AIMS Geosciences

AIMS Geosciences, 12(2): 335–359.
DOI: 10.3934/geosci.2026013
Received: 29 August 2025
Revised: 17 December 2025
Accepted: 19 January 2026
Published: 20 March 2026

https://www.aimspress.com/journal/geosciences

Research article

A multi-scale factor feature fusion modeling method for landslide susceptibility mapping

Guofang Wang, Jun Cao, Jianwen Sun*, Yifan Wang, Hao Geng and Shouhong Ye

Electric Power Research Institute, Yunnan Power Grid Company ltd., 650217, China

* **Correspondence:** Email: yndwlzl@outlook.com; Tel: +8618578225238.

Abstract: Landslide susceptibility mapping (LSM) is a critical technique for geological hazard risk assessment, aiming to quantitatively evaluate and spatially delineate potentially susceptible areas. Recently, convolutional neural networks (CNNs) and Transformer architectures have been widely applied in remote sensing and geospatial contexts, significantly enhancing automatic feature extraction capabilities. However, single-model approaches often struggle to capture fine-grained local characteristics and long-range spatial dependencies in landslide-prone regions, limiting their representation of multi-scale and irregular spatial structures. To address this challenge, we propose a progressive local-to-global synergy network (PLGS-Net), which integrates the local spatial perception of CNNs with the global dependency modeling of Transformers to achieve hierarchical, collaborative representation of landslide susceptibility features. Empirical experiments conducted in a typical landslide-prone region of Ya'an, Sichuan Province, demonstrated that PLGS-Net outperforms conventional CNNs, pure Transformers, and other mainstream deep learning models in classification accuracy, generalization ability, and regional adaptability, providing an efficient and practical approach for landslide susceptibility assessment in complex terrains.

Keywords: landslide susceptibility mapping; deep learning; PLGS-Net

1. Introduction

Landslides are a common and destructive type of geological hazard that pose serious threats to

transportation infrastructure, human lives and property, and the stability of ecosystems [1,2]. Therefore, conducting high-precision landslide susceptibility mapping (LSM) is of great significance for landslide hazard prevention and spatial planning [3].

With improvements in remote sensing measurement accuracy and the accumulation of multi-source environmental data, research on LSM has progressively evolved from empirical models to data-driven methods [4]. In this context, machine learning (ML) techniques have been widely adopted for landslide susceptibility modeling and have demonstrated excellent performance. Traditional ML algorithms such as logistic regression [5], decision trees [6], support vector machines [7], and random forest [8] have garnered significant attention due to their stability and interpretability [9–12]. For instance, Pham, B.T. et al. employed the rotation forest model for landslide susceptibility assessment in a section of the Indian Himalayas and achieved favorable results [13]. Yesilnacar and Topal, in their LSM study of the Hendek region in Turkey, showed that the predictive accuracy of an artificial neural network model was significantly superior to that of a logistic regression model [14]. Budimir, M.E.A. et al. employed logistic regression for landslide susceptibility evaluation [15]. Abdelaziz Merghadi et al. wrote a review that detailed ML models such as support vector machines, decision trees, random forest, naive Bayes, and k-nearest neighbors and applied them to produce well-performing susceptibility zonation maps [11]. However, these models generally rely on manual feature engineering, limiting their ability to capture the implicit nonlinear and high-dimensional correlations in landslide formation mechanisms, which constrains their accuracy and generalization capability [16].

In recent years, deep learning (DL) methods have achieved remarkable progress in fields such as image recognition and remote sensing analysis, due to their powerful automatic feature extraction capabilities [17,18]. CNNs can effectively extract local structural information, such as edges and textures of landslide areas, through local receptive fields and weight-sharing mechanisms [19,20]. Wang Yi et al. were the first to apply CNNs to landslide susceptibility assessment, achieving excellent results [21]. Maher Ibrahim Sameen et al. employed a 1D-CNN with a Bayesian-based model architecture and hyperparameter optimization to create a landslide susceptibility zonation map for the southern part of Yangyang Province in northeastern South Korea; the results indicated that the CNN model outperformed traditional machine learning models [22]. Huang Faming et al. utilized a fully connected sparse autoencoder, which leveraged the high sparsity and nonlinear correlations of landslide features, to provide a better feature representation than the original data for landslide susceptibility assessment in Sinan County, China [23].

However, the limited receptive field depth of CNNs still poses challenges in characterizing large-scale spatial distributions and global dependencies between landslide areas [24–26]. In contrast, the Transformer architecture, based on a self-attention mechanism, can model long-range dependencies between distant pixels and possesses strong global perception capabilities. It has been preliminarily applied to geological hazard identification tasks in remote sensing scenarios [27,28]. In recent years, some researchers have introduced it into the field of landslide susceptibility assessment to explore the feasibility of Vision Transformer (ViT) models. For instance, Bao Shuai et al. integrated CNN and ViT models for landslide susceptibility mapping in Pingwu County, while Zhao Zeyang et al. used both CNN and ViT models to comprehensively consider local and global features of the data for landslide susceptibility mapping; both studies achieved favorable learning outcomes [29,30].

In landslide susceptibility modeling, landslide-prone conditions often exhibit highly irregular and complex spatial structures [31]. On one hand, susceptible areas tend to be concentrated in specific

geomorphic units, such as steep slopes or regions near faults, displaying pronounced spatial clustering and continuity [32]. On the other hand, local characteristics—such as micro-topography, lithological textures, slope surface variations, and vegetation heterogeneity—are highly diverse and contain rich predisposing information. The interplay between these local and global spatial features renders landslide susceptibility modeling a typical multi-scale modeling problem.

Conventional methods often struggle to simultaneously capture fine-grained local features of conditioning factors and long-range spatial dependencies, thereby limiting the representation of localized predisposing conditions and multi-factor coupling relationships [33]. Therefore, designing a neural network architecture capable of jointly representing local structural features and modeling long-distance spatial dependencies is critical for improving the accuracy and robustness of landslide susceptibility modeling. Specifically, the model needs to possess the ability to encode local features such as slope variations, micro-topographic fluctuations, lithological textures, and vegetation heterogeneity, while also capturing semantic relationships between landslide-prone areas and their surrounding environment at larger spatial scales, enabling a comprehensive characterization of predisposing conditions and multi-factor interactions.

To address the aforementioned challenges, this study proposes a progressive local–global collaborative modeling approach, PLGS-Net. The method employs a multi-stage progressive mechanism (local feature extraction → spatial dependency modeling → feature fusion and discrimination) to achieve hierarchical integration of local predisposition information and regional multi-factor semantic relationships. This enhances the model's comprehensive perception of the spatial distribution characteristics of landslide susceptibility.

We conducted empirical experiments in a typical landslide-prone area and performed extensive comparisons with various mainstream deep learning and traditional machine learning models. The results demonstrate PLGS-Net's advantages in susceptibility classification accuracy, model generalization, and the quality of sensitivity zoning. These findings provide an efficient and generalizable modeling framework for landslide susceptibility analysis and risk prediction in complex regions.

2. Study area and data

2.1. Overview of the study region

This study selected Ya'an City in Sichuan Province as the research area, as shown in Figure 1. The study area is located on the western margin of the Sichuan Basin and experiences a humid subtropical monsoon climate characterized by mild winters, cool summers, four distinct seasons, and abundant rainfall. The average annual temperature is approximately 17.4 ℃, and the average annual precipitation is approximately 1820 mm, which is concentrated during the summer months.

The region boasts a well-developed river network, belonging to the Min River system of the Yangtze River Basin. Major rivers, including the Dadu River and the Qingyi River, traverse the area, featuring high discharge with significant seasonal variability. Groundwater resources are plentiful but unevenly distributed, primarily concentrated along riverbanks and in areas with favorable geological conditions.

In terms of topography and geomorphology, the study area is situated in the transition zone between the Tibetan Plateau and the Sichuan Basin, with terrain sloping from high in the northwest to

low in the southeast. The landscape is predominantly mountainous, featuring high mountains and deep valleys in the northwest, which transition into hilly areas in the southeast.

Geologically, the area is located at the intersection of the Yangtze Block and the Tibetan Plateau collision zone, resulting in complex and diverse geological structures with frequent tectonic activity. Large-scale northwest–southeast trending fault zones, such as the Xianshuihe and Longmenshan fault zones, are developed here. The lithology is diverse: sandstone, shale, and limestone dominate the basin margins and low mountainous hills; metamorphic rocks are common in high mountain areas; and granite is developed around the fault zones. Rock masses are generally fragmented, with significant weathering and erosion observed.

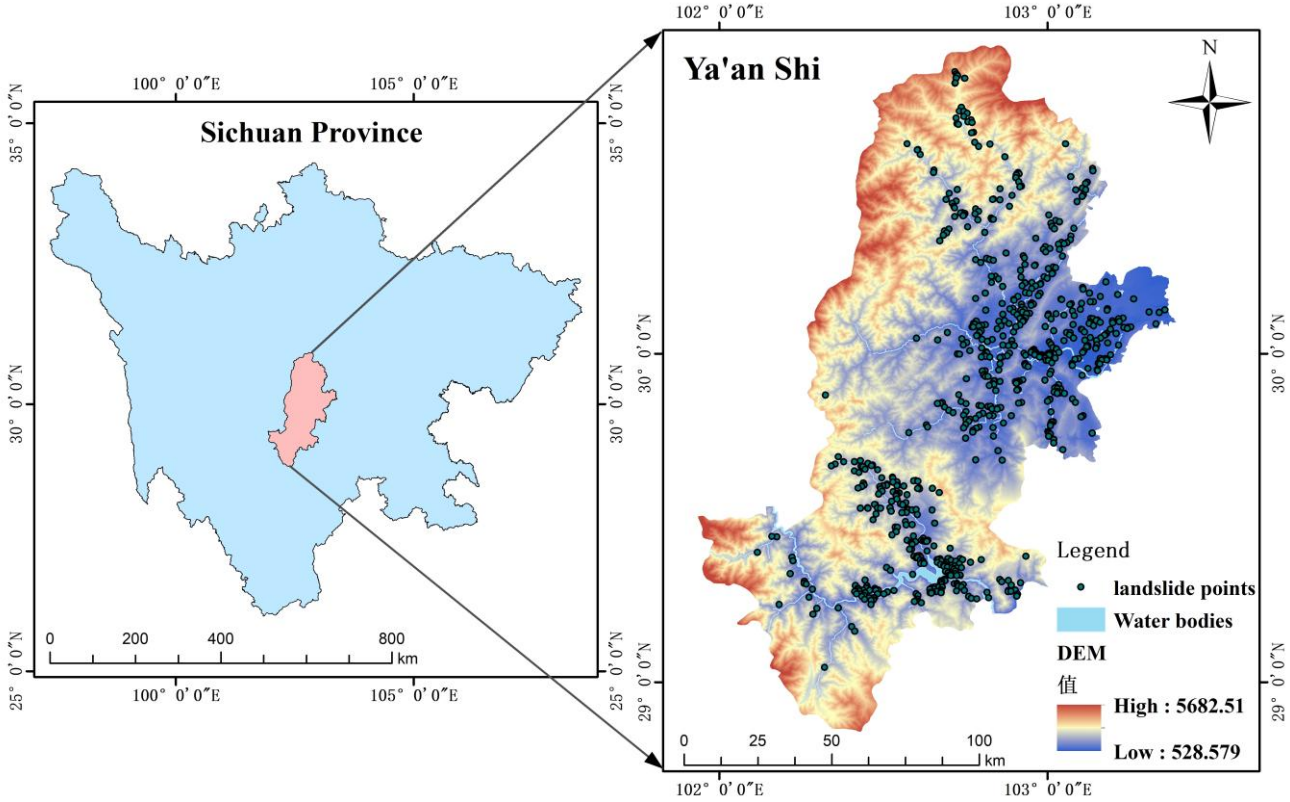


Figure 1. Thematic map of the Ya'an study area.

2.2. Introduction to the data

2.2.1. Landslide inventory

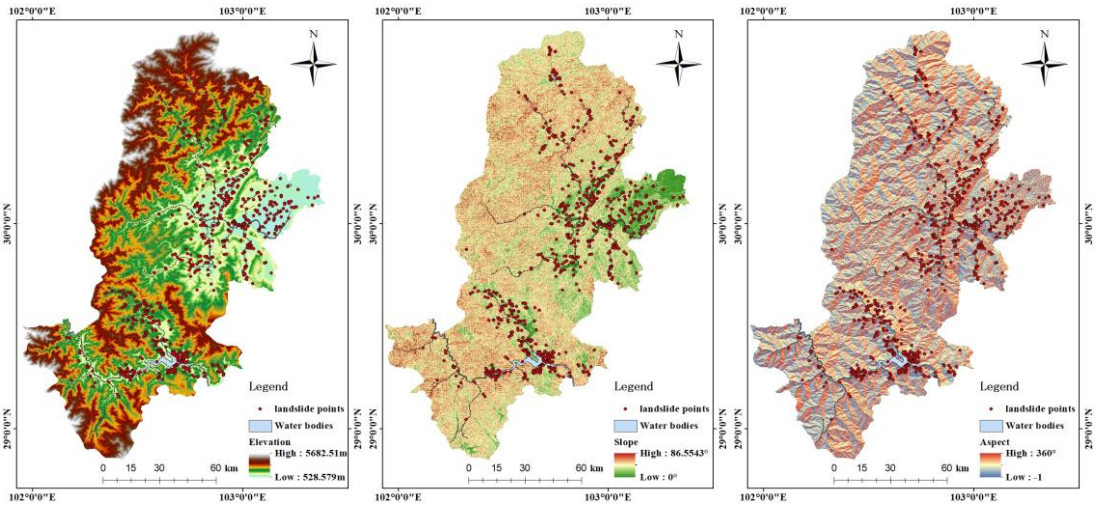
Owing to its unique natural conditions and geological background, Ya'an City frequently experiences rainfall-induced landslide hazards. A total of 1382 documented historical landslides have been documented within the study area, among which 737 were rainfall-induced. The corresponding landslide hazard information was obtained from the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences, for landslide susceptibility assessment research (https://www.resdc.cn.). These landslides are primarily distributed along riverbanks and in the eastern low-elevation hilly areas [34].

2.2.2. Conditioning factors

The selection of conditioning factors is a critical step in landslide susceptibility assessment, requiring a comprehensive consideration of landslide formation mechanisms and the actual conditions of the study area. Based on an in-depth analysis of the geological environment and historical data of the study area, the following 11 indicators were ultimately selected as conditioning factors (as shown in Figure 2): elevation, slope, aspect, plane curvature, profile curvature, lithology, distance to faults, land use, Normalized Difference Vegetation Index (NDVI), mean annual rainfall, and distance to rivers.

Table 1. Data sources for the landslide conditioning factors.

Data type	Evaluation factor	Data sources	Resolution/proportion
Landform	Elevation	ALOS 12.5m DEM	30 m
	Slope	DEM	
	Aspect	DEM	
	Plane curvature	DEM	
	Profile curvature	DEM	
Geological conditions	Fault distance	Geologic map	1:2500000
	Lithology	Geologic map	1:2500000
Hydrometeorology	Rainfall	National Meteorological Data Center	-
	Distance to rivers	2-m Google Image manual interpretation	-
Land cover	Land use	ESA land cover data	30 m
	NDVI	Sentinel-2	



(a) (b) (c)

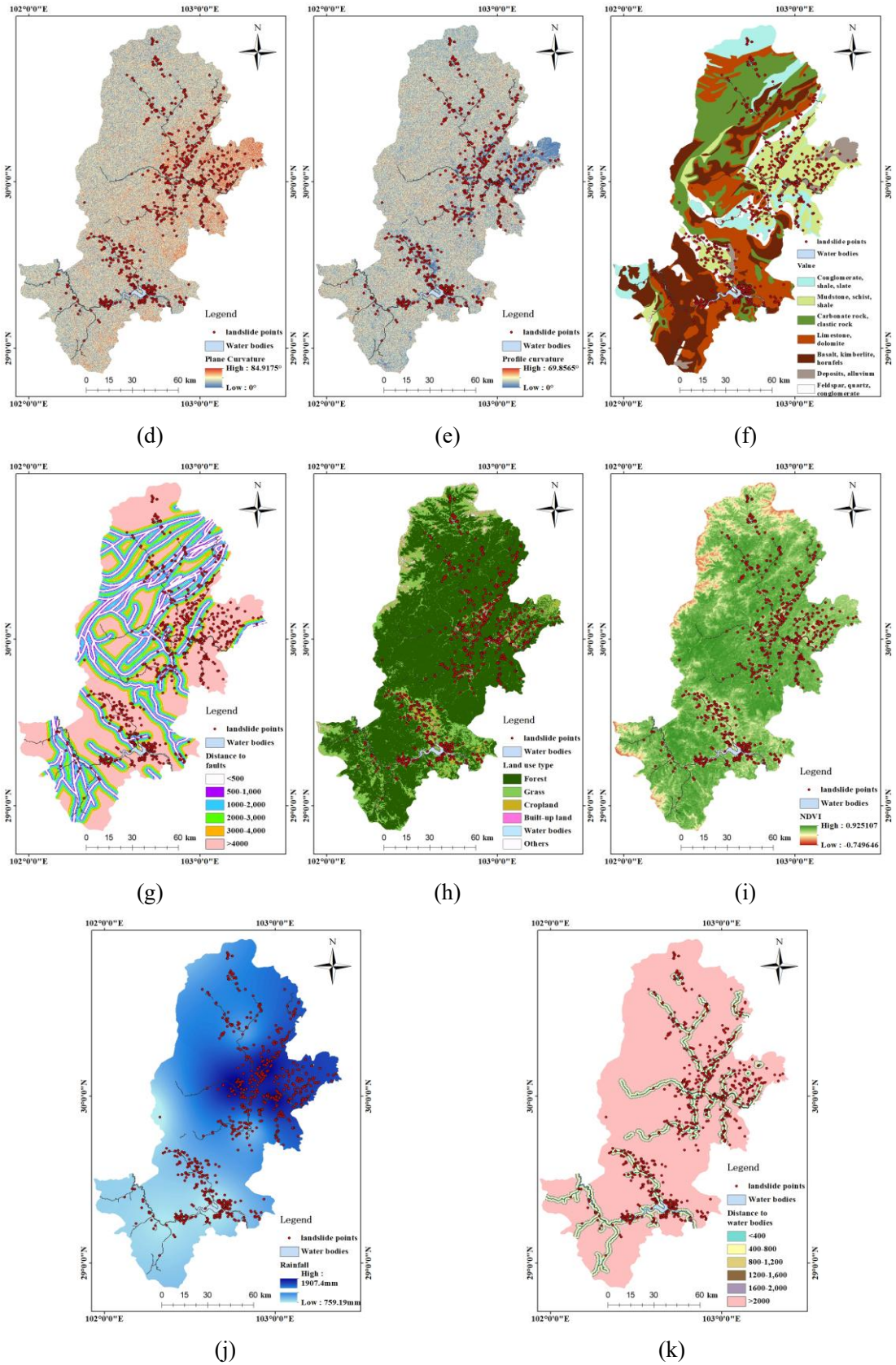


Figure 2. Landslide conditioning factors in the study area: (a) DEM; (b) slope; (c) aspect; (d) plane curvature; (e) profile curvature; (f) lithology; (g) distance to faults; (h) land use; (i) NDVI; (j) mean annual rainfall; (k) distance to rivers.

High-elevation areas exhibit significant gravitational potential energy accumulation, which substantially increases landslide susceptibility. Slope is a key parameter determining the stress state of slopes; steeper slopes are more prone to instability under gravitational forces. Differences in slope aspect lead to variations in solar radiation intensity and weathering degree, thereby regulating moisture evaporation rates and vegetation coverage density, ultimately affecting slope stability. Topographic curvature profoundly influences slope stability by controlling water flow paths, soil saturation zones, and stress concentration points. Weak rock layers (e.g., loosely structured, low-strength lithologies) experience a significant reduction in friction coefficient and shear strength due to water absorption and softening, greatly increasing sliding risk. Areas adjacent to fault zones feature highly fractured rock-soil structures with severely degraded mechanical properties, leading to a sharp increase in landslide hazard. Human activities, such as excessive development through land use practices, can weaken slope stability, while good vegetation coverage helps stabilize soil and retain water, effectively reducing landslide probability. Hydrological factors, including groundwater elevation that increases pore water pressure and slope weight while softening soil, coupled with surface runoff from heavy rainfall that causes significant scouring and erosion, collectively exacerbate landslide risk. The data sources for each landslide conditioning factor are listed in Table 1.

2.2.3. Correlation analysis

Before conducting landslide susceptibility modeling, it is necessary to perform correlation analysis on the conditioning factors. Correlation analysis helps identify relationships among factors, eliminate redundant or irrelevant variables, reduce data noise, and avoid model overfitting. If strong correlations exist among factors, they can affect the model's weight allocation for each feature during the learning process, thereby degrading model performance. Additionally, correlation analysis can reveal associations among different factors to ensure reasonable factor combinations that more accurately reflect the mechanisms of landslide occurrence. This study employs the Spearman correlation coefficient for correlation analysis, with correlation strength categorized as follows: very strong (0.9–1.0), high (0.7–0.9), moderate (0.4–0.7), low (0.2–0.4), and very weak or negligible (0.0–0.2). The correlation indices among the conditioning factors in the study area are presented in Figure 3. Figure 3 shows that the maximum absolute correlation value among the 11 factors is 0.4519, indicating low overall correlation among them. Therefore, we retained all factors for the landslide susceptibility assessment.

3. Method

To systematically enhance the comprehensive characterization of the spatial details and regional correlations of landslide-triggering factors in susceptibility modeling, this study develops a deep learning–based local-global feature collaborative modeling framework, whose entire technical workflow is illustrated in Figure 4. The framework consists of two main components: a data preparation module and a progressive local-global collaborative modeling method, PLGS-Net, as illustrated in Figure 3. This study aims to address two key limitations of traditional modeling approaches: first, the insufficiently detailed representation of local features of landslide-triggering factors (e.g., abrupt slope changes, lithological textures, and vegetation cover variations), which

reduces the accuracy of identifying local landslide-prone conditions; second, the difficulty in capturing latent long-range dependencies and spatial clustering among triggering factors, which limits the systematic modeling of multi-factor coupling mechanisms underlying landslide occurrence.

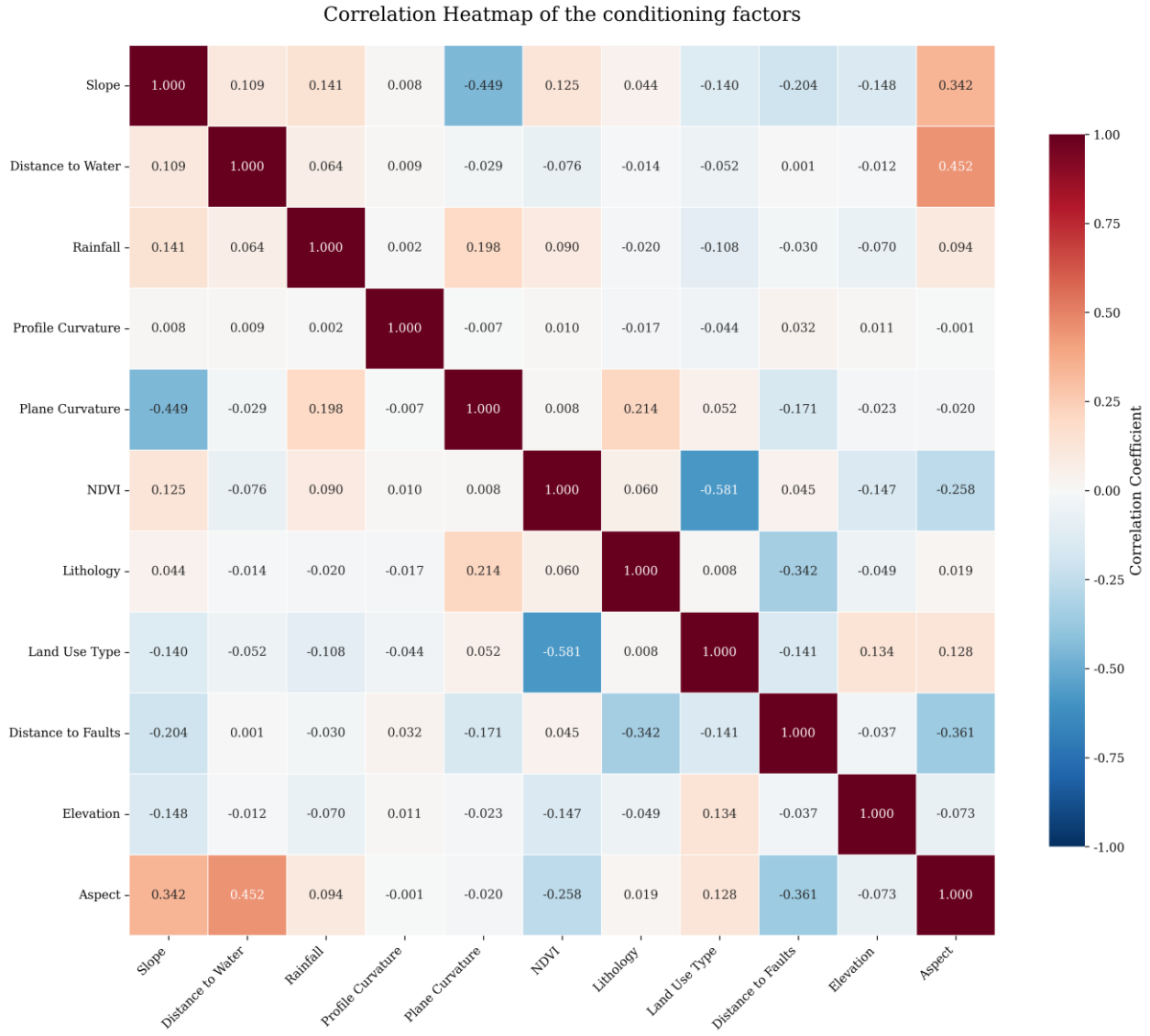


Figure 3. Correlation heatmap of evaluation factors.

PLGS-Net employs a progressive architecture to achieve a multi-level perception of landslide spatial information. From local feature extraction to global correlation modeling and feature fusion for classification, it progressively explores the associations of triggering factors at different scales and outputs a landslide probability distribution map. This approach can more accurately uncover the intrinsic relationships between geological environmental factors and landslide occurrence, significantly improving the capture rate of historical landslide points in high-susceptibility areas and providing a stable and efficient modeling tool for landslide susceptibility zonation and accuracy assessment.

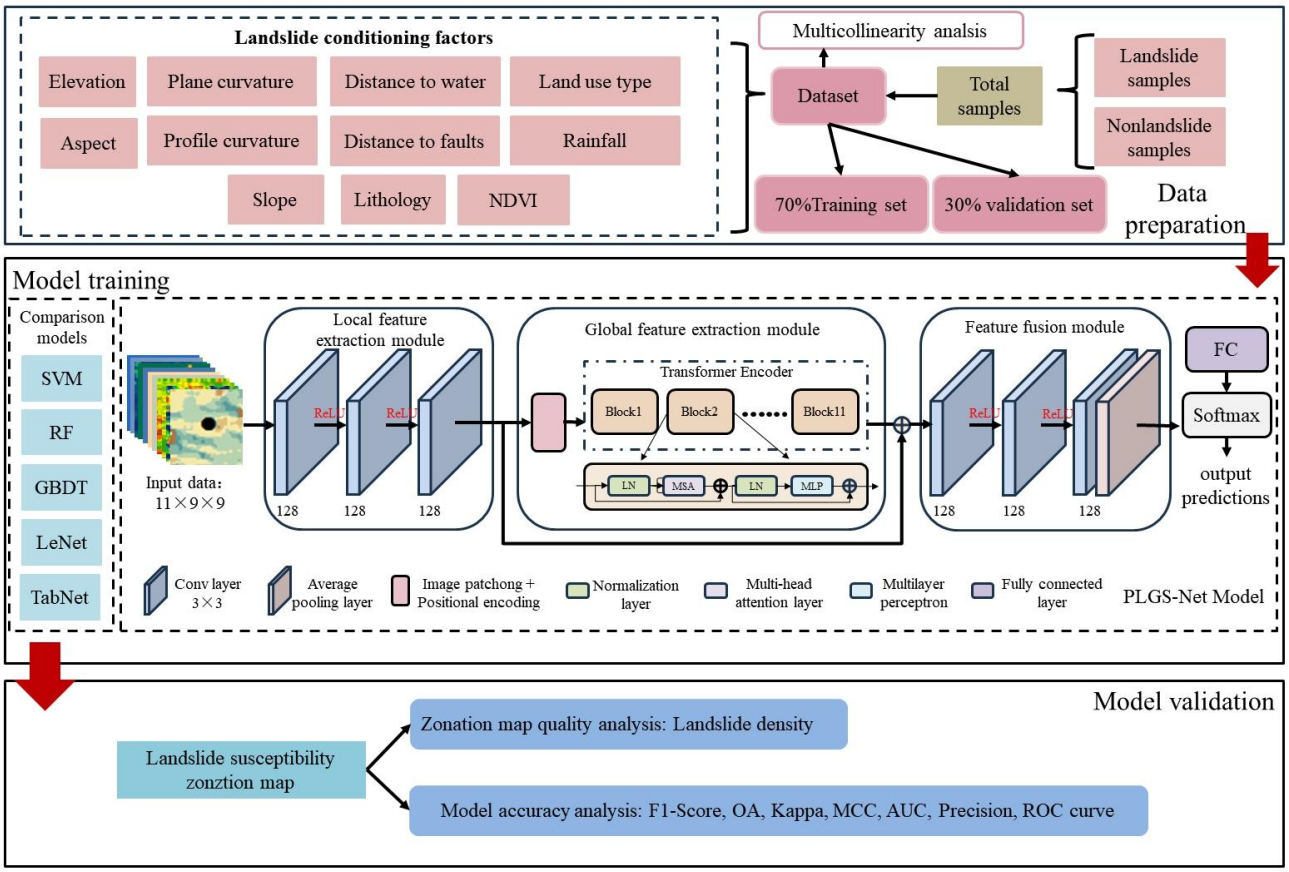


Figure 4. Flowchart of this study.

3.1. Data preparation

In this study, eleven conditioning factors were selected as input features for the model, including elevation, slope, aspect, plane curvature, profile curvature, lithology type, distance to faults, land use type, Normalized Difference Vegetation Index (NDVI), mean annual rainfall, and distance to rivers. All factor rasters were uniformly resampled to a spatial resolution of 30 m and clipped to the extent of the study area.

To enhance the model's ability to perceive local spatial context, a sliding window method was employed to extract 9 × 9 neighborhood samples around both landslide and non-landslide points. Each sample consists of 11 factor layers, forming a three-dimensional tensor with dimensions of 11 × 9 × 9, where 11 represents the number of factor channels and 9 × 9 corresponds to the local spatial window. All numerical factors were normalized using min-max normalization, scaling the values to the range [0, 1] to improve training stability and convergence efficiency. Sample labels were encoded in a binary classification manner, with landslide points labeled as 1 and non-landslide points as 0. Subsequently, the sample set was divided into training and validation sets in a 7:3 ratio, laying the foundation for subsequent model training and performance evaluation.

3.2. PLGS-Net network architecture

PLGS-Net adopts a progressive architecture design. The model input consists of 9 × 9 spatial neighborhood raster data derived from 11 conditioning factors. The architecture begins with stacked convolutional layers to extract low-level spatial features. Then, intermediate feature representations are partitioned into image patches and embedded with positional encoding. These are passed through multiple Transformer encoder layers to capture spatial dependencies and semantic correlations among landslide-prone areas. Finally, the encoded results are fused and compressed, followed by fully connected layers and a Softmax classifier to generate pixel-level landslide probability outputs. This architecture enables hierarchical spatial representation learning and strong reasoning capabilities, offering a robust solution for landslide susceptibility modeling.

The network consists of three primary modules: (1) a local feature extraction module for capturing textural edges and morphological structures; (2) a global feature encoding module for modeling long-range dependencies and spatial context relationships; and (3) a feature fusion and output module for compressing the features and predicting landslide probabilities. The detailed architectural design of the proposed PLGS-Net is visually presented in Figure 5. The synergy of these three modules allows PLGS-Net to achieve a unified representation of both spatial detail and broader regional structure.

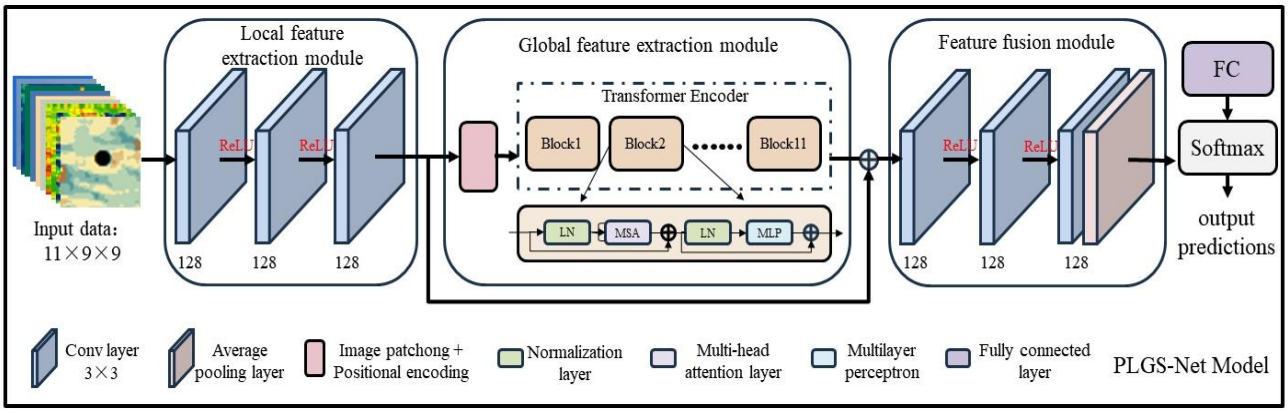


Figure 5. Architecture of the proposed PLGS-Net.

3.2.1. Local feature extraction module

In landslide susceptibility analysis, local environmental features—such as abrupt slope changes, micro-topographic variations, lithological textures, and vegetation heterogeneity—play a critical role in characterizing localized predisposing conditions. To effectively capture these fine-grained features, PLGS-Net incorporates a local feature extraction module at the initial stage, employing a multi-layer convolutional architecture to encode spatial characteristics of input samples.

The input samples are 11 × 9 × 9 three-dimensional tensors centered on landslide or non-landslide points, where 11 represents the number of factor channels, and 9 × 9 denotes the local neighborhood window. This module adopts three stacked 3 × 3 convolution layers, each with 128 channels, activated by ReLU functions. The goal is to extract spatial boundaries, textural details, and subtle variations within the limited receptive field.

Let the input tensor be $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$, where $C = 11$, $H = W = 9$. Each convolution operation is defined as follows:

$$F^{(l)} = \sigma\left(W^{(l)} * F^{(l-1)} + b^{(l)}\right) \tag{1}$$

Here, $F^{(0)} = \mathbf{X}$, * denotes the convolution operation, $\sigma(\cdot)$ is the ReLU function, and l is the current layer index.

3.2.2. Global feature encoding module

After convolutional feature extraction, the local features will be divided into multiple image patches (Patch). Each image patch is flattened into a one-dimensional vector and then transformed into a unified high-dimensional feature embedding through linear mapping. Additionally, to preserve the relative positional information between image patches, learnable position embeddings (position embedding) are added to each patch, constructing a sequence input that includes spatial order:

$$z_0 = [x_p^1 E; x_p^2 E; \ldots; x_p^n E] + E_{pos} \tag{2}$$

Here, x_p^i represents the i-th image patch, E is the embedding matrix, and E_{pos} is the position embedding. The constructed sequence input is fed into a stacked structure composed of 11 layers of Transformer encoders.

Each layer of the Transformer consists of a multi-head self-attention mechanism (MSA) and a feedforward network (MLP), establishing dynamic relationships between image patches using the following computational approach:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

$$\text{MSA}(X) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)W^O \tag{4}$$

$$\text{MLP}(X) = \text{ReLU}(XW_1 + b_1)W_2 + b_2 \tag{5}$$

Where Q, K, V represent the query, key, and value, respectively, and d_k is the dimension of the key vector.

This process is equivalent to dynamically allocating attention weights for each image patch relative to other patches, thereby calculating a weighted sum over all patches based on their semantic relevance. In landslide susceptibility modeling tasks, this enables the model to automatically perceive potential spatial causality links between adjacent or distant regions, such as whether a high-slope area has a higher risk of landslides due to its relationship with nearby fault zones or whether there is a potential coupling relationship between heavy rainfall accumulation zones and distant exposed rock layers.

Through the repeated modeling of non-local dependencies between these regions using 11 stacked Transformer encoders, the model can integrate spatial causal features at different scales, constructing a structurally hierarchical global semantic representation. The final output will be used as input for the fusion module, further optimizing the discrimination with local features.

3.2.3. Feature fusion and output module

The sequence of global features output by the Transformer encoder, after completing the long-range spatial relationship modeling, needs to be further reconstructed into a two-dimensional structure suitable for spatial discrimination tasks. For this purpose, first reconstruct the encoder output as a feature map (feature map) arranged in space, with dimensions consistent with the initial convolutional module output, ensuring spatial alignment for subsequent feature integration.

This feature map is then fed into three consecutive 3 × 3 convolutional layers for fusion and compression. Each convolutional layer has 128 channels and uses ReLU activation functions to enhance nonlinear modeling capabilities. This process aims to further integrate the spatial dependency information obtained from the previous stage's modeling and filter out redundant or irrelevant features, thereby strengthening the spatial response related to landslide discrimination.

The fusion module not only serves as a semantic compression function but also acts as an aggregation stage for local structural information and globally dependent features. Through stacked convolutional operations, the model compresses the long-range semantic dependencies captured by the Transformer and the edge-texture features retained by the local convolutional module into a unified representation vector. This representation is then flattened and fed into a fully connected layer for classification judgment:

$$\hat{y} = \text{Softmax}(W_f \cdot \mathbf{z} + b_f) \tag{6}$$

Here, $\mathbf{z} \in \mathbb{R}^d$ is the flattened one-dimensional feature vector after fusion, and $W_f \in \mathbb{R}^{2\times d}$ and $b_f \in \mathbb{R}^2$ are the weight matrix and bias term of the fully connected layer, respectively, with an output dimension of 2, representing the probability values for the categories "landslide" and "non-landslide".

This module effectively integrates local and global information while maintaining spatial precision, thereby enhancing the model's response capability to small-scale landslides and ambiguous boundary areas. The final output is a pixel-level landslide susceptibility prediction map.

4. Experiments and analysis

4.1. Experimental design

To comprehensively evaluate and compare the performance of different machine learning models in landslide susceptibility assessment, as well as to verify the effectiveness of the proposed PLGS-Net model, a systematic experimental framework was designed in this study. Under the condition of using the same training and validation datasets (training:validation = 7:3), six models—support vector machine (SVM), random forest (RF), gradient boosting decision tree (GBDT), TabNet, LeNet, and PLGS-Net—were trained separately. A unified set of evaluation metrics was employed to assess the performance of all models, including generalization ability (AUC), balance (F1), robustness (Kappa and MCC), and overall accuracy (OA), thereby providing a comprehensive assessment of each model's predictive capacity, balance, robustness, and generalization performance. Subsequently, the landslide susceptibility probability maps predicted by each model were visualized, and the probability values

were classified into five susceptibility levels—very low, low, moderate, high, and very high—using the natural breaks method, resulting in corresponding susceptibility zonation maps. Finally, to quantify the likelihood of landslide occurrence within different susceptibility levels and evaluate the reliability of each model's zonation results, three indicators were calculated: (1) the proportion of the area of each susceptibility level relative to the total study area, (2) the proportion of historical landslide counts within each level relative to the total number of landslides, and (3) the landslide point density within each level (i.e., the ratio of the number of landslide pixels to the total number of pixels in the corresponding susceptibility zone).

4.2. Parameter settings for comparative methods

All experiments in this study were implemented using Python 3.6 and run on a Windows operating system. The experimental hardware platform was equipped with an Intel(R) Core(TM) i7-8750H processor and an NVIDIA GeForce GTX 1060 GPU.

For model construction, the SVM, RF, and GBDT models were implemented using the SVC, RandomForestClassifier, and GradientBoostingClassifier classes from the scikit-learn machine learning library, respectively. TabNet was constructed using the TabNetClassifier from the pytorch_tabnet library. LeNet was adapted from the classical LeNet-5 convolutional neural network framework, comprising two convolutional layers, two pooling layers, and three fully connected layers.

During model training, this study further standardized the basic training settings across all models to ensure reproducibility and fairness in the comparative experiments:

1. Deep learning models (PLGS-Net, TabNet, LeNet) training settings:

Optimizer: Adam optimizer was used for all models. The initial learning rates were set to 0.001 for LeNet and TabNet, and 0.0001 for PLGS-Net. Adam's default β parameters were used ($\beta_1 = 0.9$, $\beta_2 = 0.999$), and a weight decay of 1×10^{-5} was added for PLGS-Net and TabNet.

- Loss function: Binary cross-entropy (BCE) was used to suit the binary classification task of landslide vs. non-landslide.
- Epochs: PLGS-Net was trained for 100 epochs; TabNet for 300 epochs; LeNet for 400 epochs.
- Early stopping: PLGS-Net and LeNet used early stopping based on validation loss. For LeNet, patience was set to 50 with a delta of 0.001; training stopped if validation loss did not improve by more than 0.001 over 50 consecutive epochs, retaining the best weights. PLGS-Net used a patience of 10 with a delta of 0.001.

2. Traditional machine learning models (SVM, RF, GBDT) training settings: SVM, RF, and GBDT were implemented using scikit-learn, so gradient-based optimizers like SGD or Adam were not used. Specifically:

- SVM: Hyperparameters (penalty factor C and kernel width γ) were determined via 5-fold cross-validation and grid search.
- RF: Internal optimization was achieved through random feature selection and Bagging. The number of trees was set to 100, with maximum features using the sqrt strategy.
- GBDT: Parameters such as learning rate, number of weak learners (n_estimators), and tree depth (max_depth) directly controlled the model iteration process, functioning as its internal optimization mechanism.

All three models used scikit-learn's default binary classification objectives: SVM used the maximum margin loss, while RF and GBDT used tree-based objectives based on information gain or residual fitting. The optimal hyperparameter configurations for all models are listed in Table 2. Below is a detailed description of the basic principles, implementation approaches, and key parameter configurations for the five comparative models.

4.2.1. SVM

SVM is a supervised classification model based on the maximum margin principle, which constructs an optimal hyperplane in the feature space to effectively separate two classes of samples [35]. For linearly inseparable data, SVM maps the samples into a higher-dimensional space using a kernel function, allowing them to become linearly separable in the new space. In this study, the radial basis function (RBF) kernel was employed, with its form given by:

$$K(x_i, x_j) = exp(-\gamma \| x_i - x_j \|^2) \tag{7}$$

where the penalty parameter C and the kernel width γ were selected via grid search to obtain the optimal combination. SVM was implemented using sklearn.svm.SVC, demonstrating good stability for nonlinear classification tasks on small- to medium-scale datasets.

4.2.2. RF

RF is an ensemble learning method that constructs multiple randomized decision trees and aggregates their predictions by voting to obtain the final classification result. Its ensemble prediction function can be expressed as follows:

$$\hat{y} = majority\_vote(h_1(x), h_2(x), \ldots, h_T(x)) \tag{8}$$

where $h_t(x)$ represents the prediction of the t-th tree, and T is the total number of trees.

RF introduces randomness both in sample selection and feature selection, which effectively reduces overfitting and improves generalization [36]. In this study, RF was implemented using RandomForestClassifier, with the number of trees set to 100 and the maximum number of features for each split set using the square root (sqrt) strategy.

4.2.3. GBDT

GBDT builds an additive model within the boosting framework by sequentially fitting residuals. In each iteration, a weak learner (CART tree) is trained to correct the errors of the previous iteration. In each round, GBDT fits the negative gradient of the loss from the previous round, thereby continuously refining the model error.

If the objective function is $L\left(y, f(x)\right)$, the m-th weak learner in GBDT can be expressed as:

$$h_m(x) = \arg\min_h \sum_{i=1}^{n} \left[-\frac{\partial L(y_i, f_{m-1}(x_i))}{\partial f_{m-1}(x_i)} - h(x_i) \right]^2 \tag{9}$$

The model is then updated as:

$$f_m(x) = f_{m-1}(x) + \nu h_m(x) \tag{10}$$

where ν is the learning rate. In this study, GBDT was implemented using GradientBoostingClassifier with 100 weak learners, tree depth of 4, and learning rate 0.05. GBDT can effectively capture nonlinear relationships between features and has wide applications in disaster prediction tasks.

4.2.4. TabNet

TabNet is a deep learning architecture designed for tabular data, which achieves dynamic feature selection and transformation through a sequential attention mechanism [37]. Its feature selection mask can be expressed as:

$$M^{(i)} = \text{Sparsemax}\left(P^{(i-1)} \cdot h_a^{(i)}(X)\right) \tag{11}$$

The model uses a multi-step decision structure to progressively extract discriminative features. In this study, the number of decision steps was set to 5, the feature and attention dimensions were set to 64, and the model was trained for 150 epochs using the Adam optimizer. TabNet can adapt to complex feature structures and performs well in disaster prediction tasks.

4.2.5. LeNet

Table 2. Parameter configurations for each model.

Model	Parameters
SVM	C=1.0; gamma = "scale"; kernel= "rbf"; class_weight = "balanced"
RF	n_estimators = 100; max_depth = 6; max_features = "sqrt"
GBDT	Lr = 0.06; n_estimators = 100; max_depth = 4
TabNet	n_d = 16; n_a = 16; n_steps = 3; gamma = 1.1; lr = 0.001; epochs = 300
LeNet	batch_size = 64; epoches = 400; patience = 50; lr = 0.001; (patience = 50, delta = 0.001)
PLGS-Net	patch_size = 3; embedding dimension = 128; mlp_ratio = 4; lr = 0.0001; epochs = 100; (patience = 10, delta = 0.001)

LeNet is a CNN model, originally applied to the handwritten digit recognition task [38]. Its structure is composed of convolutional layers, pooling layers, and fully connected layers, and it can effectively capture local spatial patterns through mechanisms like local receptive fields, weight sharing, and hierarchical feature extraction. For the grid window-based sample input in landslide susceptibility prediction, LeNet is capable of extracting valid texture and topographic information from local spatial neighborhoods.

Let the input window be $X \in \mathbb{R}^{H \times W}$; the convolution operation with convolution kernel K can be expressed as:

$$Y_{i,j} = \sum_m \sum_n X_{i+m, j+n} \cdot K_{m,n} \tag{12}$$

Subsequently, the pooling layer reduces the resolution of the feature map and enhances translation invariance. Finally, the network outputs class scores via the fully connected layer:

$$\hat{y} = \text{Softmax}(W \cdot z + b) \tag{13}$$

Here, z is the feature vector obtained after stacking convolution and pooling layers, while W and b are learnable parameters.

In this study, a LeNet structure suitable for 9 × 9 grid segments is reconstructed based on input sample features, and it is implemented on the PyTorch platform. During training, the Adam optimizer is used, with a learning rate of 0.001 and a total of 400 training epochs.

4.3. Result analysis

4.3.1. Landslide susceptibility zonation maps

The landslide susceptibility zonation maps generated by each model are shown in Figure 6. It can be observed that the predicted high-to-very-high susceptibility zones from all models exhibit a high degree of spatial consistency with historical landslide events, showing strong spatial clustering. These high-risk areas are primarily concentrated in the Qingyi River watershed, lake-adjacent radiation zones, and the low-elevation hilly regions in eastern Ya'an.

Based on the evaluation statistics of each model (as shown in Table 3), the quality of the zonation maps produced by different models is analyzed. In terms of landslide density trends, the landslide density in each model increases significantly with rising susceptibility levels across all zones. This indicates that all models effectively capture the spatial distribution patterns of landslides, and the resulting susceptibility maps are highly consistent with the defined classification criteria.

Further detailed comparative analysis reveals the following performance characteristics:

The RF model achieves the best performance in high-to-very-high susceptibility zones (occupying only 27% of the area but achieving a landslide density as high as 5.8). However, it performs poorly in very-low-to-low susceptibility zones (covering less than 60% of the area while still containing over 4% of landslides).

The LeNet model achieves the lowest landslide density in high-to-very-high zones (0.09), but these zones occupy a relatively large area (over 60%), and the model's performance in high-risk zones is mediocre.

The SVM, GBDT, and TabNet models do not show significant advantages across susceptibility levels.

In contrast, the PLGS-Net model demonstrates superior overall performance: In very-low-to-low susceptibility zones, it covers 63% of the study area, with landslides accounting for only 3.7% of the total and a low landslide density of 0.21. In high-to-very-high susceptibility zones, it captures 92% of all landslides within just 28% of the area, maintaining a high landslide density of 5.0.

Compared to other models, PLGS-Net significantly improves the reliability and predictive stability of the zonation map by optimizing the balance between low- and high-risk zones—maintaining a broad coverage with low hazard probability in low-risk areas and achieving precise localization with high hazard concentration in high-risk areas. This demonstrates its superior comprehensive performance.

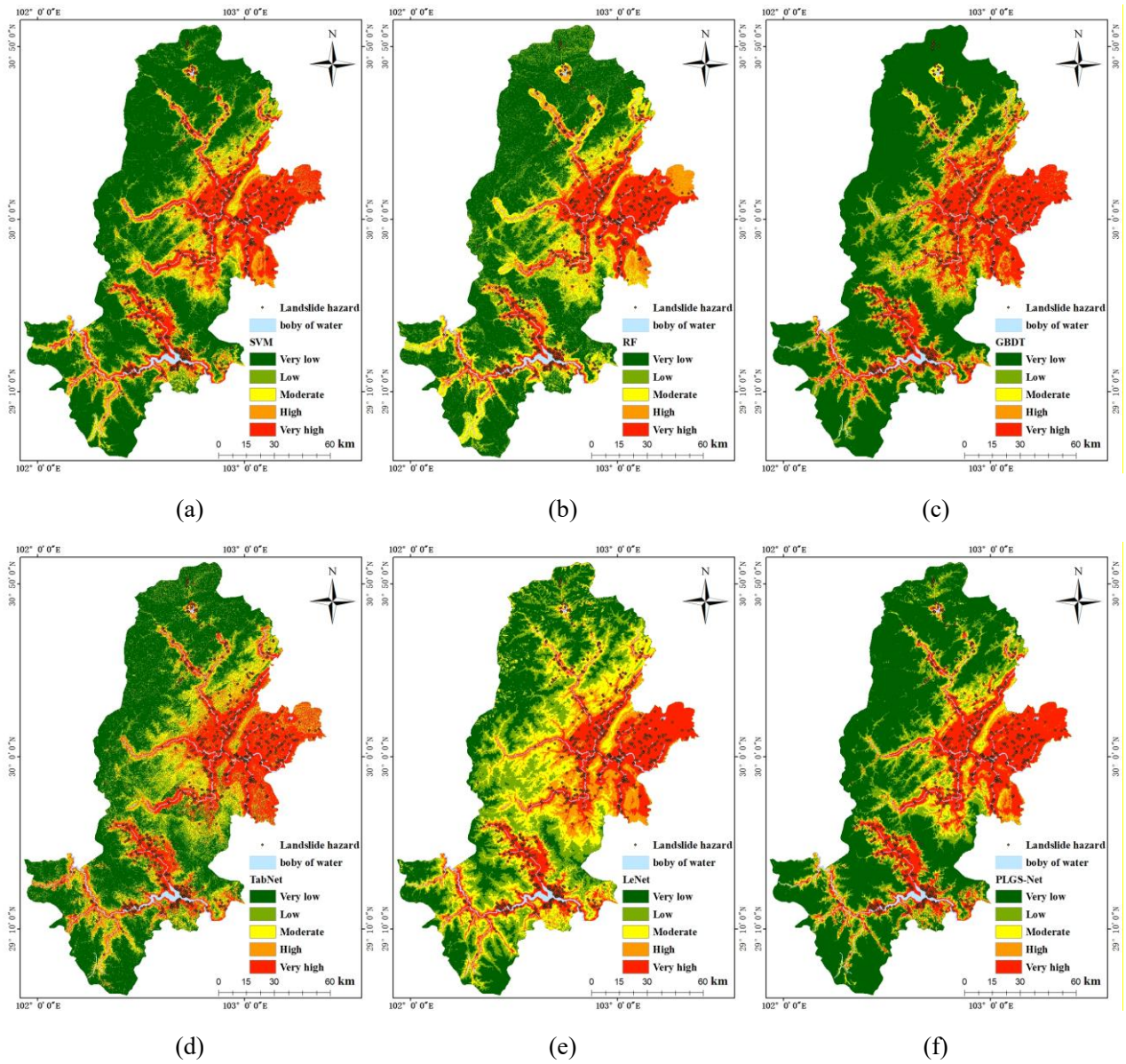


Figure 6. (a)–(f): Landslide susceptibility zonation maps for the study area based on the SVM, RF, GBDT, TabNet, LeNet, and PLGS-Net models, respectively.

4.3.2. Model performance evaluation

The evaluation metrics for each model are presented in Table 4. Comparative analysis shows that PLGS-Net, with its local–global feature fusion architecture, demonstrates comprehensive performance advantages. Among traditional models, RF achieves high precision (precision = 0.9189) and good generalization (AUC = 0.9699) but suffers from lower recall (recall = 0.9127).

SVM and GBDT exhibit an imbalance between recall and precision (recall > 0.93, precision < 0.89).

LeNet, limited by the local receptive field of CNNs, has the lowest precision (0.8448) and weaker generalization ability (AUC = 0.9542).

Table 3. Statistical evaluation of landslide susceptibility for each model.

Model	Susceptibility level	Zone area	Percentage of total area	Landslide count	Percentage of total landslides	Landslide density
SVM	Very low	7031721	0.4762	19	0.0251	0.0528
	Low	1899755	0.1286	21	0.0278	0.2159
	Moderate	1235018	0.0836	30	0.0397	0.4745
	High	1613852	0.1093	94	0.1243	1.1377
	Very high	2986798	0.2023	592	0.7831	3.8716
RF	Very low	6515842	0.4412	13	0.0172	0.0390
	Low	2248705	0.1523	22	0.0291	0.1911
	Moderate	1968440	0.1333	65	0.0860	0.6450
	High	1629530	0.1103	137	0.1812	1.6422
	Very high	2404627	0.1628	519	0.6865	4.2159
GBDT	Very low	7443598	0.5041	17	0.0225	0.0446
	Low	1638275	0.1109	17	0.0225	0.2027
	Moderate	1048065	0.0710	27	0.0357	0.5032
	High	1255047	0.0850	66	0.0873	1.0272
	Very high	3382159	0.2290	629	0.8320	3.6327
TabNet	Very low	6168704	0.4177	12	0.0159	0.0380
	Low	2302457	0.1559	18	0.0238	0.1527
	Moderate	1609980	0.1090	40	0.0529	0.4853
	High	1582071	0.1071	110	0.1455	1.3581
	Very high	3103932	0.2102	576	0.7619	3.6248
LeNet	Very low	3604147	0.2441	1	0.0013	0.005
	Low	3444214	0.2332	16	0.0212	0.0907
	Moderate	2719610	0.1842	35	0.0463	0.2514
	High	2052255	0.1390	114	0.1508	1.0850
	Very high	2946918	0.1996	590	0.7804	3.9107
PLGS-Net	Very low	7345736	0.4974	7	0.0093	0.0186
	Low	2041839	0.1383	21	0.0278	0.2009
	Moderate	1110521	0.0752	31	0.0410	0.5453
	High	1104103	0.0748	63	0.0833	1.1146
	Very high	3164945	0.2143	634	0.8386	3.9129

In contrast, PLGS-Net leads in all six evaluation metrics. It achieves a high recall of 0.9310 (second only to LeNet) while significantly improving precision to 0.9147 (an 8.3% improvement over LeNet). Its innovative feature fusion paradigm effectively resolves the core contradiction between local features (which may overlook macroscopic patterns) and global features (which may weaken fine-grained responses). PLGS-Net is the first model to achieve simultaneous improvements in AUC, F1, MCC, and OA, offering a new multi-scale collaborative modeling paradigm for geological hazard prediction.

Table 4. Evaluation metrics for each model.

Model	Evaluation metrics						
	OA	Kappa	Precision	Recall	F1	MCC	AUC
SVM	0.9038	0.8076	0.8824	0.9319	0.9064	0.8089	0.9629
RF	0.9161	0.8322	0.9189	0.9127	0.9158	0.8322	0.9699
GBDT	0.9079	0.8158	0.8913	0.9290	0.9081	0.8166	0.9683
TabNet	0.9127	0.8254	0.9127	0.9126	0.9128	0.8254	0.9692
LeNet	0.8876	0.7752	0.8448	0.9496	0.8941	0.7812	0.9542
PLGS-Net	0.9216	0.8425	0.9147	0.9310	0.9227	0.8427	0.9742

5. Discussion

5.1. Ablation study

To verify the necessity of the synergy between the local feature extraction and global dependency modeling modules in PLGS-Net, this study conducts three sets of controlled experiments: (1) retaining only the CNN branch, (2) retaining only the ViT branch, and (3) the complete model. Under strictly controlled conditions—including identical training and test sets, hyperparameters, and input features—the evaluation metrics shown in Table 5 are obtained.

The experimental results show that:

The CNN-only model predicts nearly all samples as landslides, leading to severely compromised metrics such as overall accuracy (OA = 0.5007) and indicating extreme overfitting. This demonstrates that the CNN framework, when used as a standalone module, loses fundamental generalization capability.

The ViT-only model exhibits basic predictive ability, but its overall performance is significantly inferior to that of the complete model.

Therefore, the performance improvement achieved by the full PLGS-Net through feature fusion confirms the essential and complementary roles of CNN and ViT in landslide susceptibility prediction. The integration of local spatial details (captured by CNN) and long-range contextual dependencies (modeled by ViT) is critical for robust and accurate modeling.

Table 5. Evaluation metrics from ablation experiments.

Model	Evaluation metrics						
	OA	Kappa	Precision	Recall	F1	MCC	AUC
CNN	0.5007	0.0013	0.5003	1.0000	0.6670	0.0261	0.9494
ViT	0.9053	0.8106	0.8961	0.9168	0.9064	0.8108	0.9639
PLGS-Net	0.9216	0.8425	0.9147	0.9310	0.9227	0.8427	0.9742

5.2. Extended comparison with recent state-of-the-art studies

Drawing on representative studies published between 2022 and 2024—covering ensemble learning, attention mechanisms, Transformer variants, and multiple CNN-based models—we further

compare PLGS-Net with cutting-edge approaches from five perspectives: model architecture innovation, feature-processing capability, performance, interpretability, and computational efficiency, thereby highlighting its core advantages and application value. Among the latest studies, hybrid ensemble learning and Transformer variants are the two mainstream innovation trends; however, the progressive local–global fusion mechanism of PLGS-Net remains distinct:

1. Ensemble learning models: Lv et al. proposed a heterogeneous ensemble framework based on DBN, CNN, and ResNet (Stacking/Boosting), achieving AUC = 0.984 and OA = 94.17% in the Three Gorges Reservoir Area. Its strength lies in leveraging complementary base learners to reduce variance [39]. However, the model requires parallel training of multiple models and output fusion, which leads to exponentially increased computational cost. Moreover, it does not incorporate a dedicated multi-scale spatial feature fusion mechanism tailored for landslides, relying instead on ensemble strategies to compensate for single-model limitations. In contrast, PLGS-Net employs an integrated "local extraction—global modeling—progressive fusion" architecture within a single network, achieving AUC = 0.9742 and OA = 92.16% while avoiding the redundant computation inherent in multi-model ensembles, making it more suitable for large-scale, efficient regional modeling.

2. Pure CNN comparison models: Ge et al. compared five CNN architectures (AlexNet, Xception, ResNet-101, among others) and found that Xception (AUC = 0.994) and AlexNet (AUC = 0.992) performed best for landslide assessment along transmission lines [40]. However, these models inherently rely on convolutions for local feature extraction and cannot effectively model long-distance spatial dependencies. PLGS-Net, through the synergy of CNN and Transformer modules, retains CNN's sensitivity to local structural information while overcoming the limitations of pure CNNs in global-dependency modeling, resulting in more stable performance in the complex geological environment of Ya'an.

3. Transformer-variant models: Chen et al. proposed the LCFSTE framework, which integrates Swin Transformer's local window attention with landslide-conditioning factors, achieving AUC = 0.9444 in Jiuzhaigou County [41]. The local-window mechanism reduces computational cost, but the model still adopts static factor inputs and lacks hierarchical optimization of local–global feature interactions. In contrast, PLGS-Net's progressive fusion mechanism is more consistent with the physical formation logic of landslides: local details (e.g., boundary morphology) and global dependencies (e.g., fault-zone clustering). This design allows PLGS-Net to more effectively capture complex terrain characteristics.

5.3. Limitations and future work

This study has certain limitations in terms of data and model interpretability. At the data level, modeling relies on static geological environmental factors and does not incorporate key dynamic triggering factors such as temporal variations in rainfall intensity or the evolution of human engineering activities. Recent studies highlight two promising directions for incorporating temporal dimensions: quantifying slope rainfall sensitivity to refine thresholds [42] and optimizing landslide inventory timeframes [43]. Future work could integrate such dynamic rainfall response characteristics and optimal temporal sampling strategies with static susceptibility models. Moreover, samples are constructed solely based on the historical static distribution of landslides, lacking expansion in the temporal dimension and optimization of class balance. At the model level, the multi-scale feature

fusion process of PLGS-Net still exhibits a "black-box" nature. It does not quantify the contribution of features at different scales or specific triggering factors to the susceptibility prediction results, nor does it thoroughly characterize the intrinsic relationships between factor threshold effects and disaster occurrence mechanisms.

Future improvements will focus, on one hand, on constructing a multi-source data system combining static geological environmental factors and dynamic triggering factors, adding NDVI time series, rainfall temporal data, etc., and optimizing sample balance by combining SMOTE with stratified sampling, thereby supplementing temporal samples to enhance the model's adaptability to dynamic changes; on the other hand, on introducing interpretability methods such as SHAP values to quantify the contribution weights of each factor and features at different scales and using visualization to reveal the internal logic of feature fusion. Additionally, embedding a factor threshold identification module to explore critical conditions of key factors, such as slope and distance to water systems, will strengthen the model's explanatory capability regarding landslide occurrence mechanisms.

6. Conclusions

This study proposes PLGS-Net, a progressive local–global collaborative modeling framework designed to address two long-standing challenges in landslide susceptibility analysis. First, traditional methods exhibit limited capability in capturing fine-scale local characteristics of landslide conditioning factors—such as abrupt slope changes, micro-topographic variations, lithological textures, and vegetation heterogeneity—resulting in insufficient representation of localized predisposing conditions. Second, they struggle to model the latent long-range spatial dependencies and multi-factor coupling effects among conditioning variables, thereby limiting the capacity to characterize spatial clustering patterns and structural drivers inherent in landslide formation processes.

PLGS-Net integrates the complementary strengths of convolutional neural networks and Transformer architectures. The convolutional component enhances the extraction of local spatial structures, topographic gradients, and environmental textures, while the Transformer module enables effective modeling of multi-scale spatial dependencies and semantic relationships. Through a unified feature fusion mechanism, PLGS-Net generates landslide susceptibility probability maps with improved spatial continuity and discriminative capability. Empirical experiments conducted in a typical landslide-prone region of Ya'an, Sichuan Province, demonstrate that PLGS-Net outperforms traditional machine learning models (SVM, RF, GBDT), classical convolutional networks (LeNet), and hybrid structured models (TabNet) in terms of classification accuracy, spatial consistency of susceptibility zoning, and generalization performance. These findings highlight the importance of jointly modeling local structural features and global semantic dependencies to improve landslide susceptibility assessment in complex terrain environments.

Furthermore, the ablation experiments confirm the complementary and indispensable roles of the convolutional and Transformer modules within PLGS-Net. The convolutional module enhances sensitivity to localized predisposing environments, whereas the Transformer module strengthens the representation of cross-scale interactions among topographic, geological, and environmental factors. Overall, PLGS-Net demonstrates strong feature-learning capability, spatial adaptability, and transferability, providing a promising modeling paradigm and technical pathway for landslide susceptibility mapping in heterogeneous landscapes.

Author contributions

Guofang Wang: Investigation, Formal analysis, Writing–original draft. Jun Cao and Yifan Wang: Data curation, Software, Validation. Jianwen Sun: Conceptualization, Methodology, Supervision, Writing–review & editing. Hao Geng and Shouhong Ye: Visualization, Validation. All authors have read and approved the final manuscript.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by Yunnan Power Grid Technology Project under Grant 056200KK52222193.

Conflict of interest

The authors declare no conflict of interest.

References

1. Gariano SL, Guzzetti F (2016) Landslides in a changing climate. *Earth-Sci Rev* 162: 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011
2. Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. *Nat Hazards Earth Syst Sci* 18: 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018
3. Ngo PTT, Panahi M, Khosravi K, et al. (2021) Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geosci Front* 12: 505–519. https://doi.org/10.1016/j.gsf.2020.06.013
4. Fang Y, Gan C, Cao W (2023) Undersampling Ensemble and Deep Learning-Based Landslide Susceptibility Mapping Method for Geological Hazard Warning. *China Automation Congress (CAC)*, IEEE. https://doi.org/10.1109/CAC59555.2023.10450744
5. Zhao Y, Wang R, Jiang Y, et al. (2019) GIS-based logistic regression for rainfall-induced landslide susceptibility mapping under different grid sizes in Yueqing, Southeastern China. *Eng Geol* 259: 105147. https://doi.org/10.1016/j.enggeo.2019.105147
6. Guo Y, Dou J, Xiang Z, et al. (2024) Susceptibility evaluation of Wenchuan coseismic landslides by gradient boosting decision tree and random forest based on optimal negative sample sampling strategies. *Bull Geol Sci Technol* 43: 251–265. https://doi.org/10.19509/j.cnki.dzkq.tb20230037
7. Sarkar P, Mondal M, Sarkar A, et al. (2025) Landslide susceptibility assessment for the Darjeeling Toy Train route: a GIS and machine learning approach. *Stoch Environ Res Risk Assess* 39: 613–637. https://doi.org/10.1007/s00477-024-02885-y

8. Sun D, Chen D, Mi C, et al. (2023) Evaluation of landslide susceptibility in the gentle hill-valley areas based on the interpretable random forest-recursive feature elimination model. *J Geomech* 29: 202–219. https://doi.org/10.12090/j.issn.1006-6616.2022128
9. Liu S, Wang L, Zhang W, et al. (2023) A comprehensive review of machine learning-based methods in landslide susceptibility mapping. *Geol J* 58: 2283–2301. https://doi.org/10.1002/gj.4666
10. Juyal A, Sharma S (2021) A study of landslide susceptibility mapping using machine learning approach. *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*. IEEE, 1523–1528. https://doi.org/10.1109/ICICV50876.2021.9388379
11. Merghadi A, Yunus AP, Dou J, et al. (2020) Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Sci Rev* 207: 103225. https://doi.org/10.1016/j.earscirev.2020.103225
12. Wei R, Ye C, Sui T, et al. (2022) Combining spatial response features and machine learning classifiers for landslide susceptibility mapping. *Int J Appl Earth Obs Geoinf* 107: 102681. https://doi.org/10.1016/j.jag.2022.102681
13. Pham BT, Prakash I, Dou J, et al. (2020) A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. *Geocarto Int* 35: 1267–1292. https://doi.org/10.1080/10106049.2018.1559885
14. Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng Geol* 79: 251–266. https://doi.org/10.1016/j.enggeo.2005.02.002
15. Budimir MEA, Atkinson PM, Lewis HG (2015) A systematic review of landslide probability mapping using logistic regression. *Landslides* 12: 419–436. https://doi.org/10.1007/s10346-014-0550-5
16. Huang F, Chen B, Mao D, et al. (2023) Landslide susceptibility prediction modeling and interpretability based on self-screening deep learning model. *Earth Sci* 48: 1696–1710.
17. Feng Y, Xu H, Jiang J, et al. (2022) ICIF-Net: Intra-scale cross-interaction and inter-scale feature fusion network for bitemporal remote sensing images change detection. *IEEE Trans Geosci Remote Sens* 60: 1–13. https://doi.org/10.1109/TGRS.2022.3168331
18. Ding L, Hong D, Zhao M, et al. (2025) A survey of sample-efficient deep learning for change detection in remote sensing: Tasks, strategies, and challenges. *IEEE Geosci Remote Sens Mag* 13: 164–189. https://doi.org/10.1109/MGRS.2025.3533605
19. Youssef AM, Pradhan B, Dikshit A, et al. (2022) Landslide susceptibility mapping using CNN-1D and 2D deep learning algorithms: comparison of their performance at Asir Region, KSA. *Bull Eng Geol Environ* 81: 165. https://doi.org/10.1007/s10064-022-02657-4
20. Gao Z, Hu A, Chen B, et al. (2024) A hierarchical geometry-to-semantic fusion GNN framework for earth surface anomalies detection. Advances in Brain Inspired Cognitive Systems. BICS 2023. Lecture Notes in Computer Science, 14374. Springer, Singapore. https://doi.org/10.1007/978-981-97-1417-9_6
21. Wang Y, Fang Z, Hong H (2019) Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci Total Environ* 666: 975–993. https://doi.org/10.1016/j.scitotenv.2019.02.263

22. Sameen MI, Pradhan B, Lee S (2020) Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. *Catena* 186: 104249. https://doi.org/10.1016/j.catena.2019.104249
23. Huang F, Zhang J, Zhou C, et al. (2020) A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. *Landslides* 17: 217–229. https://doi.org/10.1007/s10346-019-01274-9
24. Chen C, Fan L (2022) CNN-LSTM-attention deep learning model for mapping landslide susceptibility in Kerala, India. *ISPRS Ann Photogramm Remote Sens Spatial Inf Sci* 10: 25–30. https://doi.org/10.5194/isprs-annals-X-3-W1-2022-25-2022
25. Zhang B, Tang J, Huan Y, et al. (2024) Multi-scale convolutional neural networks (CNNs) for landslide inventory mapping from remote sensing imagery and landslide susceptibility mapping (LSM). *Geomat Nat Hazards Risk* 15: 2383309. https://doi.org/10.1080/19475705.2024.2383309
26. Huang W, Ding M, Li Z, et al. (2023) Landslide susceptibility mapping and dynamic response along the Sichuan-Tibet transportation corridor using deep learning algorithms. *Catena* 222: 106866. https://doi.org/10.1016/j.catena.2022.106866
27. Wang D, Yang RH, Wang X, et al. (2023) Evaluation of deep learning algorithms for landslide susceptibility mapping in an alpine-gorge area: a case study in Jiuzhaigou County. *J Mt Sci* 20: 484–500. https://doi.org/10.1007/s11629-022-7326-5
28. Jamali A, Roy SK, Ghamisi P (2023) WetMapFormer: A unified deep CNN and vision transformer for complex wetland mapping. *Int J Appl Earth Obs Geoinf* 120: 103333. https://doi.org/10.1016/j.jag.2023.103333
29. Bao S, Liu J, Wang L, et al. (2022) Application of transformer models to landslide susceptibility mapping. *Sensors* 22: 9104. https://doi.org/10.3390/s22239104
30. Zhao Z, Chen T, Dou J, et al. (2024) Landslide susceptibility mapping considering landslide local-global features based on CNN and transformer. *IEEE J Sel Top Appl Earth Obs Remote Sens* 17: 7475–7489. https://doi.org/10.1109/JSTARS.2024.3379350
31. Zhu Q, Zhang M, Ding Y, et al. (2021) Fuzzy Logic Approach for Regional Landslide Susceptibility Analysis Constrained by Spatial Characteristics of Environmental Factors. *Geomatics Inf Sci Wuhan Univ* 46: 1431–1440.
32. Bian J, Li X, Hu K (2018) Study on distribution characteristics and dynamic evolution of mountain hazards in Hengduan Mountains area. *J Eng Geol* 26: 6–13. https://doi.org/10.13544/j.cnki.jeg.2018030
33. Chang Z, Huang F, Jiang S, et al. (2023) Slope unit extraction and landslide susceptibility prediction using multi-scale segmentation method. *Adv Eng Sci* 55: 184–195.
34. Data Center for Resources and Environmental Sciences, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. Spatial Distribution Data of Geological Disaster Points [DB/OL]. Beijing, 2025. Available from: http://www.resdc.cn/.
35. Guo Z, Wang H, He J, et al. (2025) PSLSA v2. 0: An automatic Python package integrating machine learning models for regional landslide susceptibility assessment. *Environ Model Softw* 186: 106367. https://doi.org/10.1016/j.envsoft.2025.106367
36. Wang Y, Sun D, Wen H, H. et al. (2020) Comparison of random forest model and frequency ratio model for landslide susceptibility mapping (LSM) in Yunyang County (Chongqing, China). *Int J Environ Res Public Health* 17: 4206. https://doi.org/10.3390/ijerph17124206

37. Aalianvari A, Jahanmiri S (2025) A Comparative Study of TabNet and Classical Machine Learning Models for Landslide Prediction [Preprint]. bioRxiv: https://doi.org/10.21203/rs.3.rs-7694229/v1
38. Aslam B, Zafar A, Khalil U (2023) Comparative analysis of multiple conventional neural networks for landslide susceptibility mapping. *Nat Hazards* 115: 673–707. https://doi.org/10.1007/s11069-022-05570-x
39. Lv L, Chen T, Dou J, et al. (2022) A hybrid ensemble-based deep-learning framework for landslide susceptibility mapping. *Int J Appl Earth Obs Geoinf* 108: 102713. https://doi.org/10.1016/j.jag.2022.102713
40. Ge Y, Liu G, Tang H, et al. (2023) Comparative analysis of five convolutional neural networks for landslide susceptibility assessment. *Bull Eng Geol Environ* 82: 377. https://doi.org/10.1007/s10064-023-03408-9
41. Chen T, Wang Q, Zhao Z, et al. (2024) LCFSTE: Landslide conditioning factors and swin transformer ensemble for landslide susceptibility assessment. *IEEE J Sel Top Appl Earth Obs Remote Sens* 17: 6444–6454. https://doi.org/10.1109/JSTARS.2024.3373029
42. Zhu Y, Liu S, Yin K, et al. (2025) Impact of negative sampling strategies on landslide susceptibility assessment. *Adv Space Res* 76: 592–613. https://doi.org/10.1016/j.asr.2025.04.070
43. Zhu Y, Yin K, Yang H, et al. (2026) Temporal validity of landslide inventories in hazard mapping: insights from Hubei Province, China. *Landslides* 23: 417–432. https://doi.org/10.1007/s10346-025-02651-3

AIMS Press