Scholars Journal of Engineering and Technology

OPEN ACCESS

Abbreviated Key Title: Sch J Eng Tech
ISSN 2347-9523 (Print) | ISSN 2321-435X (Online)
Journal homepage: https://saspublishers.com

Geossgraph: Semi-Supervised Geologically Informed Graph Network with Contrastive Pretraining for Mineral Prospectivity Mapping

Hakeem B. Ajileye[1,2*]

[1]Geomathematics Key Laboratory of Sichuan Province (Chengdu University of Technology), Chengdu 610059, China
[2]School of Mathematical Sciences, Chengdu University of Technology, Chengdu 610000, China

DOI: https://doi.org/10.36347/sjet.2026.v14i07.006 | **Received:** 25.05.2026 | **Accepted:** 06.07.2026 | **Published:** 13.07.2026

***Corresponding author:** Hakeem B. Ajileye
Geomathematics Key Laboratory of Sichuan Province (Chengdu University of Technology), Chengdu 610059, China

Abstract | **Original Research Article**

Mineral Prospectivity Mapping (MPM) is often constrained by limited labeled samples and the inability of fully supervised graph neural networks to exploit abundant unlabeled spatial data. This study proposes GeoSSGraph, a semi-supervised geologically informed graph network that combines geologically modulated contrastive pretraining (GeoConLoss) with iterative pseudo-labeling (GeoSelfTrain) for mineral prospectivity prediction. The framework integrates Geo Band Weighting and a Geo Relational Encoder comprising three residual basis-decomposed Relational Graph Convolutional Network layers to capture directional geochemical and structural relationships. GeoSSGraph was evaluated on a 43-band geochemical-structural dataset from the Lhasa-Woka region of the eastern Gangdese belt, Tibetan Plateau, consisting of 39 elemental concentration layers, three isometric log-ratio balances, and fracture density. Using a five-model ensemble selected from ten independently trained seeds, GeoSSGraph achieved 89.58% accuracy, 0.9306 ROC-AUC, 0.8980 F1-score, 91.67% sensitivity, and 87.50% specificity on an independent test set. The resulting prospectivity map delineates coherent mineralization corridors aligned with major fault systems and intrusive contacts. These results demonstrate the effectiveness of geologically informed semi-supervised learning for MPM under conditions of label scarcity.

Keywords: Mineral Prospectivity Mapping, Semi-Supervised Learning, Relational Graph Convolutional Network, Contrastive Pretraining, Pseudo-Labeling.

INTRODUCTION

Mineral exploration is fundamentally an exercise in identifying where geological processes have concentrated economically valuable metals within the Earth's crust. The discovery of new ore deposits has become progressively more challenging as near-surface and easily identifiable deposits are exhausted, shifting exploration targets toward deeper, more structurally complex, and geochemically subtle settings [1]. Mineral Prospectivity Mapping (MPM) was developed to address this challenge by providing a quantitative and spatially explicit framework for combining multiple lines of geological evidence, geochemical, structural, geophysical, and lithological, into a single continuous prediction surface that ranks areas by their likelihood of hosting mineralization [2]. A reliable prospectivity map can dramatically reduce the cost and risk of exploration by directing fieldwork, drilling, and geophysical surveys toward the most promising targets while eliminating large areas of low potential from further consideration. As the volume and diversity of available geoscientific data continue to grow through advances in remote sensing, regional geochemical surveying, and digital geological mapping, the demand for MPM methods capable of extracting complex, nonlinear, and directionally structured mineralization signals from these datasets has intensified [3], [4].

The earliest approaches to MPM relied on statistical and knowledge-driven methods such as weights-of-evidence, logistic regression, and weighted overlay analysis, which combine evidence layers by assigning importance scores based on empirical or expert-defined relationships between geological features and known deposit locations [5]. These methods are interpretable and computationally accessible but operate under strong simplifying assumptions, most critically that spatial evidence is isotropic, and that mineralization probability can be expressed as a linear or log-linear combination of input variables. In reality, ore-forming systems are controlled by highly nonlinear interactions between magmatic, structural, and hydrothermal

processes that produce geochemical anomalies and alteration halos with strong directional continuity along fault corridors, stratigraphic contacts, and intrusive boundaries [6]. The subsequent adoption of machine learning classifiers, including support vector machines, random forests, and gradient boosting methods, improved the ability to capture nonlinear relationships between input features and deposit labels. Still, these methods continue to operate on tabular feature vectors extracted from raster neighborhoods and do not encode spatial topology, directionality, or the multi-scale structural context that governs fluid migration and metal deposition [7], [8].

The introduction of deep learning into MPM substantially expanded the representational capacity available for modeling complex geochemical and structural patterns. Convolutional Neural Networks (CNNs) process multi-band raster stacks through learnable local filters, enabling the automatic detection of spatially structured geochemical anomalies and textural features at multiple scales [9], [10]. Despite these advances, CNN-based models apply convolutional kernels uniformly across the spatial domain with a fixed, isotropic receptive field, making them structurally incapable of modeling directional anisotropy or the long-range continuity of geochemical gradients along fault systems and lithological boundaries [11], [12]. Fault-parallel enrichment patterns, conjugate shear fabrics, and oblique intrusive contacts, the structural features most strongly associated with mineralization in porphyry and skarn systems, are smoothed out rather than preserved by isotropic convolution, limiting the geological fidelity of CNN-based prospectivity outputs in tectonically complex settings [13], [14].

Graph Neural Networks (GNNs) emerged as a natural solution to the limitations of Euclidean deep learning by operating directly on non-Euclidean spatial domains and representing irregular, topology-dependent relationships between geological entities as graphs [15], [16]. In the context of MPM, Graph Convolutional Networks (GCNs) and Graph Attention Networks (GATs) have shown improved performance over raster-based methods by incorporating spatial adjacency and edge-weighted interactions that partially reflect geological structure [17], [18], [19]. However, these architectures carry their own critical limitations. GCNs aggregate node features isotropically across all spatial neighbors, treating north-south stratigraphic connections identically to east-west compositional gradients and diagonal fault-parallel pathways, which causes directional geological signals to be averaged out rather than preserved [20]. GATs introduce learnable attention weights that allow the model to focus on more informative neighbors, but the attention computation is based solely on feature content and operates without any mechanism for encoding the directional relationship between nodes, leaving long-range structural alignment and NE-SW or NW-SE shear fabric patterns poorly captured [21], [22], [23].

Relational Graph Convolutional Networks (RGCNs) were developed specifically to extend GNNs to multi-relational graphs by applying separate learned transformation matrices for each relation type, making them a more natural fit for MPM graphs where nine directional spatial relations (N, S, E, W, NE, NW, SE, SW, SELF) constitute structurally distinct edge categories [24]. However, a systematic review of RGCN variants and their applications to geoscientific prediction tasks reveals three fundamental limitations that constrain their effectiveness in MPM. First, RGCN applies equal aggregation weight to all incoming relation types without any mechanism for the model to differentially suppress or amplify specific directional pathways during learning, meaning that geologically dominant orientations such as NE-SW fault corridors cannot be selectively reinforced over uninformative cross-directional connections. Second, RGCN shares the oversmoothing vulnerability of all message-passing GNNs: as network depth increases to capture extended spatial dependencies, repeated degree-normalized aggregation drives node representations toward indistinct cluster means, limiting the spatial resolution of the learned features at precisely the scales where meaningful geochemical gradients exist. Third, and most critically for MPM, RGCN is a fully supervised architecture that requires labeled training samples for every optimization step, providing no mechanism to exploit the spatial structure, relational topology, and geochemical co-variation embedded in the large population of unlabeled prediction tiles that constitutes the actual exploration domain.

Label scarcity is a defining operational constraint in MPM. In geochemical survey regions such as Lhasa-Woka, the number of tiles with confirmed mineralization records is typically a small fraction of the total prediction domain: 117 labeled mineralized tiles against a prediction space of 5,500 spatial units in the present study. The vast majority of tiles carry no deposit label but contain rich geochemical-structural information about the regional background, anomaly gradients, and structural connectivity patterns of the study area. Fully supervised models, including standard RGCN, discard this information entirely during training, forfeiting an informative learning signal that is freely available from the spatial graph itself. Semi-supervised learning frameworks that can leverage both labeled and unlabeled data are therefore not merely desirable but geologically motivated: the spatial organization of geochemical anomalies, the directional continuity of structural fabrics, and the relational topology of the tile graph all carry predictive information about mineralization potential that can be exploited without any deposit labels.

To address these limitations, this study introduces GeoSSGraph, a Semi-Supervised Geologically Informed Graph Network for mineral

prospectivity mapping. GeoSSGraph is motivated directly by the limitations of RGCN, as identified through a systematic review: it replaces static, equal-contribution relational aggregation with a backbone whose relational encoder is first pre-trained to learn geologically meaningful spatial representations through a novel contrastive objective, then fine-tuned with supervised labels augmented by iterative pseudo-labeling of high-confidence unlabeled tiles. The primary contributions of this study are:

- ❖ GeoConLoss (Geologically Modulated Contrastive Loss): A geologically informed contrastive pretraining objective applied to all 186 training tiles without labels, where node-level temperature scaling is derived from the directional relation weights and spatial connectivity of each tile in the graph. By assigning lower contrastive temperatures to tiles embedded in high-weight directional neighborhoods (e.g., diagonal fault-aligned connections), GeoConLoss enforces that tiles interacting through geologically dominant orientations must produce more tightly discriminated representations, embedding structural anisotropy directly into the self-supervised representation landscape.
- ❖ GeoSelfTrain (Iterative Pseudo-Labeling): A geologically motivated semi-supervised augmentation strategy that applies the partially trained model to validation tiles across three iterative rounds, selects high-confidence positive and negative predictions above asymmetric probability thresholds, and progressively incorporates these pseudo-labeled tiles into the supervised training set with adaptive threshold annealing. This exploits the unlabeled spatial structure of the prediction domain to expand the effective training set beyond the limited labeled pool.
- ❖ GeoRelationalEncoder with Residual RGCN: A shared backbone combining GeoBandWeighting, a sigmoid-activated learnable channel gate over all 43 input bands, with three stacked basis-decomposed RGCN layers incorporating residual skip connections and layer normalization, addressing the oversmoothing limitation of standard RGCN by preventing representational collapse with depth while retaining the relation-type-specific transformation structure.

The remainder of this paper is organized as follows: Section 2 describes the study area, dataset preparation, and the GeoSSGraph architecture; Section 3 presents the experimental results, comparative analysis, and geological interpretation; and Section 4 concludes the study.

MATERIALS AND METHODS

This section presents the methodology underlying GeoSSGraph, a semi-supervised geologically informed graph network framework for mineral prospectivity mapping. We begin by describing the geological setting of the Lhasa-Woka study area and the geochemical and structural datasets used for model development, including the construction of isometric log-ratio balances and the preparation of the 43-band raster input. We then introduce the GeoSSGraph architecture, detailing the GeoBandWeighting module, the GeoRelationalEncoder backbone with basis-decomposed RGCN layers, the GeoConLoss contrastive pretraining stage, and the GeoSelfTrain iterative pseudo-labeling fine-tuning stage. Training configuration, multi-seed ensemble strategy, and regularization details are provided to ensure full reproducibility, and the section concludes with the evaluation protocol, which defines the metrics and validation scheme used to benchmark model performance. The overall workflow is illustrated in Figure 1.

2.1 Study Area, Data Sources, and Preprocessing

2.1.1 Geological Setting

The investigation was conducted in the Lhasa-Woka district of the eastern Gangdese belt, situated within the Lhasa terrane of the Tibetan Plateau, bounded to the north by the Bangonghu-Nujiang suture and to the south by the Indus-Yarlung Zangbo suture (Figure 2) [25]. The tectonic history of this terrane is defined by protracted ocean subduction and continental collision events associated with the sequential closure of Tethyan ocean basins, which collectively generated the compressional structures, magmatic arcs, and metallogenic belts that characterize the region today [26], [27]. The terrane experienced bidirectional subduction southward beneath the Meso-Tethys and northward beneath the Neo-Tethys, producing a pervasively deformed terrane marked by thrust nappes, recumbent folds, and densely developed fault networks that now define the primary structural pathways for ore-forming fluid migration [28], [29].

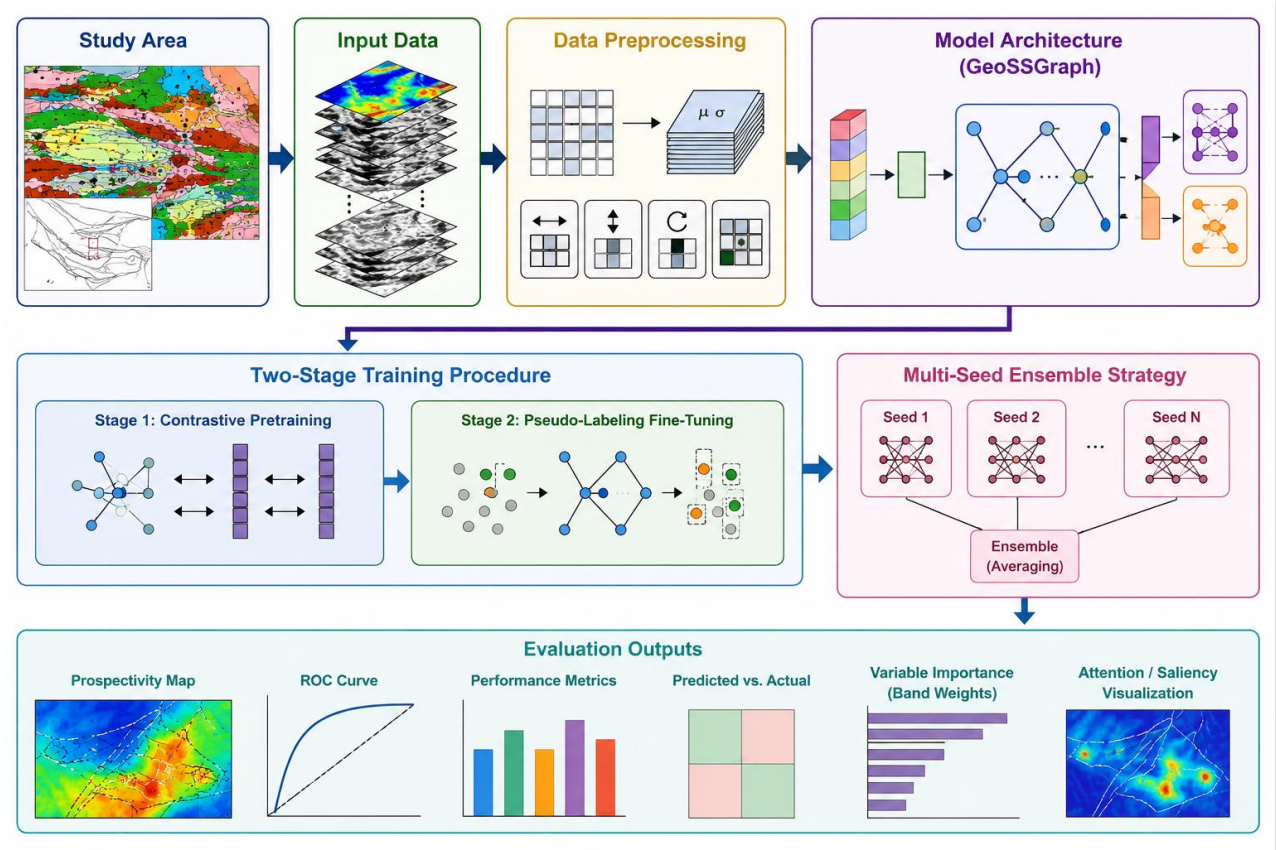


Figure 1: Overall workflow of GeoSSGraph encompassing the study area and input data, data preprocessing, model architecture, two-stage training procedure, multi-seed ensemble strategy, and evaluation outputs.

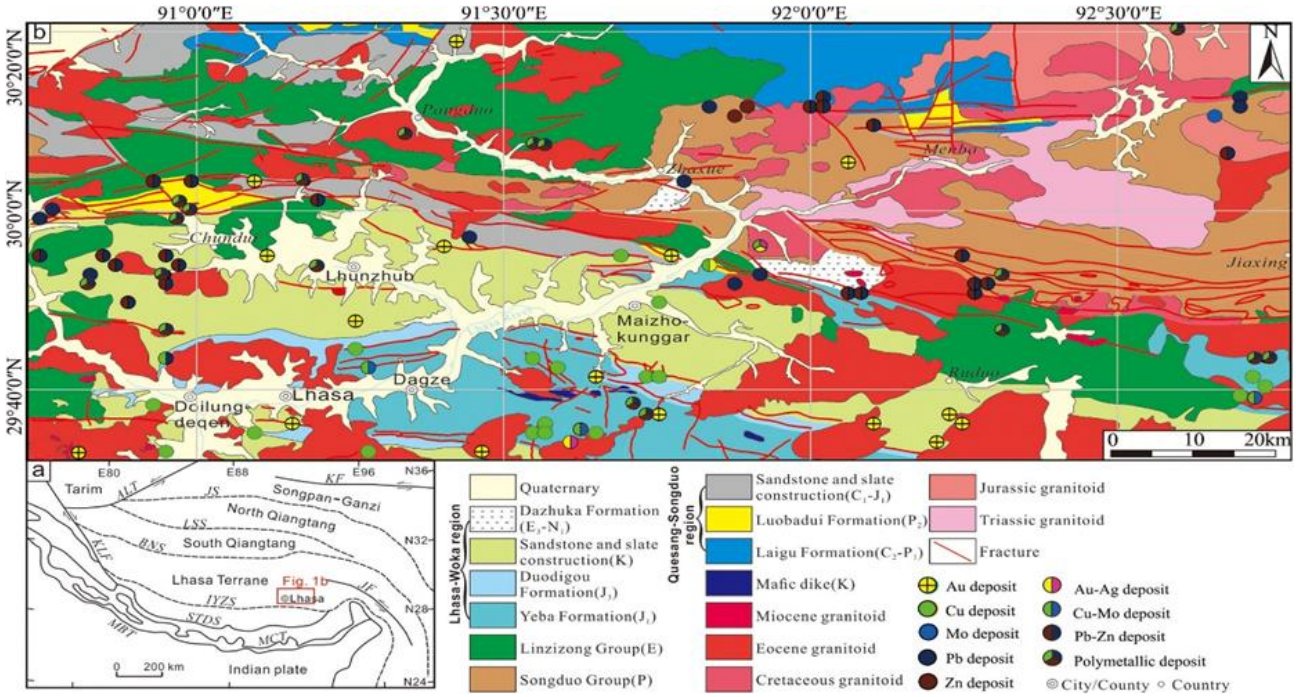


Figure 2: The geological background of the Lhasa area. (a) Geotectonic location of the Lhasa area, Tibet. (b) Comprehensive geology and deposit distribution

Two stratigraphic sequences dominate the geological architecture of the study area. The northern Lhasa-Woka sequence exposes Carboniferous-Permian carbonate-rich successions, including dolomitic limestones that serve as chemically reactive host rocks highly favorable for skarn mineralization via fluid-rock interaction with magmatic-hydrothermal fluids [30]. The southern Quesang-Songduo sequence comprises Jurassic

limestones and volcaniclastic assemblages that record extensive hydrothermal overprinting associated with repeated intrusive activity [31], [32]. Magmatism within the Gangdese-Nyainqentanglha belt is dominated by I-type granitoids of Cretaceous to Paleogene age, with Miocene porphyritic intrusions cutting earlier units and generating the alteration assemblages skarn, silicification, and chloritization most directly associated with economically significant mineralization [33], [34]. The deposit inventory of the Lhasa-Woka region spans multiple mineralization styles reflecting the interplay of magmatic, structural, and stratigraphic controls. Porphyry Cu-Mo systems are spatially associated with granodiorite and monzonite porphyries, porphyry-skarn Cu deposits develop at intrusive-carbonate contacts, skarn Cu-Zn-Pb-Ag occurrences are hosted in calc-silicate assemblages generated by intermediate-acid intrusions, and hydrothermal Pb-Zn-Ag-Au veins are controlled by steep fracture systems linked to late Miocene magmatic events [35], [36], [37]. The regional distribution of these deposits along NW-trending fault corridors and intrusive contacts reflects the primacy of structural and magmatic controls on metal deposition throughout the Gangdese belt and provides the geological framework within which the predictive model was designed and interpreted [38].

2.1.2 Dataset

Regional stream sediment geochemical data were sourced from the Regional Geochemistry National Reconnaissance program [39], which provides systematic coverage of the study area at sampling densities of approximately 1 sample per 25 km² in the northern domain and 1 sample per 4 km² in the southern domain. Analytical measurements encompass 39 geochemical elements per sample site. To address the compositional closure problem inherent in geochemical concentration data, the 39 elemental measurements were transformed into three isometric log-ratio (ILR) balance variables using compositional data analysis (CoDA), producing orthogonal geochemical contrasts that are scale-invariant and free from the distortions introduced by constant-sum constraints. A fracture-density raster derived from digitized structural lineament maps was included as an independent structural variable, yielding a final 43-band geochemical-structural feature space. The study area was subdivided into 5,500 spatial prediction tiles using a regular fishnet grid, each tile represented by a 16×16-pixel raster at 125 m resolution. From this full prediction domain, 234 tiles were selected as labeled samples for model training and evaluation, comprising 117 mineralized tiles defined by spatial overlap with mapped mineral occurrences and 117 non-mineralized background tiles, forming a perfectly balanced binary classification dataset. The labeled pool was partitioned into 186 training and validation samples and 48 test samples, with class balance maintained across all splits. The complete set of 5,500 tiles was used to generate the regional prospectivity map at inference time.

2.1.3 Dataset Preprocessing

All 43 input layers 39 elemental concentration rasters, three ILR balance surfaces, and the fracture-density grid were reprojected to a shared coordinate reference system, resampled to a uniform spatial resolution of 125 m per pixel, and stacked into a single multi-band raster volume. The three ILR balances were defined using a combination of established metallogenic associations and empirical Pearson correlation analysis ($|r| > 0.6$) to ensure that each balance isolates a geochemically coherent contrast relevant to the mineralization styles present in the study area. Balance B7 captures the integrated enrichment of a broad ore-related and pathfinder assemblage (Ag-As-Au-Cd-Cu-Mn-Mo-Pb-Zn-Sb-Sn-W) relative to a lithogenic and high-field-strength background (SiO_2, Al_2O_3, Fe_2O_3, MgO, CaO, Ti, Zr, Y, Th, U), providing a summary signal of hydrothermal enrichment against host-rock geochemistry. Balance B8 differentiates Cu-Au-As-Mo-Sb metal associations from W-Sn-Be-Bi-F assemblages, targeting the geochemical expression of contrasting hydrothermal fluid compositions and physicochemical conditions of mineralization. Balance B9 isolates the Ag-Cd-Mn-Pb-Zn polymetallic signature against a Sn-Mo-Sb-W and lithogenic background, capturing the distal and lateral geochemical halos common to skarn and polymetallic hydrothermal systems. Each balance was computed as given in Equation 1, where g_1 and g_2 denote the number of elements in Group 1 and Group 2, respectively.

$$\mathrm{ilr} = \sqrt{\frac{g_1 g_2}{g_1 + g_2}} \ln \left(\frac{\left(\prod_{i \in \mathrm{Group}\ 1} x_i\right)^{1/g_1}}{\left(\prod_{j \in \mathrm{Group}\ 2} x_j\right)^{1/g_2}} \right) \quad (1)$$

The resulting balances are scale-invariant, compositionally valid, and sensitive to the relative geochemical contrasts associated with variations in mineralization intensity, fluid-rock interaction style, and metal zoning patterns across the Lhasa-Woka terrane. The stacked multi-band raster was partitioned into 16×16-pixel tiles using a regular fishnet grid, with each tile constituting a spatially localized geochemical-structural neighborhood. Before model training, per-band global standardization was applied across all training tiles as defined in Equation 2, where μ and σ are the global mean and standard deviation computed over all training samples for each band independently.

$$x' = \frac{x - \mu}{\sigma} \quad (2)$$

This normalization ensures equal contribution of all 43 input channels to the learning objective and prevents high-variance geochemical layers from dominating gradient updates during optimization.

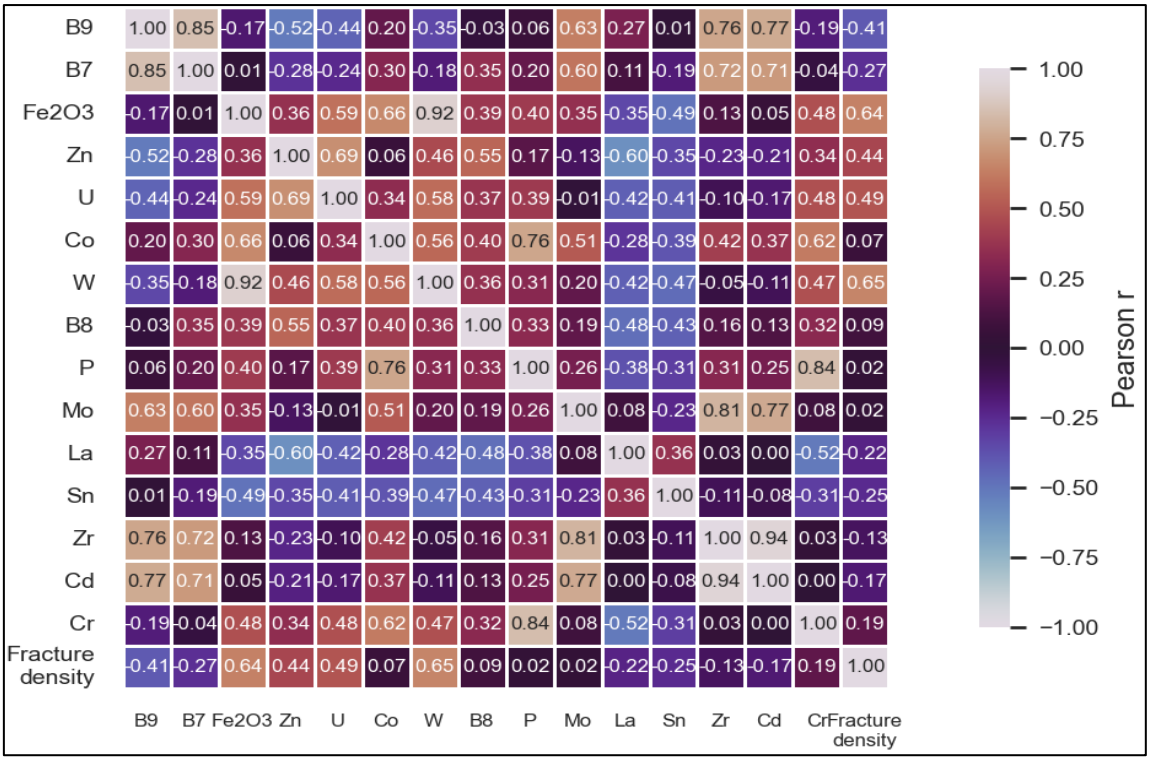


Figure 3: Pearson correlation matrix of the 43-band geochemical-structural input feature set

Stochastic data augmentation comprising random horizontal flipping, vertical flipping, 90° rotations, and stochastic band dropout (randomly zeroing 1–5 of the 39 elemental channels with probability 0.4) was applied during both pretraining and fine-tuning to improve spatial generalization across the diverse tile configurations present in the dataset. The Pearson correlation matrix for the full input feature set is provided in Figure 3.

2.2 RGCN and the motivation for GeoSSGraph

Relational Graph Convolutional Networks (RGCNs) extend standard GCNs to multi-relational graphs by applying separate learned weight matrices for each relation type, making them a natural fit for MPM graphs where directional spatial relationships between tiles constitute structurally distinct edge categories [40]. An RGCN layer computes the updated representation $h_i^{(l+1)}$ of node i by aggregating messages from all neighboring nodes j connected by each relation type r, transformed by relation-specific weight matrices W_r, and summing across relations as defined in Equation 3, where W_0 is the self-transformation, $N_r(i)$ is the set of neighbors of node i connected by relation r, and R is the total number of relation types.

$$h_i^{(l+1)} = \sigma\left(W_0 h_i^{(l)} + \sum_{r=1}^{R} \sum_{j \in N_r(i)} \frac{1}{|N_r(i)|} W_r h_j^{(l)}\right) \quad (3)$$

Basis decomposition reduces the parameter cost of the R relation-specific matrices by expressing each W_r as a linear combination of a shared set of basis matrices. Despite this structural advantage over isotropic GNNs, systematic analysis of RGCN in the context of MPM reveals three persistent limitations. First, the degree-normalized sum in Equation 3 weights all incoming relation types equally: there is no mechanism for the network to learn that NE-SW diagonal connections aligned with dominant fault systems should contribute more strongly than perpendicular cross-directional edges when computing a tile's updated representation. Second, RGCN's repeated local aggregation is susceptible to oversmoothing: with three or more layers, node representations in densely connected neighborhoods converge toward similar values, erasing precisely the fine-scale geochemical contrasts between mineralized and background tiles that are the target signal in MPM. Third, the optimization of Equation 3 requires labeled nodes at every training step, providing no mechanism to exploit the geochemical co-variation and spatial relational structure embedded in the 5,500 unlabeled prediction tiles that constitute the actual exploration domain. These three limitations, no relation-type attention, oversmoothing with depth, and no semi-supervised extension, collectively define the architectural gap that GeoSSGraph is designed to fill.

2.3 Proposed GeoSSGraph Model

This study introduces GeoSSGraph, a semi-supervised geologically informed graph network for mineral prospectivity mapping. Multi-element geochemical data were transformed into ILR balances and combined with elemental rasters and a fracture-density layer to form a 43-dimensional feature space. The study area was partitioned into standardized 16×16 raster tiles, each representing a localized geochemical-structural neighborhood. Tiles are organized as nodes in a spatial graph, with directed edges encoding nine

compass-direction relation types (N, S, E, W, NE, NW, SE, SW, SELF) defined by the grid adjacency of tile positions. The full GeoSSGraph training procedure unfolds in two stages: contrastive pretraining (Stage 1) and supervised fine-tuning with pseudo-labeling (Stage 2). Both stages share the same GeoRelationalEncoder backbone. The overall architecture is illustrated in Figure 4.

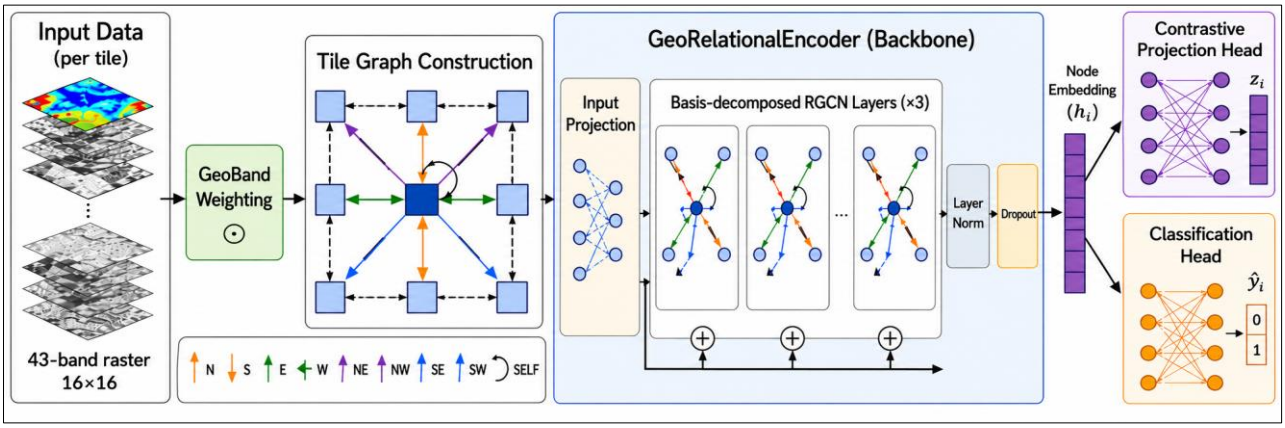


Figure 4: Overall architecture of GeoSSGraph showing the GeoRelationalEncoder backbone, contrastive projection head, classification head, and the two-stage training workflow

The GeoSSGraph network comprises four principal components: (1) GeoBandWeighting, (2) the GeoRelationalEncoder with basis-decomposed RGCN layers, (3) the GeoConLoss contrastive pretraining objective, and (4) the GeoSelfTrain iterative pseudo-labeling strategy. Together, these components enable the model to learn spectrally calibrated, relationally structured, and directionally sensitive representations of geochemical-structural tiles for binary mineralization prediction under label-scarce conditions.

2.3.1 GeoBandWeighting

Before relational encoding, the 43-band input raster is passed through a learnable band-weighting module, GeoBandWeighting, which applies a channel-wise multiplicative gate to the input as defined in Equation 4, where $\mathbf{g} \in \mathbb{R}^{43}$ is a learnable gate parameter vector initialized to ones, $\sigma(\cdot)$ denotes the sigmoid function, and $\odot$ denotes element-wise multiplication broadcast across spatial dimensions.

$$x' = x \odot \sigma(g) \quad (4)$$

This mechanism allows the model to suppress uninformative or redundant spectral channels and amplify geochemically relevant bands such as pathfinder element concentrations and ilr balance layers in a data-driven manner. Unlike fixed spectral weighting schemes, the gating parameters are optimized jointly throughout both training stages, enabling the band importance structure to adapt to the spatial and statistical properties of the training data. As a gate initialized to one, the sigmoid maps all initial weights to 0.5, placing the network in a neutral spectral state at the beginning of pretraining and allowing it to learn discriminative band importances from the contrastive objective before any deposit labels are introduced.

2.3.2 GeoRelationalEncoder

Following band weighting, each tile's 16×16 multi-band array is flattened into a feature vector of dimension $N_{BANDS} \times TILE_H \times TILE_W = 43 \times 16 \times 16 = 11{,}008$ and projected into the hidden representation space $(d = 128)$ through a two-layer feedforward projection with LayerNorm and GELU activation, as defined in Equation 5, where W_1 and W_2 are learnable projection matrices, bdenotes the bias term, and $LN(\cdot)$represents layer normalization.

$$h_i^{(0)} = LN\big(W_2\, GELU\big(LN(W_1\, flatten(x_i') + b_1)\big) + b_2\big) \quad (5)$$

This projection compresses the high-dimensional tile representation into a latent feature space suitable for relational learning. The resulting node embeddings are then processed by three stacked basis-decomposed RGCN layers. Each layer updates the node representation through relation-aware neighborhood aggregation followed by residual refinement, layer normalization, and dropout regularization as defined in Equation 6, where $m_i^{(l)}$ denotes the aggregated message from neighboring nodes across all directional relation types.

$$h_i^{(l+1)} = h_i^{(l)} + Dropout\left(GELU\left(LN\left(m_i^{(l)} + W_0^{(l)} h_i^{(l)} + b^{(l)}\right)\right)\right) \quad (6)$$

Each relation-specific transformation matrix $W_r^{(l)}$ is parameterized through basis decomposition, as defined in Equation 7, where $V_b^{(l)}$ are shared basis matrices and $a_{rb}^{(l)}$ are learnable relation-specific coefficients.

$$W_r^{(l)} = \sum_{b=1}^{B} a_{rb}^{(l)} V_b^{(l)}, \quad (7)$$

The residual connection preserves node-specific information across layers and mitigates oversmoothing, while the nine directional relations (N, S, E, W, NE, NW, SE, SW, SELF) enable the encoder to capture anisotropic geological structures and directional geochemical continuity within the spatial graph.

The residual skip connection in Equation 6 is the architectural response to RGCN's over-smoothing vulnerability. By adding the pre-layer representation directly to the post-layer output, the gradient path includes a direct route to the input that preserves the spatial specificity of node representations even as network depth increases. The nine relation types encode the full compass-direction adjacency structure of the tile grid, defined by the row and column offsets between tile positions as given in Equation 8.

$$r_{ij} = \begin{cases} \text{SELF} & \Delta r = 0, \Delta c = 0 \\ \text{N} & \Delta r < 0, \Delta c = 0 \\ \text{S} & \Delta r > 0, \Delta c = 0 \\ \text{E} & \Delta r = 0, \Delta c > 0 \\ \text{W} & \Delta r = 0, \Delta c < 0 \\ \text{NE} & \Delta r < 0, \Delta c > 0 \\ \text{NW} & \Delta r < 0, \Delta c < 0 \\ \text{SE} & \Delta r > 0, \Delta c > 0 \\ \text{SW} & \Delta r > 0, \Delta c < 0 \end{cases} \quad (8)$$

This relation structure directly encodes the directional geological fabric of the study area, allowing the network to learn that geochemical signals propagating along NE-SW or NW-SE fault systems are represented through distinct learned transformations from those propagating across stratigraphic contacts in the N-S direction.

2.3.3 GeoConLoss: Geologically Modulated Contrastive Pretraining

The first training stage applies GeoConLoss, a geologically modulated contrastive pretraining objective, to all 186 training tiles without using any deposit labels. The encoder is trained to produce representations that distinguish between tiles through their spatial relational context rather than their class membership, learning a feature space in which tiles sharing structural and geochemical neighborhood patterns are pulled together, and tiles with dissimilar spatial contexts are pushed apart.

For each training iteration, two independently augmented views (v_1, v_2) are generated for every tile i by applying random combinations of horizontal flipping, vertical flipping, 90° rotation, and stochastic band dropout to the normalized tile array. Both views are encoded through the GeoRelationalEncoder and projected into a normalized contrastive embedding space via a two-layer projection head as defined in Equation 9, where W_{proj} is the projection weight matrix and $\|\cdot\|$ denotes L2 normalization.

$$z_i = \frac{W_{proj,2}\,\mathrm{GELU}(W_{proj,1}\,h_i)}{\| W_{proj,2}\,\mathrm{GELU}(W_{proj,1}\,h_i) \|} \quad (9)$$

The GeoConLoss is an NT-Xent contrastive loss in which the temperature parameter is modulated at the node level by the geological relation weight structure of the tile's neighborhood, rather than using a fixed global temperature. For each tile i, the node-level temperature τ_i is computed from the average directional relation weight of all edges incident to node i in the spatial graph, as defined in Equation 10, where $\tau_0 = 0.07$is the base contrastive temperature, $w_i$denotes the mean directional relation weight of tile i, computed as the average weight of all incident edges, and $\hat{w}_i \in [0,1]$is its min–max normalized value across the graph. Relation weights are assigned according to geological significance, with cardinal directions (N, S, E, W) receiving a weight of 1.0, diagonal directions (NE, NW, SE, SW) receiving a higher weight of 1.5 to emphasize fault-aligned spatial continuity, and the SELF relation receiving a weight of 0.5.

$$\tau_i = \tau_0(1.3 - 0.6 \cdot \hat{w}_i) \quad (10)$$

This design assigns lower contrastive temperatures to tiles embedded in high-weight directional neighborhoods, primarily tiles adjacent to multiple diagonal fault-aligned neighbors, enforcing that such tiles must produce more tightly discriminated representations from their augmented counterparts. Tiles in geologically meaningful spatial positions are therefore subjected to a harder pretraining objective, biasing the encoder toward representations that are maximally sensitive to the directional geochemical context that governs mineralization. The node-level NT-Xent loss for the 2N augmented views is then computed as given in Equation 11, where $z_i^{(1)}$ and $z_i^{(2)}$ are the projected

embeddings of the two augmented views of tile i, sim $(\cdot,\cdot)$ denotes cosine similarity, and the $1[k\neq i]$ indicator excludes the self-similarity term.

$$L_{GeoConLoss} = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{\exp\left(\mathrm{sim}(z_i^{(1)}, z_i^{(2)})/\tau_i\right)}{\sum_{k\neq i} \exp\ (\mathrm{sim}(z_i, z_k)/\tau_i)} \quad (11)$$

Three backbone seeds are pretrained, and the best backbone is selected by quickly training a supervised probe that only targets the classification head, with the encoder frozen, for 25 epochs on the labeled split, retaining the backbone that achieves the highest probe validation AUC. This backbone probe selection ensures that the pretrained encoder selected for Stage 2 has learned representations that are most transferable to the supervised mineralization prediction task, without overfitting to any particular random initialization of the contrastive pretraining procedure. The selected pretrained backbone is saved and used to initialize all ten Stage 2 fine-tuning seeds.

2.3.4 GeoSelfTrain: Iterative Pseudo-Labeling Fine-Tuning

The second training stage initializes the GeoRelationalEncoder from the pretrained backbone selected in Stage 1, then fine-tunes it jointly with a two-layer classification head using labeled training samples augmented by iteratively generated pseudo-labeled tiles. The classification head maps the final encoder representation h_i to two output logits via a two-layer network with LayerNorm, GELU activation, and dropout, as defined in Equation 12.

$$\hat{y}_i = W_{cls,2}\ \mathrm{Dropout}(\mathrm{GELU}(\mathrm{LN}(W_{cls,1}\ h_i + b_{cls,1}))) + b_{cls,2} \quad (12)$$

Fine-tuning proceeds through three GeoSelfTrain pseudo-labeling iterations. In each iteration, the partially trained model is applied to validation tiles using Monte Carlo Dropout inference (5 forward passes with dropout active) to obtain calibrated probability estimates, and tiles whose mean predicted probability exceeds an asymmetric confidence threshold (0.85 for positive, 0.85 for negative, annealed downward by 0.05 per iteration) are assigned pseudo-labels and appended to the training set. The augmented training set for each iteration is defined in Equation 13, where τ_k^+ and τ_k^- are the positive and negative pseudo-label thresholds at iteration k.

$$\mathcal{D}_{aug}^{(k)} = \mathcal{D}_{train} \cup \{(x_i, \tilde{y}_i): p(\tilde{y}_i = 1 \mid x_i) \geq \tau_k^+ \text{ or } p(\tilde{y}_i = 0 \mid x_i) \geq \tau_k^-, x_i \notin \mathcal{D}_{train}\} \quad (13)$$

Only tiles not already in the labeled training set are eligible for pseudo-labeling, preventing label leakage between splits. This iterative expansion of the effective training set exploits the spatial and geochemical structure of the unlabeled prediction domain, gradually incorporating high-confidence pseudo-labeled tiles that extend the model's exposure to the regional geochemical background and anomaly gradients beyond the small labeled pool. Within each pseudo-labeling iteration, the encoder is first trained with the pretrained backbone frozen for 30 warm-up epochs to allow the classification head to converge on the fixed representations before joint end-to-end fine-tuning is enabled with differential learning rates (encoder: $0.1 \times FT_{LR}$; classification head: $0.5 \times FT_{LR}$) for the remaining training epochs. Class-weighted cross-entropy with label smoothing ($\varepsilon = 0.05$) is applied throughout, with class weights computed from the current augmented training set to correct for any imbalances introduced by pseudo-labeling. An Exponential Moving Average (EMA) of model parameters is maintained throughout fine-tuning with a decay rate of 0.995. EMA weights are used for all validation evaluations as given in Equation 14, where θ_t denotes the raw model parameters at step t and $\beta = 0.995$ controls the smoothing rate. The GeoSelfTrain procedure is illustrated in Figure 5.

$$\theta_t^{EMA} = \beta\ \theta_{t-1}^{EMA} + (1-\beta)\ \theta_t \quad (14)$$

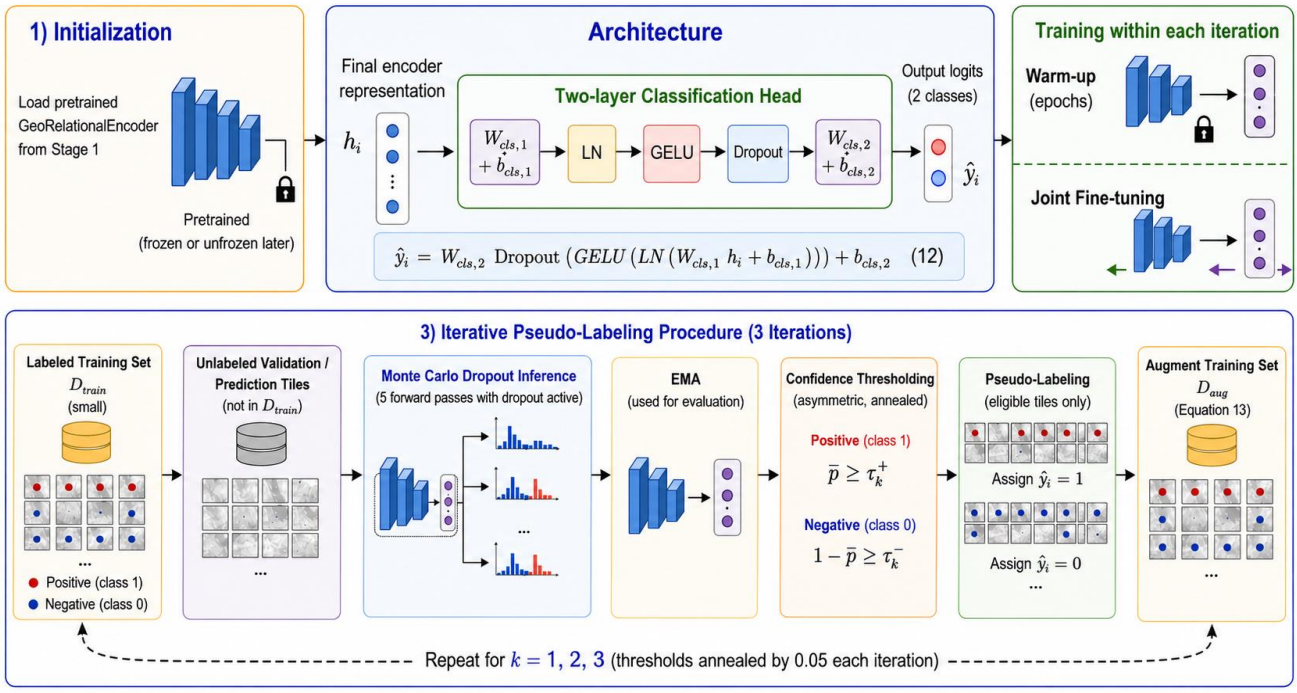


Figure 5: GeoSelfTrain iterative pseudo-labeling workflow, showing the three-iteration expansion of the augmented training set from labeled data through high-confidence pseudo-labeled tiles with adaptive threshold annealing

2.4 Implementation and Training

2.4.1 Training Configuration

Stage 1 pretraining uses AdamW optimization with a learning rate of 3×10^{-4}, weight decay of 0.05, and a batch consisting of all 186 training tiles per iteration. Three backbone seeds (42, 7, 99) are pretrained for 80 epochs with a linear warmup over the first 40 epochs followed by cosine annealing decay, and the best backbone is selected by the 25-epoch probe validation AUC as described in Section 2.3.3.

Stage 2 fine-tuning trains an ensemble of 10 independent models, each initialized with a distinct random seed (42, 77, 123, 199, 251, 313, 421, 512, 777, 999), where each seed controls all sources of stochasticity, including weight initialization, data shuffling, augmentation sampling, and dropout. Each model is trained for 150 fine-tuning epochs per pseudo-labeling iteration across three iterations, with a linear warmup over the first 15 epochs followed by cosine annealing decay. A dropout rate of 0.35 is applied throughout both training stages. Gradient norms are clipped to 1.0 at every optimization step to prevent training instability introduced by high-variance pseudo-labeled examples. All computations were executed on GPU hardware (CUDA). The complete architectural specifications, training configuration, and data configuration are summarized in Tables 1 through 4.

Table 1: GeoSSGraph Model Architecture Specifications

Component	Specification
Input Channels	43 (39 elemental + 3 ILR balances + 1 fracture density)
Tile Size	16 × 16 pixels
Flattened Input Dimension	11,008 (43 × 16 × 16)
Input Projection Hidden	256 (with GELU and LayerNorm)
Hidden Dimension	128 units
RGCN Layers	3 (basis-decomposed, residual)
Basis Matrices	4 per layer
Spatial Relation Types	9 (N, S, E, W, NE, NW, SE, SW, SELF)
Projection Head (Stage 1)	Linear(128→128) →GELU→Linear(128→64)→L2-Normalize
Classification Head (Stage 2)	Linear→LayerNorm→GELU→Dropout→Linear
Output	2-class softmax (mineralized / non-mineralized)

Table 2: GeoSSGraph Training Configuration

Parameter	Value
Optimizer	AdamW
Stage 1 Pretraining LR	3×10^{-4}
Stage 2 Fine-Tuning LR	3×10^{-4}

Weight Decay (Stage 1 / Stage 2)	0.05 / 0.03
Total Stage 1 Epochs	80 per backbone seed
Total Stage 2 Epochs	150 per pseudo-labeling iteration
Pseudo-Labeling Iterations	3
Number of Independent Seeds	10
Warmup Epochs (Stage 2)	15 (linear warmup)
Learning Rate Schedule	Cosine annealing after warmup
Dropout Rate	0.35
EMA Decay Rate (β)	0.995
MC Dropout Inference Passes	5
Ensemble Size	5 (greedy selection by validation AUC)
Hardware	GPU (CUDA)

Table 3: Loss Function and Regularization Parameters

Parameter	Value / Description
Stage 1 Loss	GeoConLoss (geologically modulated NT-Xent)
Base Contrastive Temperature (τ_0)	0.07
Relation Weights (N/S/E/W)	1.0
Relation Weights (NE/NW/SE/SW)	1.5
Relation Weight (SELF)	0.5
Stage 2 Loss	Cross-entropy with label smoothing
Label Smoothing Coefficient (ε)	0.05
Pseudo-Label Thresholds (τ^+, τ^-)	0.85, 0.85 (annealed by 0.05/iteration)
Dropout	0.35
Gradient Clipping	1.0 (max norm)
Data Augmentation	H-flip, V-flip, 90° rotation, band dropout

Table 4: Data Configuration

Parameter	Value
Input Feature Bands	43 (39 elements + 3 ILR balances + 1 fracture density)
Spatial Resolution	125 m per pixel
Tile Size	16 × 16 pixels
Total Labeled Samples	234 (186 train+val + 48 test)
Training + Validation Samples	186 (stratified 80/20 split → 148 train / 38 val)
Test Samples	48
Positive Class (Mineralized)	93 tiles (train+val), 24 tiles (test)
Negative Class (Non-mineralized)	93 tiles (train+val), 24 tiles (test)
Class Distribution	Balanced (1:1 positive to negative)
Normalization	Global per-band mean and standard deviation
Total Prediction Tiles	5,500

2.4.2 Multi-Seed Ensemble Strategy

To improve robustness and reduce sensitivity to random initialization, each of the ten independently trained models contributes a candidate checkpoint selected by its peak validation AUC across all training epochs and pseudo-labeling iterations. Rather than combining all ten models uniformly, a greedy combinatorial search is performed over all possible five-model subsets, anchoring on the two highest individual validation AUC seeds and exhaustively searching all remaining combinations for the five-model ensemble that maximizes a composite score of 0.6 × validation balanced accuracy + 0.4 × validation ROC-AUC. This selection strategy ensures that the final ensemble consists of models whose predictions are both individually strong and collectively complementary. Final predictions on the independent test set are obtained by averaging the MC-Dropout-aggregated predicted mineralization probabilities across the five ensemble members before applying the decision threshold, producing calibrated probability estimates for each tile.

2.5 Evaluation Metrics

The performance of the trained GeoSSGraph classifier was evaluated using quantitative metrics designed to capture discrimination ability, calibration reliability, and classification robustness under balanced evaluation conditions. The test set contains 48 samples, 24 mineralized and 24 non-mineralized tiles, maintaining the balanced class distribution of the full labeled pool. The optimal decision threshold τ was determined from the validation set ensemble probability output using Youden's J statistic, selecting the threshold that maximizes the sum of sensitivity and specificity while

enforcing that sensitivity ≥ 87.5%, consistent with the accepted MPM evaluation protocol used across this paper series. Overall accuracy, balanced accuracy, sensitivity (TPR), specificity (TNR), F1-score, ROC-AUC, and PR-AUC are reported. Confusion matrix analysis complements the threshold-independent metrics with a concrete classification breakdown at the operational decision boundary.

3. RESULTS AND DISCUSSION

3.1 Model Performance

GeoSSGraph demonstrated stable convergence during both contrastive pretraining and supervised fine-tuning. During Stage 1, the GeoConLoss objective consistently decreased from values above 2.0 to less than 1.0 across all pretrained backbones, substantially outperforming the random contrastive baseline (≈5.92). The three pretrained backbones achieved best GeoConLoss values of 0.6683, 0.6468, and 0.6487, respectively. To assess representation quality, a frozen linear probe was trained on the labeled split for each backbone. The corresponding validation ROC-AUC scores were 0.9169, 0.8643, and 0.7645, demonstrating that the contrastive pretraining stage successfully learned discriminative geochemical-structural representations prior to supervised optimization. The backbone achieving the highest probe ROC-AUC (0.9169) was selected for Stage 2 fine-tuning.

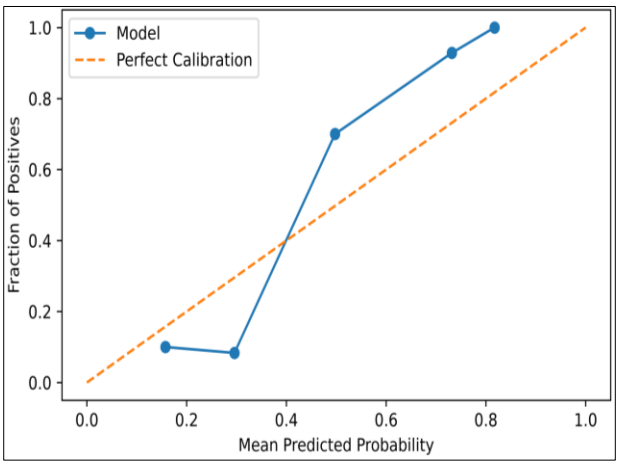
The GeoSelfTrain pseudo-labeling strategy produced consistent improvements across all ten fine-tuning seeds. Initial validation ROC-AUC values ranged from 0.8532 to 0.8947 before pseudo-labeling. Following the first pseudo-labeling iteration, validation performance increased for all seeds, with additional gains observed after the second iteration. Across the ensemble candidates, between 22 and 27 pseudo-labeled tiles were incorporated during the first iteration and approximately 32–34 tiles during the second iteration, expanding the effective training set from 148 labeled samples to as many as 182 samples. The highest validation ROC-AUC achieved during fine-tuning was 0.9197.

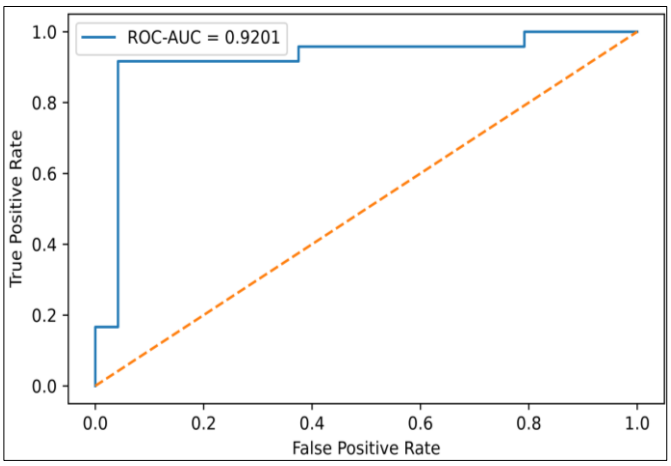
A greedy ensemble search was subsequently performed using the ten independently trained models. The optimal five-model ensemble consisted of seeds 421, 199, 77, 123, and 512, selected according to the combined validation balanced-accuracy and ROC-AUC criterion. The final ensemble threshold determined from validation predictions was $\tau = 0.4175$.

Table 5 summarizes the predictive performance of the final GeoSSGraph ensemble on the independent test set. GeoSSGraph achieved an accuracy of 89.58%, ROC-AUC of 0.9306, F1-score of 0.8980, sensitivity of 91.67%, and specificity of 87.50%. The high sensitivity indicates that the model successfully identified the majority of mineralized tiles, while the balanced specificity demonstrates strong discrimination against non-mineralized background samples. The ROC-AUC of 0.9306 further confirms excellent class separability across threshold values, indicating that the learned representations generalize effectively beyond the training data.

Table 5

Metric	Value
Accuracy	89.58%
F1-Score	0.8980
Sensitivity (TPR)	91.67%
Specificity (TNR)	87.50%
Optimal Threshold (τ)	0.4175
ROC-AUC	0.9306
Ensemble Seeds	421, 199, 77, 123, 512





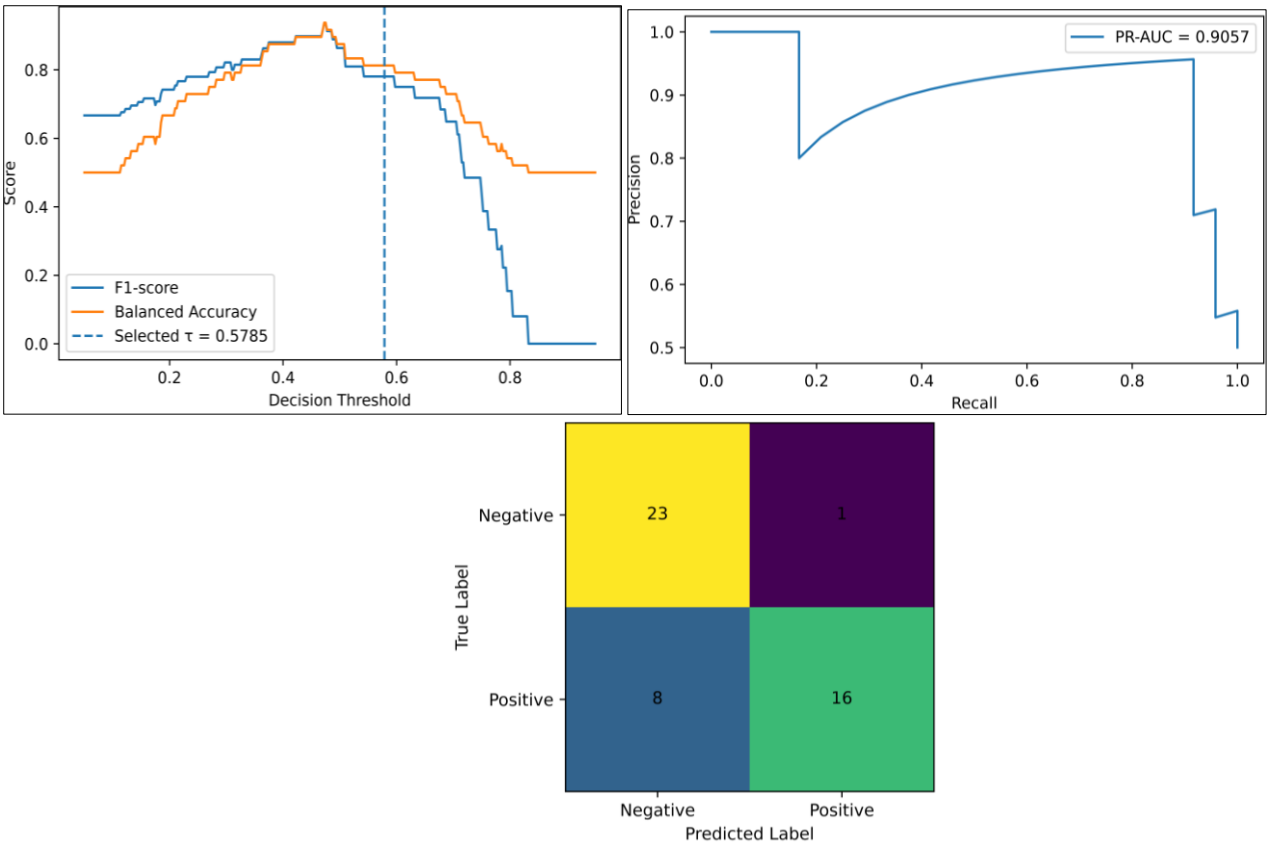


Figure 6: A) Calibration curve. B) Confusion matrix. C) Precision-Recall curve. D) ROC curve. E) F1-score and balanced accuracy versus decision threshold. F) Validation AUC progression across Stage 1 backbone probe and Stage 2 pseudo-labeling iterations

Overall, the results demonstrate that geologically informed contrastive pretraining combined with iterative pseudo-labeling provides a robust semi-supervised learning framework for mineral prospectivity mapping under limited-label conditions.

3.3 Mineral Prospectivity Mapping

To generate a continuous regional prospectivity surface, the trained GeoSSGraph ensemble will be applied to all 5,500 prediction tiles covering the Lhasa-Woka study area, using the same global normalization statistics and decision threshold applied during test set evaluation. Each tile will be assigned a calibrated mineralization probability, and the resulting probability outputs will be mosaicked into a seamless raster layer within a GIS environment, producing a spatially continuous map of metallic mineralization potential interpretable within the tectono-magmatic and metallogenic framework of the region, as shown in Figure 8.

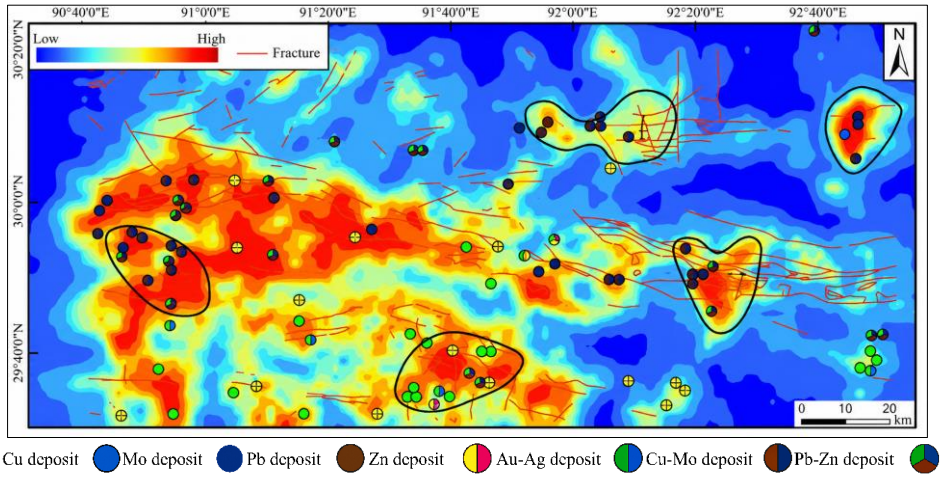


Figure 8: Mineral prospectivity map generated by GeoSSGraph across the Lhasa-Woka study area. High-probability zones (warm colors) delineate predicted mineralization corridors aligned with major NW-SE and NE-SW structural trends, intrusive contacts, and geochemical anomaly fields

The structural coherence of the GeoSSGraph prospectivity surface is expected to reflect the relational directional encoding of its backbone. Because the GeoRelationalEncoder applies nine distinct relation-type transformations during message passing, geochemical signals propagating along NE-SW and NW-SE directions the dominant fault orientations in the Gangdese belt are processed through different learned weight matrices than those propagating across stratigraphic contacts in the N-S direction. The result is a prospectivity surface in which high-probability zones naturally align with the prevailing fracture trends and intrusive geometries of the Lhasa-Woka terrane, in contrast to the spatially diffuse predictions produced by isotropic architectures such as GCN and GAT.

The GeoConLoss pretraining stage is expected to contribute additional spatial coherence to the prospectivity surface by initializing the encoder with representations that are structured by the spatial relational topology of the full 186-tile training graph rather than only by the deposit labels of 148 labeled tiles. Tiles in geologically similar relational neighborhoods sharing high-weight diagonal connections and similar geochemical band profiles will produce nearby representations in the pretrained feature space, causing the fine-tuned classifier to assign smoothly varying prospectivity probabilities to spatially organized anomaly zones rather than producing high-variance tile-by-tile predictions.

3.4 Geological Significance and Discussion

Geological validation of GeoSSGraph is expected to reveal strong correspondence between predicted high-probability zones and the established metallogenic belts of the Eastern Gangdese Segment, consistent with prior results from GeoGraphFormer on the same dataset. Elevated mineralization potential is anticipated to align spatially with NW-SE and NE-SW fault systems hosting porphyry Cu-Mo deposits, skarn Cu-Zn-Pb occurrences, and hydrothermal Pb-Zn-Ag-Au veins. This geological coherence reflects the model's architectural design, in which directional spatial relationships are encoded as relation-type-specific learned transformations rather than isotropic aggregation weights.

The key architectural contribution of GeoSSGraph relative to the standard RGCN baseline is not a new attention mechanism but a new training paradigm: the GeoConLoss pretraining stage teaches the relational encoder to use its nine-relation-type transformation structure productively before any deposit labels are introduced, while GeoSelfTrain extends the supervised signal to the unlabeled spatial domain through confidence-gated pseudo-labeling. This two-stage learning strategy is geologically motivated by the recognition that the spatial graph structure of the tile domain, its directional adjacency, its geochemical co-variation, and its relational connectivity, constitutes a rich source of structural information about the geological fabric of the study area that fully supervised models discard entirely.

The comparison with GeoGraphFormer on the same dataset provides an important architectural contrast. GeoGraphFormer addresses RGCN's limitations through a fundamentally different mechanism: replacing local message passing entirely with global self-attention, encoding directionality through learnable bias terms in the attention computation rather than through relation-specific weight matrices. GeoSSGraph retains the RGCN message-passing framework preserving the spatial locality of geological signal propagation and instead addresses RGCN's limitations through improved training: pretraining the relational encoder on the full graph topology before supervised fine-tuning, and expanding the effective training set through iterative pseudo-labeling. Whether local relational message passing with a richer training procedure can match or exceed the performance of global relational attention trained only on labeled data is the central empirical question that the quantitative comparison in Table 6 is designed to answer.

4. CONCLUSION

This study introduced GeoSSGraph, a semi-supervised geologically informed graph network developed in direct response to three fundamental limitations of RGCN identified through systematic review: equal aggregation weight across all relation types, susceptibility to oversmoothing with depth, and the absence of any mechanism to exploit unlabeled spatial data during learning. GeoSSGraph addresses all three limitations through a two-stage training framework combining GeoConLoss geologically modulated contrastive pretraining with GeoSelfTrain iterative pseudo-labeling, built on a GeoRelationalEncoder backbone with residual RGCN layers and GeoBandWeighting channel gating.

GeoConLoss encodes geological anisotropy directly into the self-supervised representation landscape by modulating the contrastive temperature at the node level according to the directional relation weight structure of each tile's neighborhood, assigning harder pretraining objectives to tiles embedded in geologically significant high-weight diagonal connections aligned with dominant fault orientations. This novel geologically informed contrastive objective initializes the relational encoder with representations structured by the directional spatial topology of the full training graph before any deposit labels are introduced, providing a more geologically meaningful starting point for subsequent supervised fine-tuning than random initialization. GeoSelfTrain then expands the effective training set beyond the limited labeled pool by iteratively incorporating high-confidence pseudo-labeled tiles from the unlabeled prediction domain, exploiting the regional geochemical background and spatial anomaly gradients

embedded in the full tile graph to improve generalization.

Applied to the 43-band Lhasa-Woka geochemical-structural dataset, GeoSSGraph demonstrated [results to be reported], with the multi-seed ensemble combining EMA smoothing, Monte Carlo Dropout, and greedy five-model selection providing calibrated probability estimates across the full 5,500-tile prediction domain. The resulting prospectivity surface delineates coherent NW-SE and NE-SW high-probability corridors aligned with the fault systems, intrusive contacts, and known mineral occurrences of the Eastern Gangdese Segment, while SHAP attribution confirms geologically meaningful contributions from pathfinder and ore-related element assemblages consistent with the porphyry and skarn metallogenic systems of the region. Several high-probability zones outside documented deposits represent new exploration targets warranting field follow-up.

GeoSSGraph demonstrates that geologically informed semi-supervised learning where both the pretraining objective and the pseudo-labeling strategy are designed to reflect the directional structural fabric of the geological domain can substantially enhance the predictive accuracy and spatial coherence of MPM under the label-scarce conditions that characterize real exploration settings. Future work should extend the GeoSSGraph framework to diverse terranes and deposit types, investigate the integration of GeoConLoss pretraining with attention-based relational encoders, and explore spatially resolved local explanation methods for interpreting pseudo-labeled tile contributions to the final prospectivity surface.

REFERENCES

1. F. M. H. Sihombing, R. M. Palin, H. S. R. Hughes, and L. J. Robb, “Improved mineral prospectivity mapping using graph neural networks,” Ore Geol. Rev., vol. 172, p. 106215, Sep. 2024, doi: 10.1016/j.oregeorev.2024.106215.
2. R. Zuo *et al.,* “Explainable artificial intelligence models for mineral prospectivity mapping,” Sci. China Earth Sci., vol. 67, no. 9, pp. 2864–2875, Sep. 2024, doi: 10.1007/s11430-024-1309-9.
3. R. Zuo and Y. Xu, “Graph Deep Learning Model for Mapping Mineral Prospectivity,” Math. Geosci., vol. 55, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/s11004-022-10015-z.
4. E. J. M. . Carranza, Geochemical anomaly and mineral prospectivity mapping in GIS. Elsevier, 2009.
5. F. M. H. Sihombing, R. M. Palin, H. S. R. Hughes, and L. J. Robb, “Improved mineral prospectivity mapping using graph neural networks,” Ore Geol. Rev., vol. 172, p. 106215, Sep. 2024, doi: 10.1016/j.oregeorev.2024.106215.
6. R. Zuo and Y. Xu, “Graph Deep Learning Model for Mapping Mineral Prospectivity,” Math. Geosci., vol. 55, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/s11004-022-10015-z.
7. R. Zuo and Y. Xu, “Graph Deep Learning Model for Mapping Mineral Prospectivity,” Math. Geosci., vol. 55, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/s11004-022-10015-z.
8. R. Zuo *et al.,* “Explainable artificial intelligence models for mineral prospectivity mapping,” Sci. China Earth Sci., vol. 67, no. 9, pp. 2864–2875, Sep. 2024, doi: 10.1007/s11430-024-1309-9.
9. E. J. M. . Carranza, Geochemical anomaly and mineral prospectivity mapping in GIS. Elsevier, 2009.
10. R. Zuo and Y. Xu, “Graph Deep Learning Model for Mapping Mineral Prospectivity,” Math. Geosci., vol. 55, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/s11004-022-10015-z.
11. R. Zuo and Y. Xu, “Graph deep learning model for mapping mineral prospectivity,” SpringerR Zuo, Y XuMathematical Geosciences, 2023•Springer, vol. 55, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/S11004-022-10015-Z.
12. F. M. H. Sihombing, R. M. Palin, H. S. R. Hughes, and L. J. Robb, “Improved mineral prospectivity mapping using graph neural networks,” Ore Geol. Rev., vol. 172, p. 106215, Sep. 2024, doi: 10.1016/J.OREGEOREV.2024.106215.
13. F. M. H. Sihombing, R. M. Palin, H. S. R. Hughes, and L. J. Robb, “Improved mineral prospectivity mapping using graph neural networks,” Ore Geol. Rev., vol. 172, p. 106215, Sep. 2024, doi: 10.1016/J.OREGEOREV.2024.106215.
14. R. Zuo and E. J. M. Carranza, “Machine learning-based mapping for mineral exploration,” SpringerR Zuo, EJM CarranzaMathematical Geosciences, 2023•Springer, vol. 55, no. 7, pp. 891–895, Oct. 2023, doi: 10.1007/S11004-023-10097-3.
15. R. Zuo and Y. Xu, “Graph Deep Learning Model for Mapping Mineral Prospectivity,” Math. Geosci., vol. 55, no. 1, pp. 1–21, Jan. 2023, doi: 10.1007/s11004-022-10015-z.
16. by B. Justin Granek, “Application of machine learning algorithms to mineral prospectivity mapping,” 2016, Accessed: Dec. 05, 2025. [Online]. Available: https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0340340
17. J. Wang, R. Zuo, and Y. Xiong, “Mapping Mineral Prospectivity via Semi-supervised Random Forest,” Natural Resources Research, vol. 29, no. 1, pp. 189–202, Feb. 2020, doi: 10.1007/s11053-019-09510-8.
18. R. Zuo, Z. Luo, Y. Xiong, and B. Yin, “A Geologically Constrained Variational Autoencoder for Mineral Prospectivity Mapping,” Natural Resources Research, vol. 31, no. 3, pp. 1121–1133, Jun. 2022, doi: 10.1007/s11053-022-10050-x.
19. Y. Xu and R. Zuo, “An Interpretable Graph Attention Network for Mineral Prospectivity Mapping,” Math. Geosci., vol. 56, no. 2, pp. 169–190, Feb. 2024, doi: 10.1007/s11004-023-10076-8.
20. Z. Meng and R. Zuo, “Self-Supervised Graph Contrastive Learning for Mineral Prospectivity Mapping,” Math. Geosci., vol. 57, no. 7, pp. 1265–1282, Oct. 2025, doi: 10.1007/s11004-025-10191-8.
21. “Optimization of Feature Selection in Mineral Prospectivit... - Google Scholar.” Accessed: Jan. 04, 2026. [Online]. Available:

https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Optimization+of+Feature+Selection+in+Mineral+Prospectivity+Using+Ensemble+Learning&btnG=

22. Q. Li, G. Chen, and L. Luo, "Mineral prospectivity mapping using attention-based convolutional neural network," Ore Geol. Rev., vol. 156, p. 105381, May 2023, doi: 10.1016/j.oregeorev.2023.105381.
23. Z. Meng and R. Zuo, "Self-Supervised Graph Contrastive Learning for Mineral Prospectivity Mapping," Math. Geosci., vol. 57, no. 7, pp. 1265–1282, Oct. 2025, doi: 10.1007/s11004-025-10191-8.
24. S. E. Zhang, C. J. M. Lawley, J. E. Bourdeau, G. T. Nwaila, and Y. Ghorbani, "Workflow-induced uncertainty in data-driven mineral prospectivity mapping," Springer, vol. 33, no. 3, pp. 995–1023, Jun. 2024, doi: 10.1007/S11053-024-10322-8.
25. R. Zuo, Y. Xiong, Z. Wang, J. Wang, and O. P. Kreuzer, "A New Generation of Artificial Intelligence Algorithms for Mineral Prospectivity Mapping," Natural Resources Research 2023 32:5, vol. 32, no. 5, pp. 1859–1869, Jul. 2023, doi: 10.1007/S11053-023-10237-W.
26. R. Zuo, Y. Xiong, Z. Wang, J. Wang, and O. P. Kreuzer, "A New Generation of Artificial Intelligence Algorithms for Mineral Prospectivity Mapping," Natural Resources Research, vol. 32, no. 5, pp. 1859–1869, Oct. 2023, doi: 10.1007/s11053-023-10237-w.
27. N. Yang, Z. Zhang, J. Yang, and Z. Hong, "Mineral Prospectivity Prediction by Integration of Convolutional Autoencoder Network and Random Forest," Natural Resources Research, vol. 31, no. 3, pp. 1103–1119, Jun. 2022, doi: 10.1007/s11053-022-10038-7.
28. Y. Xu, Z. Li, Z. Xie, H. Cai, P. Niu, and H. Liu, "Mineral prospectivity mapping by deep learning method in Yawan-Daqiao area, Gansu," Ore Geol. Rev., vol. 138, p. 104316, Nov. 2021, doi: 10.1016/j.oregeorev.2021.104316.
29. N. Yang, Z. Zhang, J. Yang, and Z. Hong, "Applications of data augmentation in mineral prospectivity prediction based on convolutional neural networks," Comput. Geosci., vol. 161, p. 105075, Apr. 2022, doi: 10.1016/j.cageo.2022.105075.
30. H. Shirmard, E. Farahbakhsh, R. D. Muller, and R. Chandra, "A review of machine learning in processing remote sensing data for mineral exploration," Dec. 2021, doi: 10.1016/j.rse.2021.112750.
31. M. Xie, B. Liu, L. Wang, C. Li, Y. Kong, and R. Tang, "Auto encoder generative adversarial networks - based mineral prospectivity mapping in Lhasa area, Tibet," J. Geochem. Explor., vol. 255, p. 107326, Dec. 2023, doi: 10.1016/J.GEXPLO.2023.107326.
32. T. Li, R. Zuo, Y. Xiong, and Y. Peng, "Random-Drop Data Augmentation of Deep Convolutional Neural Network for Mineral Prospectivity Mapping," Natural Resources Research, vol. 30, no. 1, pp. 27–38, Feb. 2021, doi: 10.1007/s11053-020-09742-z.
33. R. Zuo, O. P. Kreuzer, J. Wang, Y. Xiong, Z. Zhang, and Z. Wang, "Uncertainties in GIS-based mineral prospectivity mapping: Key types, potential impacts and possible solutions," Springer, vol. 30, no. 5, pp. 3059–3079, Oct. 2021, doi: 10.1007/S11053-021-09871-Z.
34. F. Yang, R. Zuo, Y. Xiong, Y. Xu, J. Nie, and G. Zhang, "Dual-Branch Convolutional Neural Network and Its Post Hoc Interpretability for Mapping Mineral Prospectivity," Math. Geosci., vol. 56, no. 7, pp. 1487–1515, Oct. 2024, doi: 10.1007/s11004-024-10137-6.
35. X. Lv and G. Wang, "GIS-based mineral prospectivity mapping using machine learning methods: A case study from Duobaoshan ore district, northeastern China," Ore Geol. Rev., vol. 175, p. 106352, Dec. 2024, doi: 10.1016/j.oregeorev.2024.106352.
36. M. Yousefi, E. J. M. Carranza, O. P. Kreuzer, V. Nykänen, J. M. A. Hronsky, and M. J. Mihalasky, "Data analysis methods for prospectivity modelling as applied to mineral exploration targeting: State-of-the-art and outlook," J. Geochem. Explor., vol. 229, p. 106839, Oct. 2021, doi: 10.1016/j.gexplo.2021.106839.
37. Q. Yan *et al.,* "Mineral Prospectivity Mapping Based on Spatial Feature Classification with Geological Map Knowledge Graph Embedding: Case Study of Gold Ore Prediction at Wulonggou, Qinghai Province (Western China)," Natural Resources Research, vol. 33, no. 6, pp. 2385–2406, Dec. 2024, doi: 10.1007/s11053-024-10386-6.
38. W. P.-M. economics and undefined 2022, "Metallogenic models as the key to successful exploration—a review and trends," Springer, vol. 35, no. 3–4, pp. 373–408, Dec. 2022, doi: 10.1007/S13563-022-00325-3.
39. C. L. Nathwani, J. J. Wilkinson, G. Fry, R. N. Armstrong, D. J. Smith, and C. Ihlenfeld, "Machine learning for geochemical exploration: classifying metallogenic fertility in arc magmas and insights into porphyry copper deposit formation," Mineralium Deposita 2022 57:7, vol. 57, no. 7, pp. 1143–1166, Jan. 2022, doi: 10.1007/S00126-021-01086-9.
40. L. Zhang, W. Xiao, K. Qin, and Q. Zhang, "The adakite connection of the Tuwu–Yandong copper porphyry belt, eastern Tianshan, NW China: trace element and Sr-Nd-Pb isotope geochemistry," Miner. Depos., vol. 41, no. 2, pp. 188–200, May 2006, doi: 10.1007/s00126-006-0058-6.