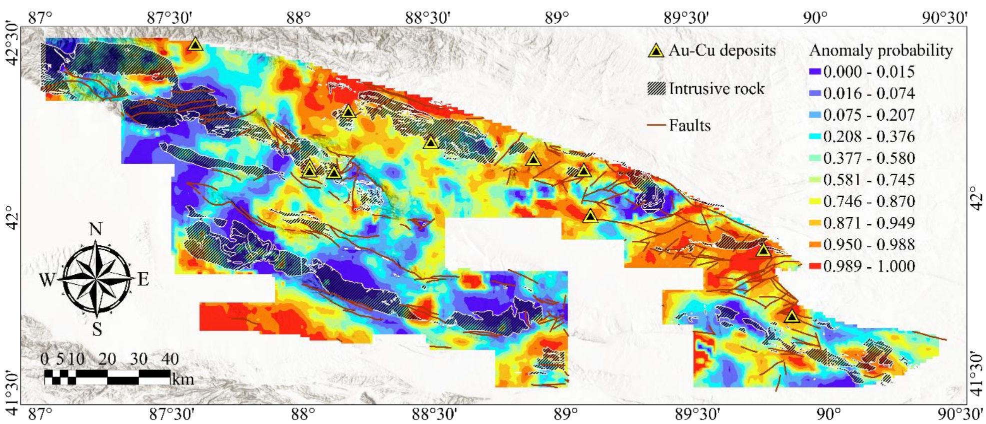
DCN and CNN exhibit significant differences in the extraction of geochemical anomalies. By introducing deformable convolution modules, DCN gains the ability to adaptively adjust the shape and size of receptive fields. Through the

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

incorporation of learnable offset parameters, the convolutional kernels of DCN can dynamically deform based on the characteristics of the input data, learning the complex spatial distribution and structural features of geochemical elements. This allows the model to actively "focus" on the spatial anisotropy of geochemical anomalies, effectively capturing irregular anomaly patterns controlled by geological factors such as lithology. The extracted anomaly boundaries show higher consistency with known ore-forming geological bodies and exhibit stronger spatial continuity (Fig. 6a). In contrast, CNN is constrained by its fixed geometric structure, leading to insufficient responsiveness to irregular boundaries. Its extraction results tend to be overly smooth, with significant loss of anomaly information (Fig. 6b). Comparative results demonstrate that DCN holds clear advantages in improving the spatial positioning accuracy of anomalies and their relevance to geological factors, providing more reliable geochemical indicators for deep mineral exploration. In summary, DCN significantly enhances the ability to represent the nonlinear and anisotropic characteristics of geochemical spatial distributions through its adaptive mechanism.



(a)DCN

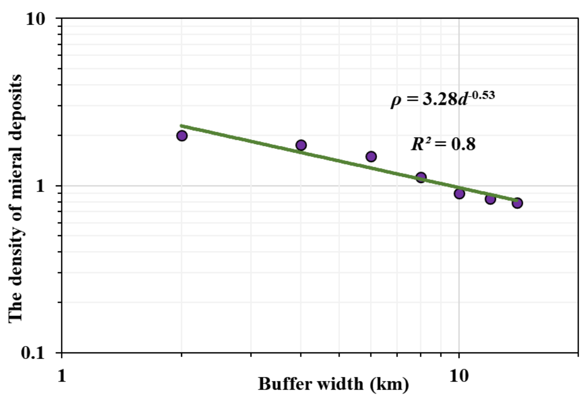


(b)CNN

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

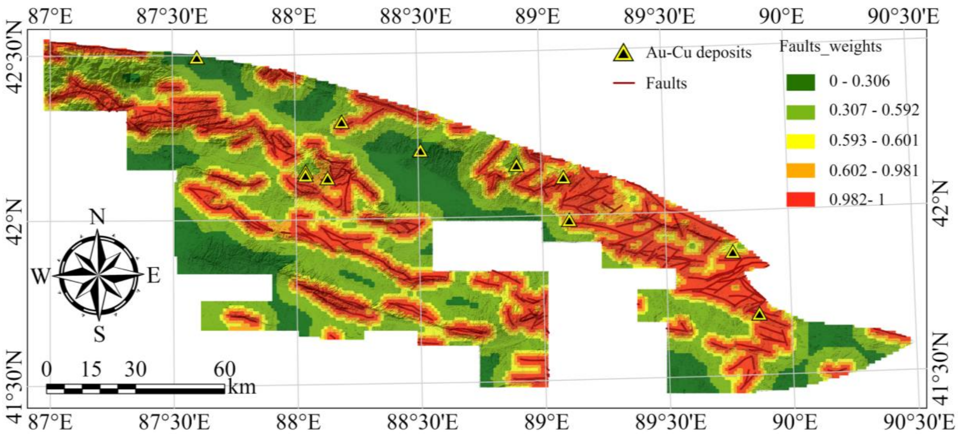
Figure 6: Geochemical anomalies associated with mineralization obtained by (a) DCN and (b) CNN

Reflecting the geological setting where faults and subsidiary fractures provided fluid pathways and deposition sites for Au-Cu mineralization, the quantified spatial relationship between ore-controlling faults and known deposits (Fig. 7) was incorporated into the DCN and CNN's loss function (Fig. 3). A non-linear controlling function between perspective density ρ and d was fitted: $\rho = 3.28d^{-0.53}$. The d was the distance, and ρ was normalized for building a geologically constrained loss term. By incorporating geological constraints constructed from prior knowledge of fault-related mineralization to guide the training of both DCN and CNN, thus generating the GK_DCN and GK_CNN models. These models not only thoroughly learn the spatial distribution patterns and combinatorial relationships of geochemical elements but also strengthen their understanding of the geological background. This effectively suppresses background and noise interference unrelated to mineralization. The results show that compared to traditional methods, the anomalies extracted by the geologically constrained models exhibit higher spatial structural consistency with known mineralized fault structures, and the anomaly concentration centers are more prominent (Fig. 8). This approach significantly reduces the multiplicity of solutions in anomaly recognition and enhances the reliability and geological interpretability of the anomaly results.



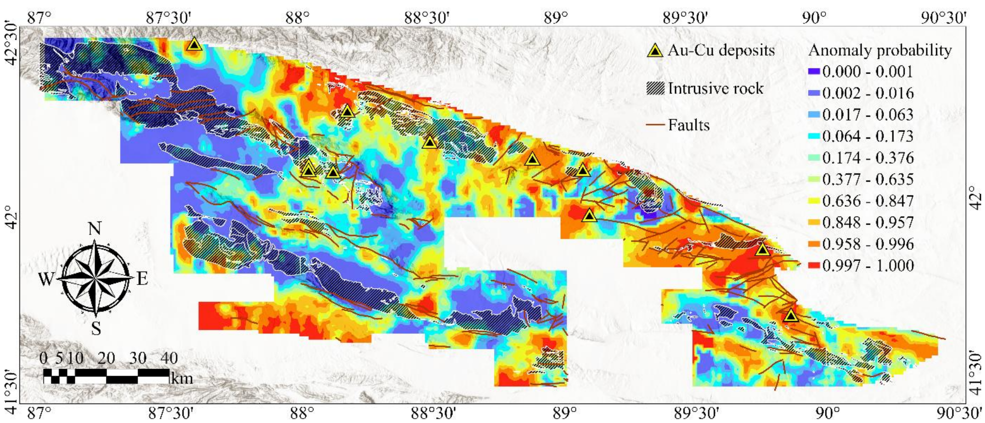
(a)

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

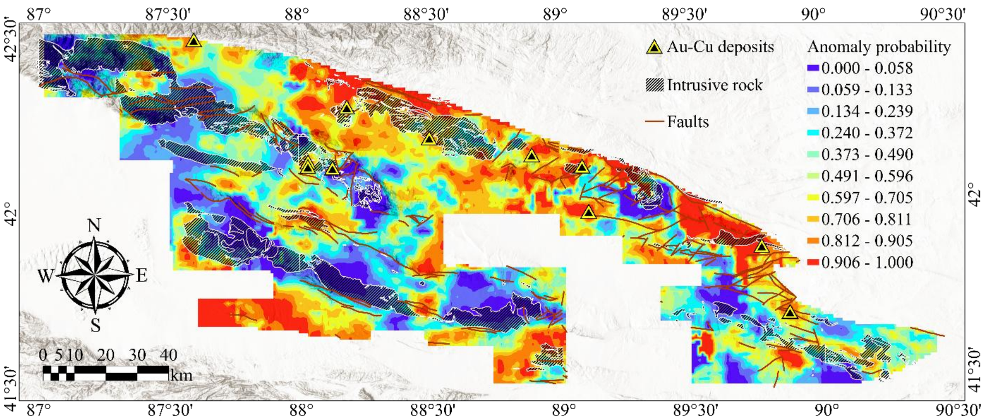


(b)

Figure 7: (a) Log–log plots between the density of mineral deposits ρ and the distance from faults. (b) Density value for faults in the case area.



(a)GK_DCN

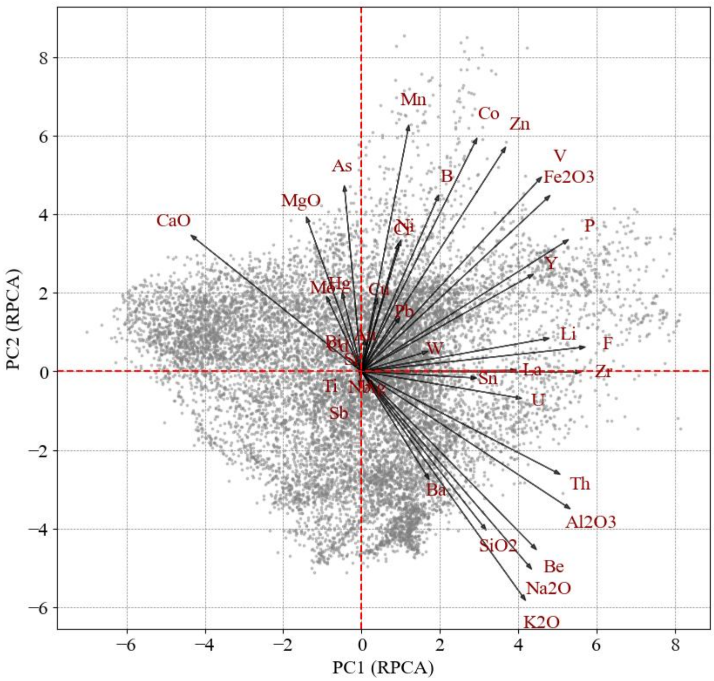


(b)GK_CNN

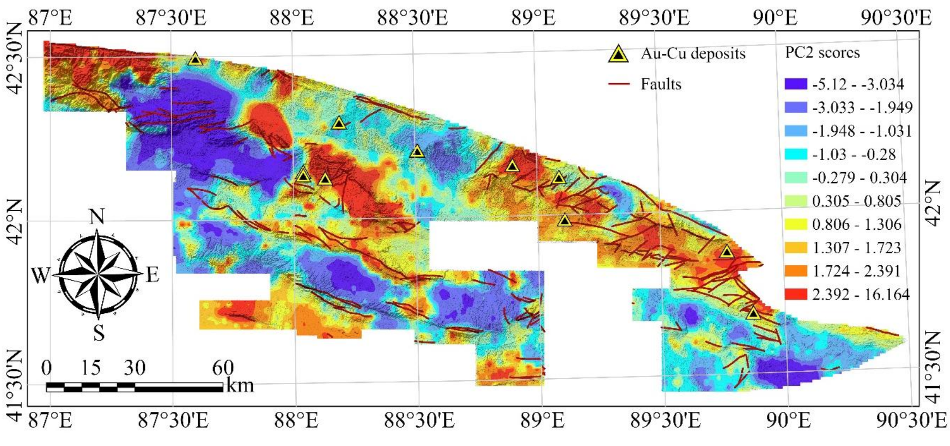
365 **Figure 8: Geochemical anomalies associated with mineralization obtained by (a) GK_DCN and (b) GK_CNN.**

To compare the feature extraction capabilities of CNN and DCN in the identification of geochemical anomalies, this study visualizes the offsets and employs Grad-CAM technology to visualize the spatial features learned by both types of models, followed by a comparison with the geochemical patterns, which can be obtained by integrating multiple geochemical variables via robust principal component analysis (RPCA). PC1 vs. PC2 plots for the 39 elements (Fig. 9a) reveal two distinct compositional assemblages. The assemblage characterized by positive loadings of PC2 (Au, Cu, As, Hg, Bi, Mo, W, Co, Pb, Zn and Ni) (Fig. 9a) corresponds to Au–Cu mineralization in the region. The spatial distribution of PC2 scores (Fig. 9b) shows that low values, associated with this mineralization-related assemblage, correlate with areas of Au–Cu mineralization.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025



(a)



(b)

Figure 9: (a) Biplots of the PC1 and PC2 obtained by the RPCA methods, (b) Map showing the spatial distribution of the second principal component related to mineralization.

As mentioned above, offsets are the core idea of deformable convolution. By introducing a parallel "offset prediction" structure, the network learns the shape and size of the receptive field on its own. For each sampling point of standard convolution, the network additionally learns two values (Δx, Δy), representing its offsets in the x and y directions. Based

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

on this, the actual sampling positions are no longer regular grid points but new positions formed by the original locations plus the predicted offsets. These new positions may distribute along the actual contours of the target object, thereby capturing more precise features. Figure 10 illustrates the geochemical patterns corresponding to ten mineral deposits clipped from PC2 score maps, as well as the offsets direction and magnitude learned by the DCN for ten irregular spatial patterns. For irregular spatial patterns, deformable convolution adjusts the sampling positions of the convolution kernel through offsets. For each position of the convolution kernel, the deformable network adds their corresponding offsets to the original grid points, resulting in new sampling positions that "pull" the originally regular sampling points to more effective locations. The arrows pointing from the original grid points to the new sampling points represent the direction and magnitude of the offsets. Both the direction and magnitude of the offsets indicate that, during the training process, the actual sampling positions of the deformable convolution significantly shift toward areas with higher concentrations of geochemical elements. This demonstrates that the network is more capable of adapting to the quantification and extraction of irregular geochemical spatial patterns.

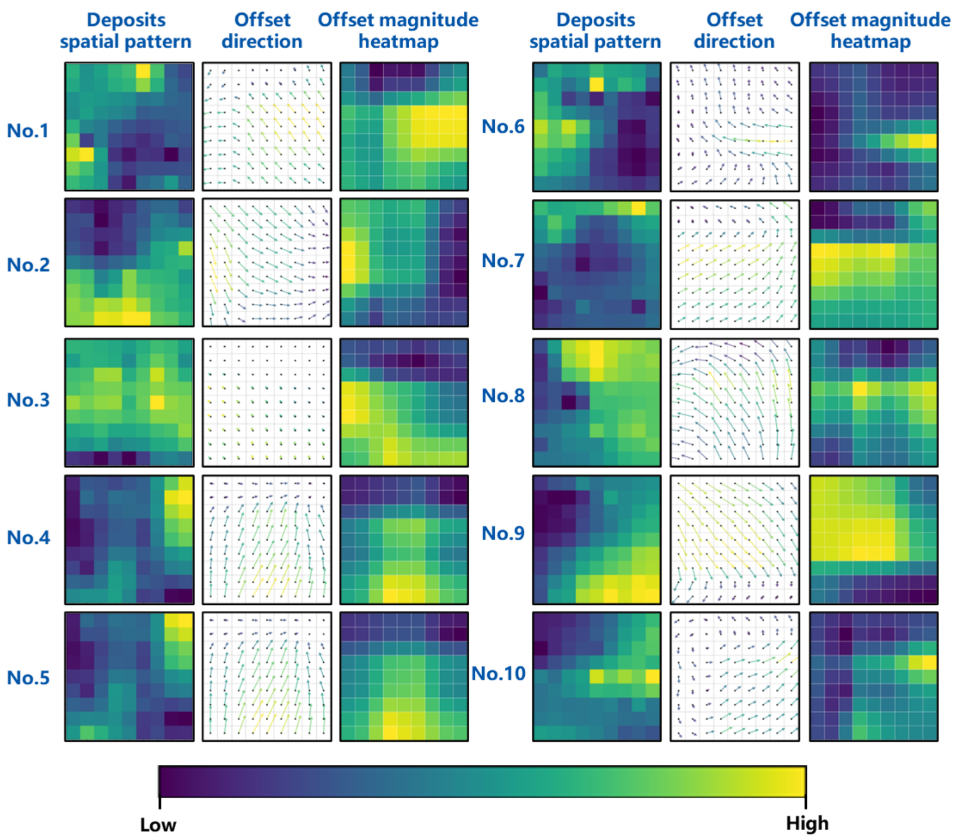


Figure 10: Comparison of offset direction and magnitude maps obtained by GK_CNN and GK_DCN with the geochemical

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

patterns of ten mineral deposits clipped from PC2 score maps in this study. The yellow in the maps represent high concentration and high offset magnitude. The longer the arrow in the offset direction maps, the greater the offset.

CAM is a visualization technique used to reveal the image regions that DCN and CNN focus on when making decisions. It generates a "heatmap" by taking the feature maps of the last convolutional layer and performing a weighted summation. Bright areas indicate regions critical for predicting a specific class. The limitation of CAM is that it requires the network architecture to include a global average pooling layer. Grad-CAM is a generalization and enhancement of CAM. It overcomes the structural constraints of CAM by computing the gradients of the target class with respect to the feature maps of the last convolutional layer to obtain weights, generating a heatmap that localizes key regions of the image. This heatmap visually demonstrates which features the model focuses on to make predictions, thereby enhancing the model's interpretability. It allows us to intuitively understand the basis of the model's decisions and verify whether it is focusing on reasonable features. Figure 10 displays the geochemical patterns corresponding to ten mineral deposits clipped from PC2 score maps, along with the Grad-CAM maps generated by the GK_CNN and GK_DCN models. As can be seen, GK_DCN, with their ability to adaptively adjust receptive fields, generate Grad-CAM maps that more accurately align with the spatial distribution patterns of actual geochemical spatial patterns.

This indicates that the deformable network's ability to adjust the sampling locations of convolutional operations through offset modulation allows it to effectively capture complex and irregular geochemical patterns. Consequently, the deformable network demonstrates greater flexibility and accuracy in identifying and extracting geochemical spatial patterns. Their heatmaps clearly outline the spatially anisotropic distribution of geochemical fields, exhibiting higher spatial coupling with actual geochemical spatial patterns, and enhance the interpretability of model decisions (Fig. 11).

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

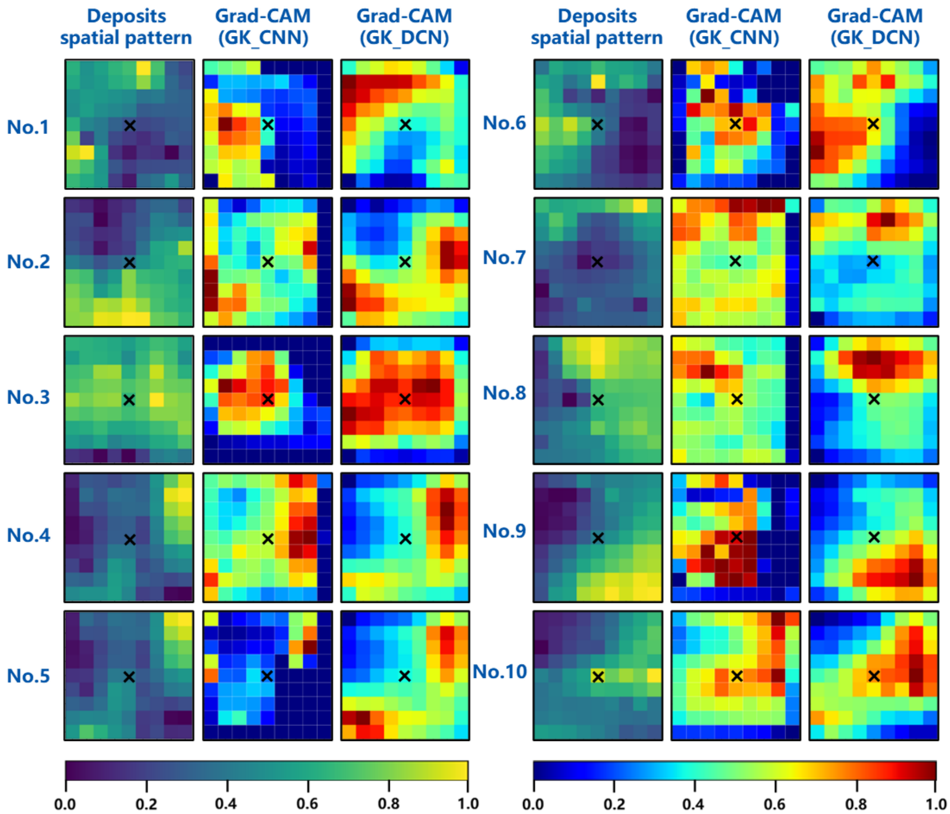


Figure 11: Comparison of Grad-CAM maps obtained by GK_CNN and GK_DCN with the geochemical patterns of ten mineral deposits clipped from PC2 score maps in this study. The yellow in the geochemical patterns of mineral deposits represent high concentration. The red highlighted regions in the Grad-CAM maps are the parts where models give more weight and contribute more the final classification. The black crosses represent the known deposits.

4.2. Comparative experiments

This section assesses the performance of our proposed model using seven metrics—Accuracy (ACC), Area Under the Curve (AUC), Kappa, Matthews Correlation Coefficient (MCC), Precision, Recall, and F1, and compares it with models that are either non-geologically constrained or do not employ deformable convolution operation. The aim is to identify and interpret the performance differences (Chicco and Jurman, 2020; Powers, 2020). The metrics are defined as follows:

$$ACC = \frac{TP+TN}{TP+FP+TN+FN}, (15)$$

$$Kappa = \frac{ACC-P_e}{1-P_e};\ P_e = \frac{\frac{n}{2}(TP+FP)+\frac{n}{2}(FN+FTN)}{n^2}, (16)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}}, (17)$$

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

$Precision = \frac{TP}{TP+FP}$, (18)

$Recall = \frac{TP}{TP+FN}$, (19)

$F1 = \frac{2\times TP}{2\times TP+FP+FN}$, (20)

Here, true positive (TP), true negative (TN), false positive (FP), and false negative (FN) represent the agreement between the actual labels (true or false) and the classifier's predictions (positive or negative), with n denoting the total number of samples. The AUC (Area Under the Curve) corresponds to the area under the Receiver Operating Characteristic (ROC) curve, expressed as a proportion of the total area of the unit square. The ROC curve plots the true positive rate (TPR, or sensitivity) against the false positive rate (FPR, or 1 − specificity) (Fawcett, 2006).

Below is a comparative performance analysis of CNN and DCN in geochemical anomaly recognition tasks based on seven performance metrics. The radar chart comparison clearly shows that the DCN outperforms the standard CNN in the vast majority of performance metrics, demonstrating superior overall performance. In terms of recognition accuracy and reliability, DCN exhibits significant advantages. Its higher accuracy indicates a stronger overall prediction correctness and greater certainty in positive class predictions. In terms of model discriminative ability and error control, DCN also leads. Its larger AUC indicates a stronger ability to distinguish between positive and negative samples and superior ranking quality. Additionally, DCN's lower false positive rate (FPR) means fewer false alarms where normal samples are misclassified as anomalies, which is crucial in practical applications emphasizing safety and efficiency. In summary, due to its deformable convolutional structure, DCN can adaptively adjust the receptive field and more accurately capture the irregular and complex spatial features of anomalies. This enables a comprehensive outperformance over traditional CNN across most of metrics, particularly in reducing missed detections (high recall) and lowering false alarms (low FPR). This demonstrates DCN's stronger applicability and robustness for complex anomaly recognition tasks (Fig. 12).

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

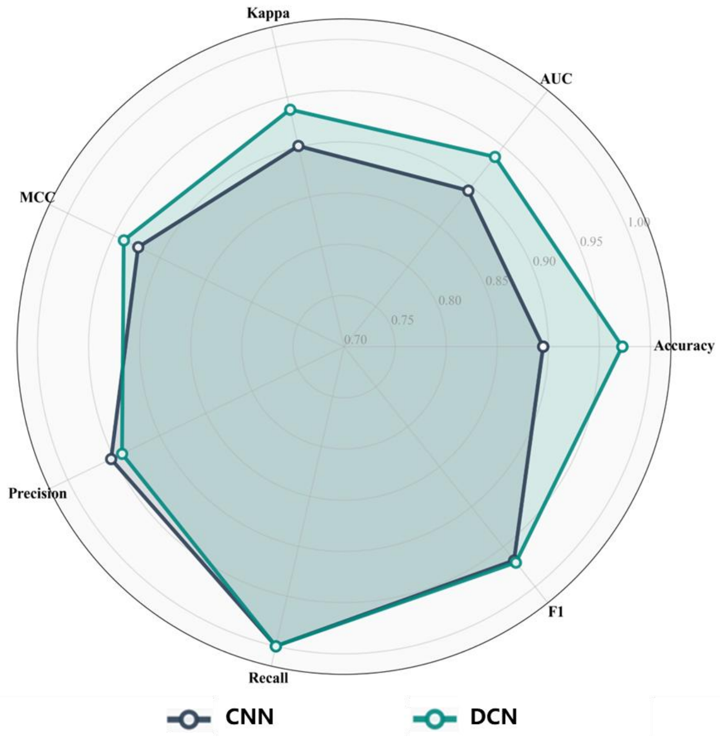


Figure 12: Evaluation of model performance between CNN and DCN in Accuracy, Precision, Recall, F1-score, AUC, Kappa, and MCC.

The model must prioritize not only accuracy but also geological consistency. The radar plot compares its geologically constrained counterpart (GK_CNN and GK_DCNN) (Fig. 13). While both models demonstrate excellent predictive capabilities, GK_CNN and GK_DCN, which incorporates geological knowledge directly within the model architecture, outperformed the unconstrained CNN and DCN. This is evident in key metrics like AUC, Recall, and F1-score, where the knowledge-enhanced model achieved higher performance while successfully integrating geological constraints. The experimental results demonstrate that incorporating geological knowledge (e.g., physical models, constraints) as a physics-based regularization term within the loss function significantly boosts pattern recognition performance and model interpretability. This geologically constrained model effectively identifies potential mineral deposits by guiding training optimization to recognize anomalies associated with ore-controlling faults, enhancing learning and generalization capabilities.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

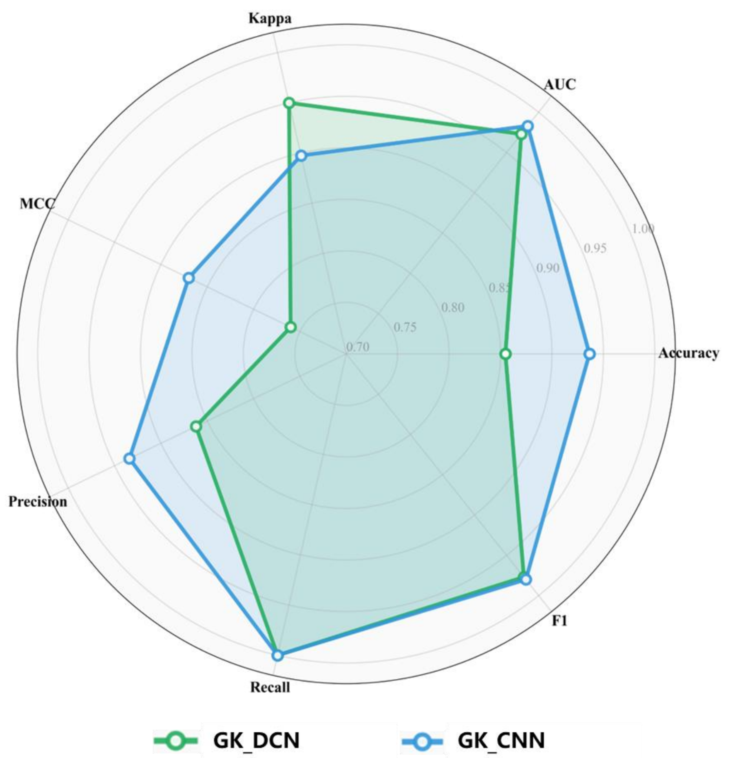


Figure 13: Evaluation of model performance between GK_CNN and GK_DCN in Accuracy, Precision, Recall, F1-score, AUC, Kappa, and MCC.

5. Conclusions

This study introduces deformable convolutional neural networks (DCN) into the field of geochemical anomaly identification to address the issues in capturing irregularly shaped anomalies within complex geological backgrounds. The adaptive receptive field adjustment capability of deformable convolution units enables more precise capture of the spatial distribution characteristics of geochemical anomalous in complex geological settings, enhancing the model's ability to learn and represent geochemical spatial distribution features, thereby achieving superior anomaly identification results. Experimental results demonstrate that, compared to conventional CNN, this method significantly improves accuracy and spatial continuity in anomaly identification, allowing more effective separation of mineralization-related anomalous information from high-dimensional, nonlinear geochemical data.

Prior knowledge of ore-controlling fault is incorporated into the model's loss function as a constraint. The fault-constrained loss function effectively guides the network's learning process, resulting in identified geochemical anomalies that exhibit higher spatial alignment with known fault structures. This enhances the geological significance of the anomalies, reduces interference from the geochemical background field, and improves the accuracy of anomaly identification.

The interpretability of the model is further examined through visualizations of the learned offsets and Grad-CAM. First, the visualization of the offset fields learned by the deformable convolution kernels clearly reveals the network's adaptive receptive field adjustment behavior. The learned offset vectors effectively point to key anomalous spatial structures and irregular trends in the geological mineralization process, serving as important quantitative indicators of anomaly irregularity. Second, Grad-CAM intuitively demonstrates the key regions focused on by the model during decision-making. The highlighted areas in the heatmap show strong overlap with known mineral deposits and high anomaly zones, providing compelling evidence from the "black-box" decision-making process and demonstrating the model's focus on geochemical response features related to mineralization. In summary, this study not only validates the effectiveness of combining deformable convolution with geological prior knowledge in geochemical anomaly identification but also provides a window into understanding the model's decision-making process through offset and Grad-CAM visualizations, significantly enhancing the accuracy and interpretability of AI models in geochemical data processing. This method offers a new tool for deep learning-driven geochemical data analysis and holds practical value for future geochemical exploration.

Code and Data Availability

The code used for geochemical pattern recognition based on the geological knowledge guided deformable convolution network are archived on Zenodo (https://zenodo.org/records/17243487; Zhang et al., 2025). Data supporting this research are available in Wang et al. (2007) from China Geological Survey.

CRediT Authorship Contribution Statement

XZ: Conceptualization, Methodology, Writing – original draft. **YX:** Conceptualization, Resources, Methodology, Writing – original draft. **ZC:** Methodology.

Declaration of Competing Interest

The contact author has declared that none of the authors has any competing interests.

Acknowledgement

This research was supported by the Key Research and Development Program of Xinjiang Uygur Autonomous Region, China (2024B03010-3), and the National Key Research and Development Program of China (NO. 22023YFC2906400, 22023YFC2906404).

References

Chen, Y., Pirajno, F., Li, N., Guo, D., and Lai, Y.: Epithermal deposits in North Xinjiang, NW China, International Journal of Earth Sciences, 101, 889-917, https://doi.org/10.1007/s00531-011-0689-4, 2012.

Chen, Z., and Zuo, R.: Geological-knowledge-guided graph self-supervised pretraining framework for identifying mineralization-related geochemical anomalies, Computers & Geosciences, 199, 105913, https://doi.org/10.1016/j.cageo.2025.105913, 2025.

Cheng, Q., Agterberg, F. P., and Ballantyne, S. B.: The separation of geochemical anomalies from background by fractal methods, Journal of Geochemical Exploration, 51(2), 109-130, https://doi.org/10.1016/0375-6742(94)90013-2, 1994.

Cheng, Q., Xu, Y., and Grunsky, E.: Integrated spatial and spectrum method for geochemical anomaly separation, Natural Resources Research, 9(1), 43-52, https://doi.org/10.1023/A:1010109829861, 2000.

Cheng, Q.: Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China, Ore Geology Reviews, 32(1-2), 314-324, https://doi.org/10.1016/j.oregeorev.2006.10.002, 2007.

Cheng, Q., and Zhao, P.: Singularity theories and methods for characterizing mineralization processes and mapping geo-anomalies for mineral deposit prediction, Geoscience Frontiers, 2(1), 67-79, https://doi.org/10.1016/j.gsf.2010.12.003, 2011.

Cheng, Q., Bonham-Carter, G., Wang, W., Zhang, S., Li, W., and Xia, Q.: A spatially weighted principal component analysis for multi-element geochemical data for mapping locations of felsic intrusions in the Gejiu mineral district of Yunnan, China, Computers & Geosciences, 37(5), 662-669, https://doi.org/10.1016/j.cageo.2010.11.001, 2011.

Cheng, Q.: Singularity theory and methods for mapping geochemical anomalies caused by buried sources and for predicting undiscovered mineral deposits in covered areas, Journal of Geochemical Exploration, 122, 55-70, https://doi.org/10.1016/j.gexplo.2012.07.007, 2012.

Chicco, D., and Jurman, G.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC genomics, 21(1), 6, https://doi.org/10.1186/s12864-019-6413-7, 2020.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

Dai, J., Qi, H., Xiong, Y., Li, Y., Zhang, G., Hu, H., and Wei, Y.: Deformable convolutional networks, In Proceedings of the IEEE International Conference on Computer Vision, pp. 764-773, 2017.

Fawcett, T.: An introduction to ROC analysis, Pattern Recognition Letters, 27(8), 861-874, https://doi.org/10.1016/j.patrec.2005.10.010, 2006.

Gao, L., Hong, H., Liang, D., and Min, F.: Surface wave suppression through deformable convolutional wavelet transform network with residual dense blocks, Acta Geophysica, 72(6), 4151-4167, https://doi.org/10.1007/s11600-024-01339-x, 2024.

Gao, J., Long, L., Klemd, R., Qian, Q., Liu, D., Xiong, X., Su, W., Liu, W., Wang, Y., and Yang, F.: Tectonic evolution of the South Tianshan orogen and adjacent regions, NW China: geochemical and age constraints of granitoid rocks, International Journal of Earth Sciences, 98(6), 1221-1238, https://doi.org/10.1007/s00531-008-0370-8, 2009.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., and Kagal, L.: Explaining explanations: An overview of interpretability of machine learning, In 2018 IEEE 5th International Conference on data science and advanced analytics (DSAA), pp. 80-89, https://doi.org/10.1109/DSAA.2018.00018, 2018.

Guan, J., Lai, R., Lu, Y., Li, Y., Li, H., Feng, L., Yang, Y., and Gu, L.: Memory-efficient deformable convolution based joint denoising and demosaicing for UHD images, IEEE Transactions on Circuits and Systems for Video Technology, 32(11), 7346-7358, https://doi.org/10.1109/TCSVT.2022.3182990, 2022.

Goldfarb, R.J., Taylor, R.D., Collins, G.S., Goryachev, N.A., and Orlandini, O.F.: Phanerozoic continental growth and gold metallogeny of Asia, Gondwana Research, 25, 48-102, https://doi.org/10.1016/j.gr.2013.03.002, 2014.

Gong, L., and Cheng, Q.: Exploiting edge features for graph neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9211-9219, 2019.

Hajihosseinlou, M., Maghsoudi, A., and Ghezelbash, R.: Intelligent mapping of geochemical anomalies: Adaptation of DBSCAN and mean-shift clustering approaches, Journal of Geochemical Exploration, 258, 107393, https://doi.org/10.1016/j.gexplo.2024.107393, 2024.

Jimenez-Espinosa, R., Sousa, A. J., and Chica-Olmo, M.: Identification of geochemical anomalies using principal component analysis and factorial kriging analysis. Journal of Geochemical Exploration, 46(3), 245-256, https://doi.org/10.1016/0375-6742(93)90024-G, 1993.

Jung, H., and Oh, Y.: Towards better explanations of class activation mapping, In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1336-1344, 2021.

Han, B., He, G., Wang, X., and Guo, Z.: Late Carboniferous collision between the Tarim and Kazakhstan–Yili terranes in the western segment of the South Tian Shan Orogen, Central Asia, and implications for the Northern Xinjiang, western China, Earth-Science Reviews, 109(3-4), 74-93, https://doi.org/10.1016/j.earscirev.2011.09.001, 2011.

LeCun, Y., and Bengio, Y.: Convolutional networks for images, speech, and time series, The handbook of brain theory and neural networks, 1998.

Liu, H., Li, Z., Lin, S., and Cheng, L.: Remote sensing image denoising based on deformable convolution and attention-guided filtering in progressive framework, Signal, Image and Video Processing, 18(11), 8195-8205, https://doi.org/10.1007/s11760-024-03461-1, 2024.

Liu, Y., Wang, W., Li, Q., Min, M., and Yao, Z.: DCNet: A deformable convolutional cloud detection network for remote sensing imagery, IEEE Geoscience and Remote Sensing Letters, 19, 1-5, https://doi.org/10.1109/LGRS.2021.3086584, 2021.

Luo, Y., Liu, X., Meng, H., Ye, Y., and Wu, B.: Multitask Seismic Inversion Based on Deformable Convolution and Generative Adversarial Network, IEEE Geoscience and Remote Sensing Letters, 21, 1-5, https://doi.org/10.1109/LGRS.2024.3388213, 2024.

Luo, Z., Zuo, R., Xiong, Y., and Zhou, B.: Metallogenic-factor variational autoencoder for geochemical anomaly detection by ad-hoc and post-hoc interpretability algorithms, Natural Resources Research, 32(3), 835-853, https://doi.org/10.1007/s11053-023-10200-9, 2023.

Na, Z., Guo, Z., and Zhu, Y.: Soil Moisture Monitoring Based on Deformable Convolution Unit Net Algorithm Combined with Water Area Changes, Electronics, 14(5), 1011, https://doi.org/10.3390/electronics14051011, 2025.

Nielsen, A. H., Iosifidis, A., and Karstoft, H.: Forecasting large-scale circulation regimes using deformable convolutional neural networks and global spatiotemporal climate data, Scientific Reports, 12(1), 8395, https://doi.org/10.1038/s41598-022-12167-8, 2022.

Nykänen, V., Lahti, I., Niiranen, T., and Korhonen, K.: Receiver operating characteristics (ROC) as validation tool for prospectivity models—A magmatic Ni–Cu case study from the Central Lapland Greenstone Belt, Northern Finland, Ore Geology Reviews, 71, 853-860, https://doi.org/10.1016/j.oregeorev.2014.09.007, 2015.

Pirajno, F.: Hydrothermal processes and mineral systems. Springer Science & Business Media, 2008.

Powers, D.M.: Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, arXiv preprint arXiv:2010.16061, 2020.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

Reis, A. P., Sousa, A. J., and Fonseca, E. C.: Application of geostatistical methods in gold geochemical anomalies identification (Montemor-O-Novo, Portugal). Journal of Geochemical Exploration, 77(1), 45-63, https://doi.org/10.1016/S0375-6742(02)00269-8, 2003.

Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nature Machine Intelligence, 1(5), 206-215, https://doi.org/10.1038/s42256-019-0048-x, 2019.

Scarselli, F., Gori, M., Tsoi, A. C., and Hagenbuchner, M., Monfardini, G.: The graph neural network model, IEEE Transactions on Neural Networks, 20(1), 61-80, https://doi.org/10.1109/TNN.2008.2005605, 2008.

Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., and Batra, D.: Grad-CAM: Why did you say that?, arXiv preprint arXiv:1611.07450, 2016.

Seltmann, R., Porter, T.M., and Pirajno, F.: Geodynamics and metallogeny of the central Eurasian porphyry and related epithermal mineral systems: A review, Journal of Asian Earth Sciences, 79, 810-841, https://doi.org/10.1016/j.jseaes.2013.03.030, 2014.

Semenov, A., Boginski, V., and Pasiliao, E. L.: Neural networks with multidimensional cross-entropy loss functions, International Conference on Computational Data and Social Networks, Cham: Springer International Publishing, 57-62, https://doi.org/10.1007/978-3-030-34980-6_5, 2019.

Soltani, Z., Hassani, H., and Esmaeiloghli, S.: A deep autoencoder network connected to geographical random forest for spatially aware geochemical anomaly detection, Computers & Geosciences, 190, 105657, https://doi.org/10.1016/j.cageo.2024.105657, 2024.

Sun, X., Mo, T., Song, J., and Wang, B.: Deformable convolution kernel and residual learning assisted irregular seismic data interpolation, IEEE Transactions on Geoscience and Remote Sensing, 62, 1-11, https://doi.org/10.1109/TGRS.2024.3360449, 2024.

Tian, J., Guo, X., Liu, W., Tao, D., and Liu, B.: Deformable convolutional network constrained by contrastive learning for underwater image enhancement, IEEE Geoscience and Remote Sensing Letters, 20, 1-5, https://doi.org/10.1109/LGRS.2023.3299613, 2023.

Tian, M., Wang, X., Nie, L., and Zhang, C.: Recognition of geochemical anomalies based on geographically weighted regression: A case study across the boundary areas of China and Mongolia, Journal of Geochemical Exploration, 190, 381-389, https://doi.org/10.1016/j.gexplo.2018.04.003, 2018.

Wang, H., and Zuo, R.: A comparative study of trend surface analysis and spectrum–area multifractal model to identify geochemical anomalies. Journal of Geochemical Exploration, 155, 84-90, https://doi.org/10.1016/j.gexplo.2015.04.013, 2015.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

Wang, H., Cheng, Q., and Zuo, R.: Quantifying the spatial characteristics of geochemical patterns via GIS-based geographically weighted statistics, Journal of Geochemical Exploration, 157, 110-119, https://doi.org/10.1016/j.gexplo.2015.06.004, 2015.

Wang, J., and Zuo, R.: An extended local gap statistic for identifying geochemical anomalies, Journal of Geochemical Exploration, 164, 86-93, https://doi.org/10.1016/j.gexplo.2016.01.002, 2016.

Wang, W., Zhao, J., and Cheng, Q.: Fault trace-oriented singularity mapping technique to characterize anisotropic geochemical signatures in Gejiu mineral district, China. Journal of Geochemical Exploration, 134, 27-37, https://doi.org/10.1016/j.gexplo.2013.07.009, 2013.

Wang, W., Cheng, Q., Zhang, S., and Zhao, J.: Anisotropic singularity: A novel way to characterize controlling effects of geological processes on mineralization, Journal of Geochemical Exploration, 189, 32-41, https://doi.org/10.1016/j.gexplo.2017.07.019, 2018.

Wang, X., Zhang, Q., and Zhou, G.: National-scale geochemical mapping projects in China, Geostandards and Geoanalytical Research, 31(4), 311-320, https://doi.org/10.1111/j.1751-908X.2007.00128.x, 2007.

Wu, Z., Paoletti, M. E., Su, H., Tao, X., Han, L., Haut, J. M., and Plaza, A.: Background-guided deformable convolutional autoencoder for hyperspectral anomaly detection, IEEE Transactions on Geoscience and Remote Sensing, 61, 1-16, https://doi.org/10.1109/TGRS.2023.3334562, 2023.

Xiao, F., Chen, J., Zhang, Z., Wang, C., Wu, G., and Agterberg, F. P.: Singularity mapping and spatially weighted principal component analysis to identify geochemical anomalies associated with Ag and Pb-Zn polymetallic mineralization in Northwest Zhejiang, China, Journal of Geochemical Exploration, 122, 90-100, https://doi.org/10.1016/j.gexplo.2012.04.010, 2012.

Xiao, F., Chen, J., Hou, W., Wang, Z., Zhou, Y., and Erten, O.: A spatially weighted singularity mapping method applied to identify epithermal Ag and Pb-Zn polymetallic mineralization associated geochemical anomaly in Northwest Zhejiang, China. Journal of Geochemical Exploration, 189, 122-137, https://doi.org/10.1016/j.gexplo.2017.03.017, 2018.

Xiao, F., Wang, K., Hou, W., and Erten, O.: Identifying geochemical anomaly through spatially anisotropic singularity mapping: A case study from silver-gold deposit in Pangxidong district, SE China, Journal of Geochemical Exploration, 210, 106453, https://doi.org/10.1016/j.gexplo.2019.106453, 2020.

Xie, X., Mu, X., and Ren, T.: Geochemical mapping in China, Journal of Geochemical Exploration, 60(1), 99-113, https://doi.org/10.1016/S0375-6742(97)00029-0, 1997.

Xiong, Y., Zuo, R., Luo, Z., and Wang, X.: A physically constrained variational autoencoder for geochemical pattern recognition, Mathematical Geosciences, 54(4), 783-806, https://doi.org/10.1007/s11004-021-09979-1, 2022.

Xu, L., Zhang, X., Yu, H., Chen, Z., Du, W., and Chen, N.: Incorporating spatial autocorrelation into deformable ConvLSTM for hourly precipitation forecasting, Computers & Geosciences, 184, 105536, https://doi.org/10.1016/j.cageo.2024.105536, 2024.

Xu, Y., Zuo, R., and Zhang, G.: The graph attention network and its post-hoc explanation for recognizing mineralization-related geochemical anomalies, Applied Geochemistry, 155, 105722, https://doi.org/10.1016/j.apgeochem.2023.105722, 2023.

Xu, Y., Shi, L., and Zuo, R.: Geologically constrained unsupervised dual-branch deep learning algorithm for geochemical anomalies identification, Applied Geochemistry, 174, 106137, https://doi.org/10.1016/j.apgeochem.2024.106137, 2024.

Xu, Y., Zuo, R., Chen, Z., Shi, Z. and Kreuzer, O. P.: Recent advances and future research directions in deep learning as applied to geochemical mapping, Earth-Science Reviews, 105209, https://doi.org/10.1016/j.earscirev.2025.105209, 2025.

Xue, C., Zhao, X., Mo, X., Gu, X., Zhang, Z., Wang, X., Zu, B., Zhang, G., Feng, B., Liu, J., Dong, L., Bakhtiar, N., and Nikolay, P.: Asian Gold Belt in western Tianshan and its dynamic setting, metallogenic control and exploration, Earth Science Frontiers, 21(5), 128-155, https://doi.org/10.19762/j.cnki.dizhixuebao.2023114, 2014.

Yin, B., Zuo, R., Xiong, Y., Li, Y., and Yang, W.: Knowledge discovery of geochemical patterns from a data-driven perspective. Journal of Geochemical Exploration, 231, 106872, https://doi.org/10.1016/j.gexplo.2021.106872, 2021.

Yu, D., Ji, S., Li, X., Yuan, Z., and Shen, C.: Earthquake crack detection from aerial images using a deformable convolutional neural network, IEEE Transactions on Geoscience and Remote Sensing, 60, 1-12, https://doi.org/10.1109/TGRS.2022.3183157, 2022.

Yu, H., Wang, R., Li, P., and Zhang, P.: Flood detection in polarimetric SAR data using deformable convolutional vision model, Water, 15(24), 4202, https://doi.org/10.3390/w15244202, 2023.

Zhang, S., Xiao, K., Carranza, E. J. M., Yang, F., and Zhao, Z.: Integration of auto-encoder network with density-based spatial clustering for geochemical anomaly detection for mineral exploration, Computers & Geosciences, 130, 43-56, https://doi.org/10.1016/j.cageo.2019.05.011, 2019.

Zhang, T., Zhang, W., Cao, D., Yi, Y., and Wu, X.: A new deep learning neural network model for the identification of InSAR anomalous deformation areas, Remote Sensing, 14(11), 2690, https://doi.org/10.3390/rs14112690, 2022.

https://doi.org/10.5194/egusphere-2025-4877
Preprint. Discussion started: 15 October 2025

Zhang, X., Xiong, Y., and Chen, Z.: Recognizing geochemical spatial patterns using deformable convolutional networks guided with geological knowledge, GitHub [code], https://github.com/Aleksiyu/Recognizing-geochemical-spatial-patterns-using-DCN-guided-with-geological-knowledge, 2025.

Zhao, C., Zhu, W., and Feng, S.: Hyperspectral image classification based on kernel-guided deformable convolution and double-window joint bilateral filter, IEEE Geoscience and Remote Sensing Letters, 19, 1-5, https://doi.org/10.1109/LGRS.2021.3084203, 2021.

Zhao, H., Zhou, Y., Bai, T., and Chen, Y.: A U-Net based multi-scale deformable convolution network for seismic random noise suppression, Remote Sensing, 15(18), 4569, https://doi.org/10.3390/rs15184569, 2023.

Zhao, X., Xue, C., Seltmann, R., Dolgopolova, A., Andersen, J. C., and Zhang, G.: Volcanic–plutonic connection and associated Au‐Cu mineralization of the Tulasu ore district, Western Tianshan, NW China: Implications for mineralization potential in Palaeozoic arc terranes, Geological Journal, 55(3), 2318-2341, https://doi.org/10.1002/gj.3750, 2020.

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M.: Graph neural networks: A review of methods and applications, AI Open, 1, 57-81, https://doi.org/10.1016/j.aiopen.2021.01.001, 2020.

Zhu, J., Fang, L., and Ghamisi, P.: Deformable convolutional neural networks for hyperspectral image classification, IEEE Geoscience and Remote Sensing Letters, 15(8), 1254-1258, https://doi.org/10.1109/LGRS.2018.2830403 2018.

Zhu, X., Hu, H., Lin, S., and Dai, J.: Deformable convnets v2: More deformable, better results, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9308-9316, 2019.

Zuo, R.: A nonlinear controlling function of geological features on magmatic–hydrothermal mineralization, Scientific Reports, 6(1), 27127, https://doi.org/10.1038/srep27127, 2016.

Zuo, R., Wang, J., Xiong, Y., and Wang, Z.: The processing methods of geochemical exploration data: past, present, and future, Applied Geochemistry, 132, 105072, https://doi.org/10.1016/j.apgeochem.2021.105072, 2021.

Zuo, R.: Machine learning of mineralization-related geochemical anomalies: A review of potential methods. Natural Resources Research, 26(4), 457-464, https://doi.org/10.1007/s11053-017-9345-4, 2017.

Zuo, R., Cheng, Q., Xu, Y., Yang, F., Xiong, Y., Wang, Z., and Kreuzer, O. P.: Explainable artificial intelligence models for mineral prospectivity mapping, Science China Earth Sciences, 67(9) 2864-2875, https://doi.org/10.1007/s11430-024-1309-9, 2024.

Zuo, R., Yang, F., Cheng, Q., and Kreuzer, O. P.: A novel data-knowledge dual-driven model coupling artificial intelligence with a mineral systems approach for mineral prospectivity mapping, Geology, 53(3), 284-288, https://doi.org/10.1130/G52970.1, 2025.