applied sciences

MDPI

Article

Potential Landslide Area Identification Method Based on Deep Generative Adversarial Reinforcement Learning (DGARL-LS)

Xiao-Nan Liu [1,*], Bao-Cai Yin [1], Jiang-Bi Hu [1] and Bing-Qi Li [2]

1 School of Information Science and Technology, Beijing University of Technology, Bejing 100124, China; m18610154187@163.com (J.-B.H.)
2 China Institute of Water Resources and Hydropower Research, Bejing 100038, China; libq@iwhr.com
* Correspondence: liuxn@emails.bjut.edu.cn; Tel.: +86-13661047462

Abstract

Landslides pose serious threats to lives and property in mountainous areas, yet large-scale preliminary screening of potential landslide areas remains costly and constrained by vegetation obstruction and data availability. Existing deep learning methods depend heavily on multi-source remote sensing data and dense annotations, limiting their applicability in data-scarce regions. This study proposes DGARL-LS, a potential landslide area identification method, which uses 10 m resolution DEMs as the sole data source and extracts slope and curvature information to construct curvature–slope (CS) stereoscopic maps. A pix2pix conditional GAN serves as the segmentation backbone, integrated with a bidirectional exploratory deep reinforcement learning module in which forward exploration localizes candidate regions and backward trajectory priors alleviate reward sparsity. GAN discriminator scores provide the reward signal, forming a closed-loop localization–segmentation–optimization framework. Training utilizes 15,000 augmented sample pairs annotated via a dual-constraint morphological–structural system in the Three Gorges Reservoir area. DGARL-LS achieves IoU of 76.8%, Dice of 83.5%, precision of 85.6%, and Recall of 80.2% on a 1800 km^2 test area, outperforming U-Net, pix2pix, Trans U-Net, and Swin U-Net, with approximately 25% faster convergence than U-Net. Cross-regional tests in Dazhou (Sichuan) adopt a small, unaugmented dataset for full retraining. DGARL-LS achieves 64.98% IoU, exceeding the second-best baseline (53.69%), while its absolute performance, especially IoU, decreases relative to the main research area. These results demonstrate that accurate, large-scale potential landslide screening is achievable using only freely available DEM data, providing a lightweight and transferable technical pathway for geohazard prevention in complex mountainous regions.

Keywords: digital elevation model; CS stereo map; potential landslide identification; conditional generative adversarial network; bidirectional exploratory reinforcement learning

Check for updates

Received: 12 May 2026
Revised: 1 July 2026
Accepted: 8 July 2026
Published: 13 July 2026

1. Introduction

Landslide refers to the geological hazard phenomenon in which a rock–soil mass on a slope undergoes overall downward sliding along a certain sliding surface under the action of gravity [1]. Its evolutionary process and scale vary with geological background, sliding mode, and morphological characteristics, featuring strong suddenness, great destructive power, and wide impact range. Although landslide formation is influenced by multiple factors, including rainfall and groundwater changes, seismic disturbances, and human engineering activities [2–5], slopes often develop a series of identifiable geomorphic response features before the occurrence of triggers, such as earthquakes and rainfall. These features can indicate that a slope is in a critical or metastable state, thus constituting precursor information for potential landslide occurrence. In engineering practice, particularly in transportation infrastructure and hydraulic engineering (e.g., highway slopes and reservoir bank slopes), such slopes are widely distributed and long subjected to the coupled effects of complex environments and engineering disturbances, making their stability especially prominent. Constrained by monitoring costs and technical conditions, it is difficult to implement comprehensive, long-term, and refined monitoring and stability analysis for all slopes. Therefore, prior to slope stability assessment and monitoring early warning, it is urgent to conduct preliminary screening and target area selection of the study region based on topographic and geomorphic features—that is, to identify key slope units with potential instability risks—so as to optimize the allocation of monitoring resources and advance risk prevention and control.

Potential landslide areas generally exhibit geomorphic patterns characterized by steep slopes, strong terrain undulation, and abrupt slope-form transitions, accompanied by a series of micro-geomorphic markers that can be quantitatively described by terrain factors—including rear-edge fractures, step-like gentle slopes, slope-toe bulging, and anomalous terrain curvature [6,7]. Effectively utilizing the geomorphic features before landslide occurrence and conducting preliminary identification of potential landslide areas through topographic and geomorphic characteristics is of great significance for hazard risk assessment, subsequent monitoring and early warning, and early prevention and control of disasters.

Digital elevation models (DEMs) are a core derivative product of the Earth observation system. Their data sources cover a variety of active and passive remote sensing platforms, including spaceborne synthetic aperture radar (SAR) interferometry (such as SRTM and TanDEM-X), airborne LiDAR scanning, and optical stereo image pair photogrammetry (such as ASTER GDEM and AW3D30) [8]. Compared with optical images, DEM data have three significant advantages: first, DEMs directly record the three-dimensional geometric shape of the Earth's surface and are not affected by vegetation cover, cloud cover, or lighting conditions, and thus have higher data reliability in complex terrain environments of southwestern mountainous areas; second, the topographic factors, such as slope and curvature derived from DEMs, are pure geometric quantities, eliminating the interference of seasonal phases and ground features caused by spectral information, and can more directly reflect the mechanical response state of the slope; and third, with the large-scale deployment of Earth observation satellite constellations, DEM data of 10 m or higher resolution can be obtained free of charge worldwide, providing a stable data foundation for large-scale, low-cost preliminary screening of potential landslides. Based on the above advantages, this paper selects 10 m resolution DEM data as the only data source and constructs CS stereo maps by extracting slope and curvature information, thus realizing the systematic identification of potential landslide areas without the need for multispectral or SAR time series data.

Numerous studies have shown that indicators such as slope, plane curvature, profile curvature and topographic change line can effectively reflect the open conditions and stress concentration state of the slope, and are the core basis for distinguishing between stable slopes and potentially unstable slopes [8–12], and can all be quantitatively expressed by DEM topographic geometric factors [13]. Among them, the slope reflects the first-order gradient of elevation and directly quantifies the shear driving force of gravity along the slope. Specifically, Zou et al. analyzed the topographic features of 46 landslide samples in Wanshan District [14] and found that the landslides with a slope in the range of 25–30° accounted for 60.8%. The soil thickness in this range is moderate, the matrix suction is strong, the groundwater effect is obvious, and the risk of landslide is high. Bakhsoliani [15] confirmed that the profile curvature characterizes the slope change rate along the direction of water flow. Positive values (convex shape) correspond to the high and steep slope of the landslide source area, while negative values (concave shape) correspond to the accumulation toe. Adding curvature information can significantly improve the identification accuracy of potential landslide areas. Planar curvature describes the degree of surface curvature perpendicular to the slope direction. The negative value area corresponds to the catchment depression. If the soil is in a high saturation state for a long time, the pore water pressure increases significantly, and the effective stress continues to decrease. The synergistic effect of the above three types of topographic factors constitutes the physical mechanism chain of landslide instability, laying the foundation for the geophysical basis of CS stereo maps as the model input. Therefore, this study uses a 10 m resolution digital elevation model (DEM) as the basic data source, systematically extracts slope and curvature information, and constructs a CS stereo map, which serves as the core terrain representation for potential landslide identification.

In recent years, with advances in remote sensing data acquisition and artificial intelligence, deep-learning-based potential landslide identification methods have progressively shifted from traditional empirical interpretation toward data-driven and intelligent approaches. Existing research mainly focuses on multi-source data fusion, spatial feature extraction, and temporal evolution modeling. The integration of InSAR technology with deep learning models has become an important direction for potential landslide identification. For example, Yang et al. [16] used PS-InSAR, SBAS-InSAR, and optical remote sensing combined with deep learning to successfully identify 56 potential landslide sites in Jianzha County, Qinghai Province, achieving an identification accuracy of 76.92%. He et al. [17] proposed a comprehensive potential landslide susceptibility assessment method integrating TD-CNN-Bi-GRU temporal feature extraction with MSCNN spatial feature extraction, achieving an AUC of 0.91 by fusing InSAR deformation dynamics and spatial features of landslide influencing factors. Li et al. [18] proposed MSFD-Net, a multi-source data fusion network using a non-weight-sharing pseudo-Siamese structure to separately extract deformation phase texture features and topographic features, identifying 254 landslides in Eastern Qinghai Province. Ma et al. [19] proposed a landslide identification method based on a Dual Graph Convolutional Network (DGCNet) using GeoEye-1 high-resolution remote sensing imagery in the mountainous area of Xinyuan County, Xinjiang. Li et al. used a CNN to integrate terrain, geological, hydrological, and climatic variables for high-precision landslide susceptibility mapping in a region of Iran [20]. More recently, generative approaches have also been explored to alleviate sample scarcity in landslide identification. Tong et al. [21] employed a deep convolutional generative adversarial network (DCGAN) to augment landslide inventories in Western Sichuan, improving the accuracy of downstream SVM, CNN, and ResNet classifiers by 2.2–5.3%, which corroborates the value of GAN-based data augmentation adopted in this study. Wang et al. [22] proposed MSTA-YOLO, incorporating a receptive-field attention module to address blurred boundaries and diminished texture features of landslides caused by vegetation cover and weathering, and achieved notable gains on multi-scale landslide detection benchmarks. Li et al. [18] further demonstrated that incorporating topographic features alongside InSAR deformation signals improves wide-area landslide detection robustness, reinforcing the rationale for using DEM-derived terrain factors as a primary identification basis. A recent comprehensive review [23] also highlighted these deep generative models—including GANs, VAEs, and diffusion models—as particularly suited to addressing limited labeled data and providing uncertainty quantification in landslide risk assessment, further motivating the generative adversarial component of the proposed DGARL-LS framework.

Although the aforementioned deep-learning-based methods improve identification accuracy to some extent, model performance is significantly influenced by input image quality, and there is a strong dependence on annotated datasets and long-term monitoring data—applicability is limited in small-sample datasets or vegetation-covered areas. To address this, generative approaches such as generative adversarial networks (GANs) provide new technical paths for landslide identification. Compared with traditional discriminative models, GANs can learn complex terrain structure features from the data distribution level. Among these, pix2pix, as a typical conditional GAN (cGAN), can model the mapping relationship from input terrain factors to target spatial distributions, showing good application prospects in terrain-driven potential landslide identification. For example, Hirokazu et al. [24] applied pix2pix to generate terrain interpretation results, with a landslide location identification accuracy up to 80%. Feng et al. [25] used GANs to generate high-quality landslide samples, effectively alleviating the problem of insufficient training data. Al-Najjar and Pradhan [26] indicated that GAN-generated samples can significantly improve class imbalance problems, thereby enhancing model generalization. Nevertheless, existing GAN research mostly focuses on data augmentation or image generation, lacking in-depth exploration of learning spatial patterns of potential landslides driven by terrain factors. More critically, pix2pix and similar GANs face three inherent limitations in potential landslide identification: (1) pix2pix adopts a fixed full-image segmentation strategy, which is highly sensitive to background interference and prone to producing numerous false detections in vegetation-covered or topographically complex areas; (2) its interpretation threshold is fixed after training and cannot adaptively adjust to the characteristic differences in different terrain regions, limiting cross-regional generalization; and (3) under extremely scarce annotated sample conditions, purely supervised learning fails to fully mine the terrain distribution patterns in unlabeled CS maps, resulting in low model learning efficiency. The essence of these problems is that potential landslide identification is not merely an image translation task but a sequential decision-making process that requires dynamic search and continuous optimization of interpretation strategies in uncertain environments.

To address the above issues, this paper introduces a reinforcement learning (RL) mechanism as a dynamic optimization layer for pix2pix. Since RL is well-suited for "learning optimal strategies through trial and error in uncertain environments," the landslide area search and identification process can be modeled as a sequential decision problem—where an agent iteratively adjusts candidate window positions and scales in the CS stereoscopic map, using the GAN discriminator output score as environmental feedback reward to progressively learn an exploration strategy for "key areas," overcoming the background-noise sensitivity of pix2pix's fixed full-image strategy. Meanwhile, the policy optimization mechanism of RL enables the model to dynamically search and reposition the candidate detection window based on different terrain features instead of relying on a single fixed segmentation region, improving cross-regional generalization. Furthermore, the bidirectional exploration mechanism (forward exploration + backward priors) can still effectively learn under scarce annotated sample conditions, compensating for the data bottleneck of purely supervised learning. Based on this, this paper proposes the DGARL-LS method, which uses DEM data to construct a slope–curvature (CS) terrain representation system, with pix2pix as the generative segmentation backbone network, and bidirectional exploration DRL as the dynamic search and policy optimization module. Through a cooperative feedback mechanism, the three are integrated into a closed-loop learning framework of "region localization–fine segmentation–feedback optimization," effectively compensating for the shortcomings of traditional methods in geomorphic structure feature representation and adaptive interpretation, providing a technical path for large-scale, efficient identification of potential landslide areas in complex mountainous regions.

2. Topographic Features of Potential Landslide Areas and Their Data Representations

2.1. Topographic Response Mechanism of Potential Landslides

Landslide is a geological process in which the slope body undergoes progressive failure along a potential sliding surface under gravity, ultimately becoming unstable. Its formation is controlled by the coupled effects of topographic structure, rock and soil mechanical properties, hydrological conditions, and external disturbances. Unlike sudden disasters, landslides often undergo an evolution process of "stable–local damage–progressive deformation–overall instability." During this process, the topographic morphology undergoes observable geometric responses first, becoming an important basis for potential landslide identification.

From a geomorphic structure perspective, landslides typically consist of source areas, sliding areas, and accumulation areas (Figure 1), with significant differences in topographic geometric features among different zones. The source area is usually located in the upper part of the slope, exhibiting elevation discontinuities and tension crack characteristics. The sliding area corresponds to continuous steep slope sections with large slope angles and strong topographic consistency. The accumulation area exhibits reduced slope angles and concave curvature characteristics. These topographic features are generally not the result of landslide occurrence; instead, they form gradually during the landslide incubation stage as "prior geometric signals." Therefore, analysis based on topographic and geomorphic structure is more suitable for early interpretation of potential landslides.

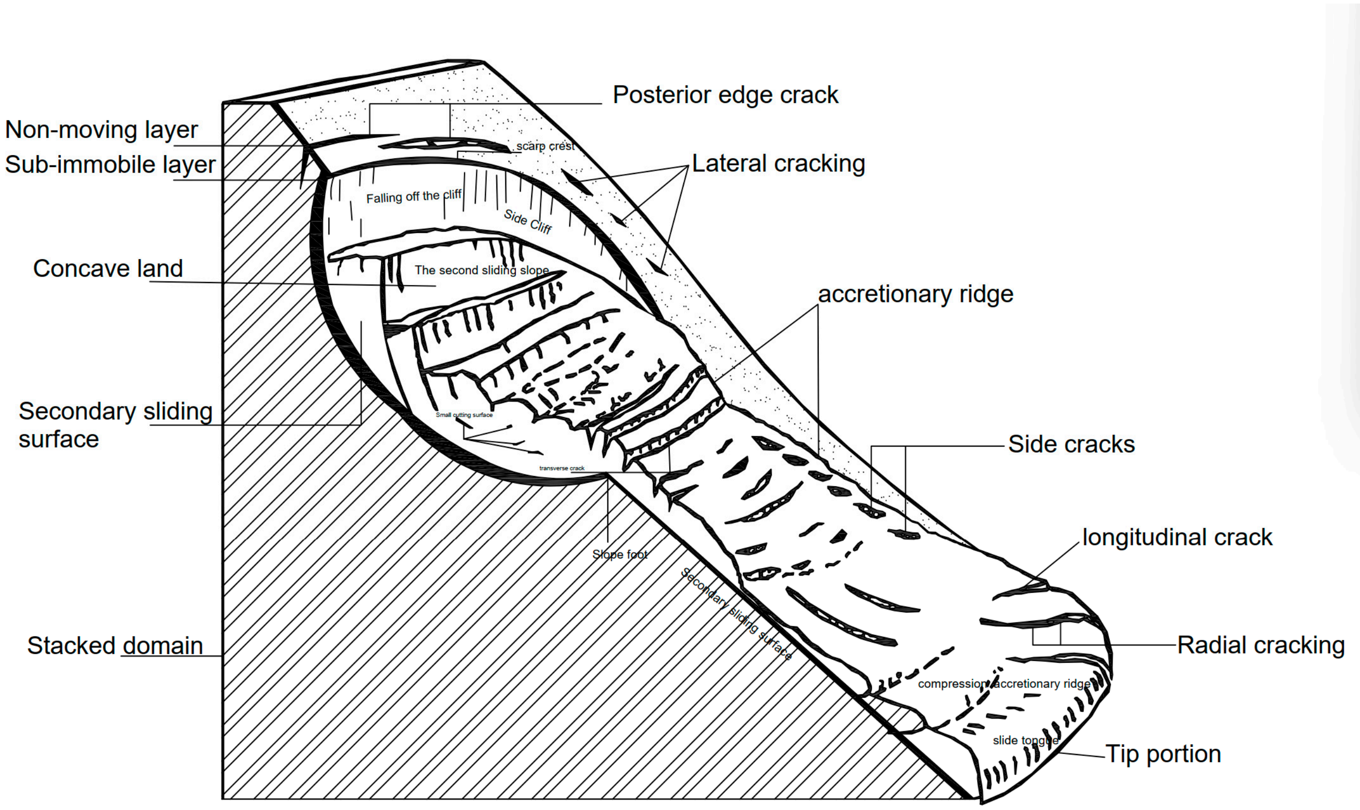

Figure 1. Schematic diagram of potential landslide cross-section.

2.2. Learnable Topographic Feature Representation of Potential Landslides

Traditional landslide identification methods mostly rely on expert experience when interpreting single terrain factors, such as slope, curvature, and aspect. However, such methods are difficult to characterize the non-linear features of multi-factor coupling under complex terrain conditions, limiting identification accuracy and generalization ability.

To adapt to the development of data-driven methods, potential landslide identification needs to shift from "single indicator interpretation" to "multi-dimensional feature representation." The basic topographic components of potential landslides include the source area, sliding body, sliding surface, sliding zone, accumulation area, and tension crack zone [27]. Ma et al. [28] analyzed topographic features before and after landslides and found that obvious erosion traces exist on steep slopes before landslide occurrence. These areas have quantifiable topographic features, with typical characteristics mainly including cracks, slope-angle mutations, curvature anomalies, slope-toe disturbances, and elevation reversals in source areas (Table 1, Figure 2). Among these, the micro-topography within the slip scarp and the interior of the landslide body are most typical. These features reveal anomalous changes in the internal structure and stress field of the slope body, providing important basis for early identification of potential landslides. This paper treats these as learnable implicit feature patterns to be automatically extracted by deep learning models in terms of their combinational relationships, avoiding uncertainties arising from empirical dependence.

Table 1. Potential landslide topographic characteristics.

Location	Description	Characteristics
Tension crack zone	Upper zone of potential landslide initiation, usually a zone of fractures and cracks	Obvious slip scarps can usually be captured in topographic maps
Sliding zone	Main sliding body, often accompanied by deformation	Large slope angle (>25°), positive curvature (convex), high elevation
Slip bed/sliding surface	Sliding surface, possibly developed along weak interlayers or geological interfaces	Moderate-to-high slope angle (15–35°), coherent downslope terrain
Accumulation zone	Final accumulation area of the sliding body, with a reduced slope angle	Negative curvature zone (concave), DEM shows a transitional or abrupt band

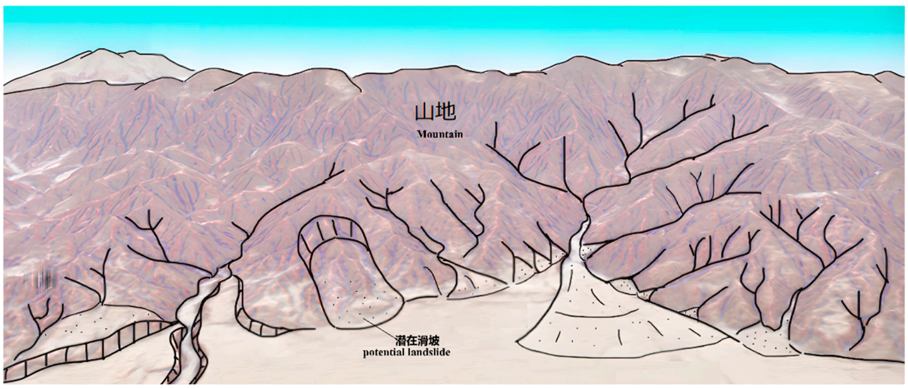

Figure 2. Schematic diagram of potential landslide topographic features.

2.3. CS Stereoscopic Map Based on DEMs and Its Value for Landslide Interpretation

The core basis for this paper to use DEM data for potential landslide area identification is that terrain morphology is a direct reflection of the incubation and evolution processes of landslides. The surface elevation variation expressed by DEMs can reveal the spatial continuity and geometric features of geomorphic structures, which are closely related to landslide formation and development; their geometric distribution often presages potential instability trends. Numerous studies have shown that parameters such as slope, curvature, and terrain relief extracted from DEMs can effectively capture the spatial morphological features of potential landslide areas and provide quantitative indicators for potential landslide identification [29–32]. Furthermore, comparative analysis of long-term landslide inventories shows that the spatial distribution of landslides in mountainous areas of Yunnan, Chongqing, and Sichuan is highly correlated with terrain factors such as slope and curvature. Therefore, DEM-based candidate area interpretation can serve as an efficient preliminary screening method.

Based on the above geomorphological mechanisms, this paper derives slope and curvature parameters from DEM data to construct a curvature–slope (CS) stereoscopic map that amplifies topographic geometric anomalies to form high-contrast spatial features for early preliminary screening and automated identification of potential landslide areas. For example, in the CS stereoscopic map, the lateral steep scarps on both sides of a sliding body appear in blue and converge at the top, with the overall morphology typically forming a horseshoe shape; numerous horizontal cracks and small ridges, as well as secondary sliding surfaces and cracks, are visible inside the landslide body; and deeply incised gullies formed by erosion on the sliding body surface are few, with the overall appearance in a light white tone—which is also a typical characteristic (Figure 3).

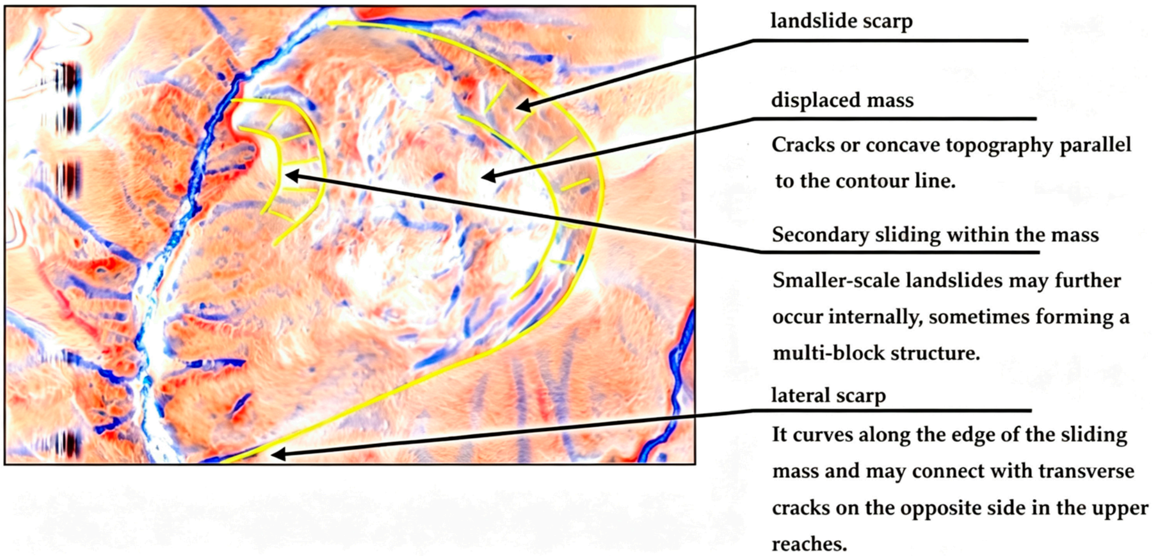

Figure 3. Schematic diagram of topographic features in potential landslide areas in CS stereoscopic maps.

For data source convenience, this paper introduces the CS stereoscopic map (curvature–slope) constructed from DEM data with vegetation layers and other information removed. It builds a visualization image integrating slope and curvature information, achieving quantitative representation of potential landslide areas through analysis of terrain morphology and topographic features, and provides a reliable basis for potential landslide area identification. Compared with single terrain factors, the CS representation couples slope and curvature information with the following advantages:

(1) Enhanced identifiability of topographic anomaly features, simultaneously characterizing both the steepness of the slope surface and the concave–convex morphology.
(2) Highlighted potential landslide geometric features, with higher sensitivity to slope-angle mutations and curvature anomalies.

(3) Reduced data dependency, as no spectral or texture information required and can be constructed solely from DEMs.
(4) Compatibility with deep learning model input that is convertible to high-contrast images suitable for convolutional neural network modeling.

2.4. CS Stereoscopic Map Production

The CS stereoscopic map is produced by first extracting slope and curvature information from DEM data through the ArcGIS platform (Version 10.8.2), then combining color mapping and layer overlay to spatially represent the slope steepness and concave–convex morphology at the image level. The CS stereoscopic map production workflow is shown in Figure 4; this section takes a landslide-prone area in Chongqing as an example (Figure 5) with the following procedure:

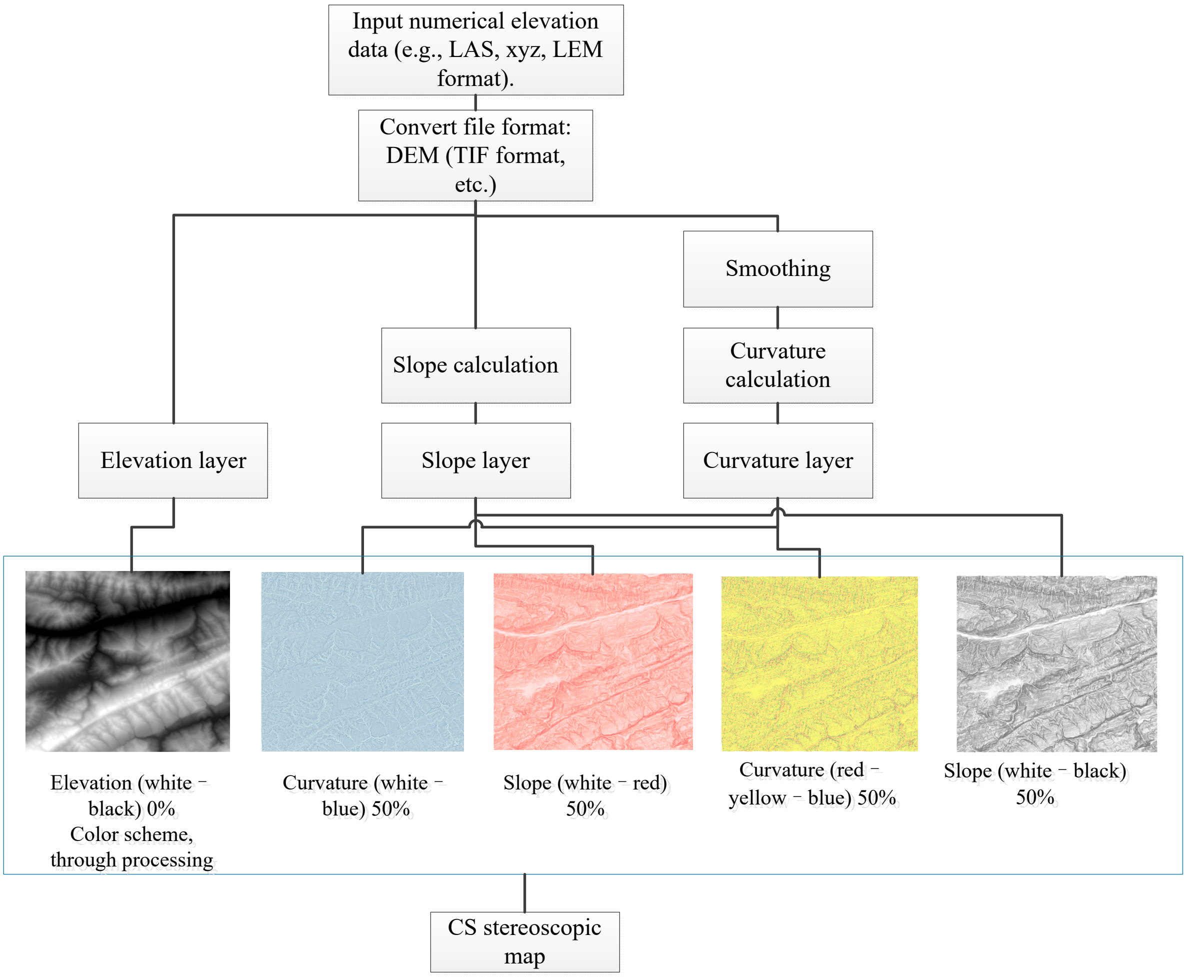

Figure 4. Production workflow of the curvature–slope (CS) stereoscopic map.

Step 1: Download DEM data (spatial resolution: 10 m) from a data portal.
Step 2: Load the DEM into ArcGIS software.
Step 3: Generate the elevation layer directly from the raw elevation data of the DEM.
Step 4: Calculate the terrain slope (inclination angle) from DEM data to generate the slope layer.
Step 5: Apply smoothing to the DEM to reduce noise, then calculate curvature to obtain the curvature layer.
Step 6: Combine the elevation, slope, and curvature layers into a 3D stereoscopic map to display detailed terrain features.

Step 7: Synthesize the slope and curvature layers into a curvature map that emphasizes slope and curvature features.

Step 8: Merge the 3D stereoscopic map from Step 6 and the curvature map from Step 7 to form the CS stereoscopic map (Figure 5), realizing visualization of detailed geomorphic structure. Colors can be adjusted to user-selected schemes. It should be noted that the 50% layer blending ratio and the specific color scheme (white–black for elevation, white–red for slope, white–blue/red–yellow-blue for curvature) were chosen empirically with reference to prior CS map visualization practices [33], keeping the slope and curvature layers visually balanced without either one visually dominating or obscuring the other; these settings were not systematically tuned or compared against alternative ratios or color schemes in this study.

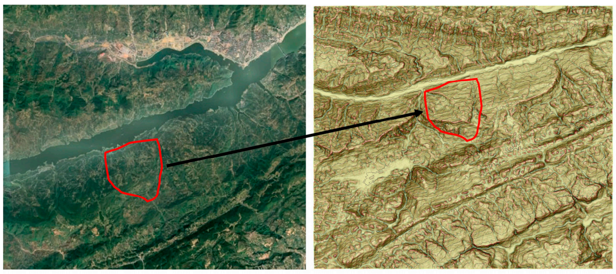

Figure 5. A landslide-prone area in Chongqing and an example of a CS image. The black arrow denotes the corresponding location, and red-circled areas indicate prominent large-scale potential landslides.

3. Potential Landslide Area Identification Method Based on DGARL-LS

3.1. Overall Framework

To leverage the advantages of the slope–curvature stereoscopic map in terrain feature representation and combine the outstanding performance of generative adversarial neural networks in image processing, this paper proposes potential landslide area identification method based on deep generative adversarial reinforcement learning (DGARL-LS). By constructing a mapping relationship between "terrain maps and potential landslide maps," efficient identification and automated extraction of landslide risk areas are achieved (Figure 6).

Since DEM-derived factors, such as slope, curvature, and elevation, have been confirmed in multiple studies in the Sichuan–Yunnan mountainous region to have significant statistical correlation with the spatial distribution of landslides, a DEM is used as a preliminary screening tool. By merging slope, curvature, and elevation information—validated by the existing literature's conclusions—"potential landslide areas" are identified. These areas are not assertions of imminent landslides but rather target zones designated for subsequent risk area monitoring, field investigation, and coupling with triggering factors in early warning systems, thereby improving monitoring efficiency and targeting when resources are limited.

This paper temporarily does not consider natural factors such as rainfall, identifying potential landslide areas solely from the terrain perspective. In subsequent research, triggering factors such as rainfall and earthquakes will be incorporated to further study the probability of landslide occurrence in potential landslide areas. This paper takes the Wushan–Fengjie–Yunyang area of the Three Gorges Reservoir region in Chongqing as the primary study area, using a DEM as the primary data source. Multi-scale terrain factors (such as slope and curvature) are used to construct a CS stereoscopic map sample dataset. A GAN and a deep reinforcement learning (DRL) module are then integrated: the generator produces segmentation masks of potential landslide areas from CS terrain images; the discriminators evaluate the consistency between generated masks and true landslide distributions; and the RL module dynamically adjusts candidate window positions to optimize the detection strategy, achieving accurate identification of potential landslide areas.

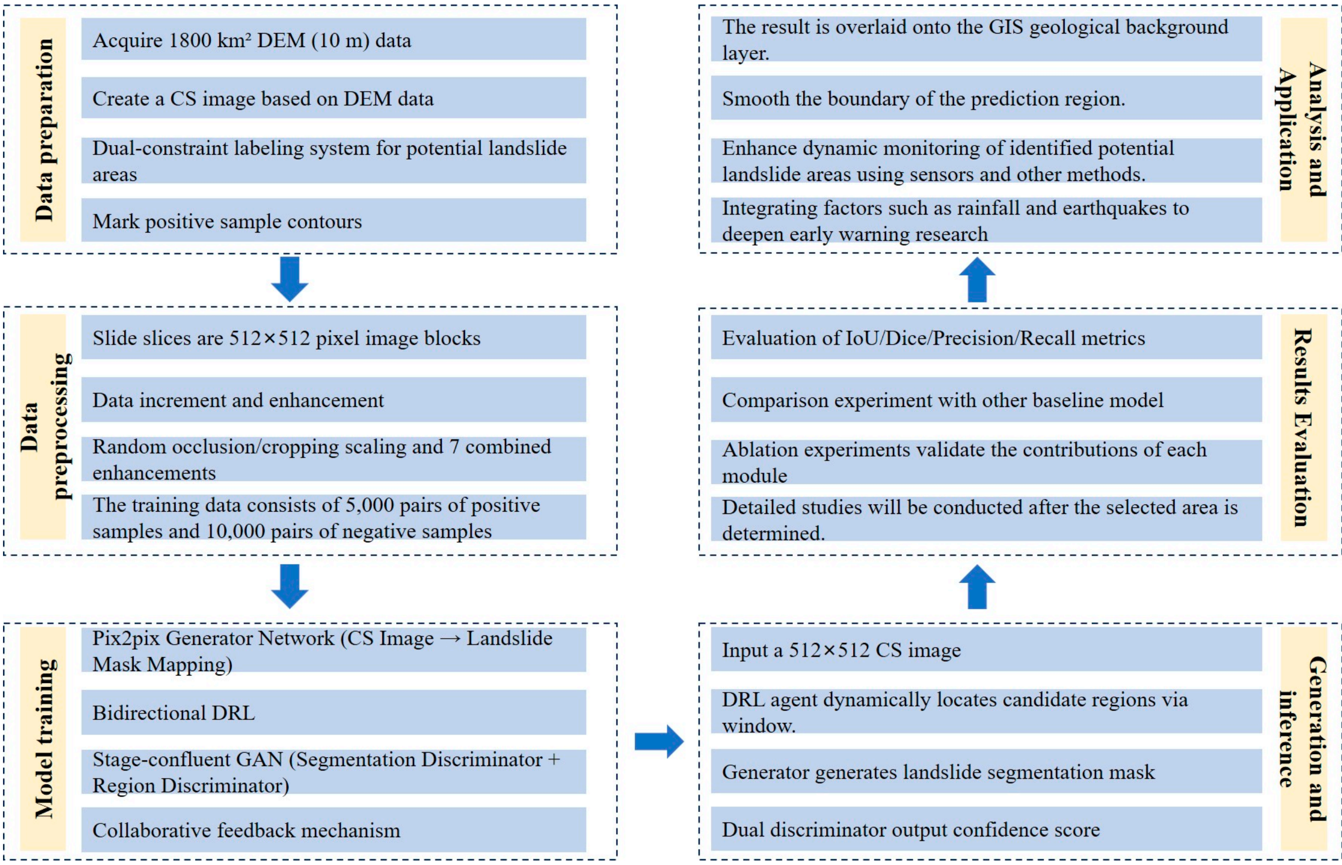

Figure 6. Potential landslide area technical identification framework based on DGARL-LS.

During training, the actor–critic architecture is adopted to design the RL agent, introducing a reward function based on terrain feature changes to enhance model sensitivity to potential landslide areas. The overall model consists of three parts: a data preprocessing and sample construction module, a terrain image generative network based on pix2pix, and an RL agent that uses the discriminator output as environmental state. Training employs a warm-up phase for the first 100 epochs (learning rate fixed at the initial value), followed by Cosine Annealing scheduling to enhance model stability and convergence speed.

3.2. Model Architecture

The DGARL-LS framework proposed in this paper for early identification and interpretation of potential landslide-prone areas consists of two modules: the bidirectional exploration DRL module and the stage-confluent GAN (Figure 7). A cooperative feedback mechanism is also incorporated, in which the GAN discrimination results serve as reward feedback to guide the RL module's dynamic adjustment of candidate windows, while the RL provides high-confidence candidate regions to reinforce GAN training, thus realizing closed-loop learning of "region localization–fine segmentation–feedback optimization".

The overall data flow is as follows: the input CS stereoscopic image I is first received by the bidirectional exploration DRL module, which outputs a candidate window C_t; the sub-image cropped by C_t is then passed to the stage-confluent GAN generator, which outputs a pixel-level segmentation mask M of the same spatial size as the cropped window; and the dual discriminators evaluate both the cropped window placement and the generated mask, and then their combined output is converted into a scalar reward that is fed back to the RL agent to update its policy for the next time step. This closed loop is repeated for a fixed number of steps per episode before the final detection window and segmentation mask are output as the potential landslide area.

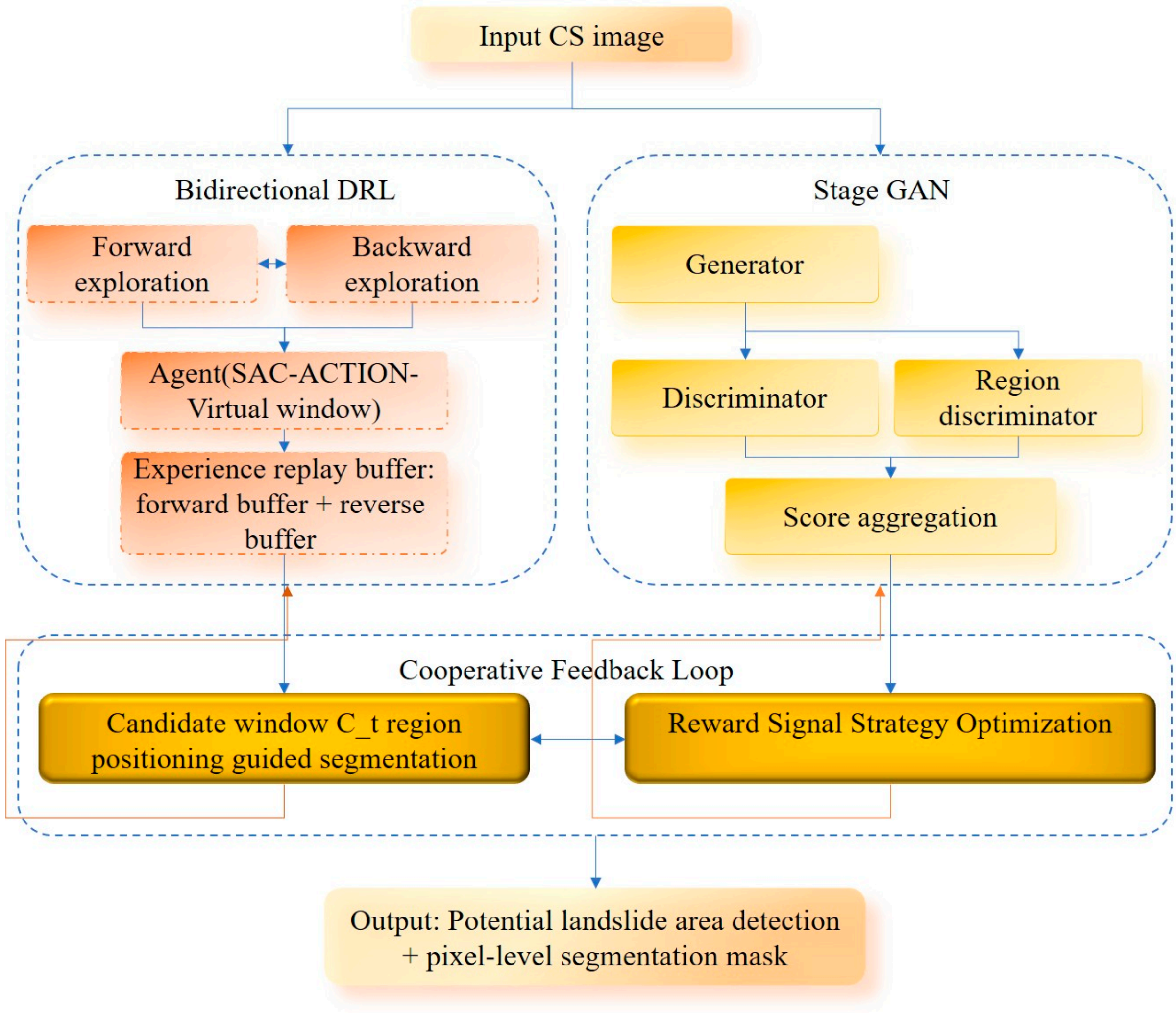

Figure 7. Schematic diagram of the deep generative adversarial reinforcement learning (DGARL-LS) model.

3.2.1. Bidirectional Exploration DRL

The topographic and structural features of potential landslide areas often suffer from blurred textures, large-scale differences, and sample scarcity. Traditional deep reinforcement learning (DRL) methods based on unidirectional forward exploration are difficult to converge in reward-sparse scenarios. To address this, this paper introduces bidirectional exploration DRL, combining forward and backward exploration so that the agent can still acquire effective experience in the absence of explicit rewards, thereby improving the stability and accuracy of landslide detection.

A. Forward Exploration

Forward exploration starts from the central region of the input CS stereoscopic image. The agent iteratively moves and scales a virtual window to approach the potential landslide area. The trajectory can be formally expressed as:

$$\tau_{forward} = (s_0, \alpha_0, \gamma_1, s_1, \alpha_1, \gamma_2, s_2, \alpha_2, \gamma_3, \ldots) \tag{1}$$

where s_t denotes the state (consisting of the current window position and the CS stereoscopic map), α_t is the action (window translation or scaling), and γ_t is the reward provided by the discrimination module.

B. Backward Exploration

Backward exploration starts from a known or high-confidence landslide candidate area, simulates the terminal detection state, and then generates exploration experience through reverse operations. The process can be expressed as:

$$\tau_{backward} = (\ldots, \gamma_{i-1}, s_{i-1}, \alpha_{i-1}, \gamma_i, s_i, \alpha_i) \tag{2}$$

where the agent executes the reverse action $-\alpha_{i-1}$ to obtain the previous state s_{i-1} and calculate the corresponding reward:

$$\gamma_i = R(s_{i-1}, \alpha_{i-1}) \tag{3}$$

This approach can still provide learnable prior experience in the absence of explicit rewards, alleviating the reward sparsity problem.

C. State, Action, and Reward Function Design

(1) State:

$$s_t = [C_t, I] \tag{4}$$

where I denotes the input CS image, and C_t is the binary mask corresponding to the window.

(2) Action:

$$a_t = \left[\alpha_t^v, \alpha_t^h, \alpha_t^{sv}, \alpha_t^{sh}\right] \tag{5}$$

where a denotes the action, representing vertical/horizontal translation and length/width scaling of the window.

(3) Reward function: The reward is provided by the generative discrimination module, comprehensively measuring the degree of match between the window region and the true distribution of potential landslides,

$$r_t = Score[I, C_t] \tag{6}$$

where *Score* is provided by the discriminator.

(4) Learning objective: The agent aims to maximize the cumulative reward through the optimal policy π^* to approach the potential landslide area,

$$\pi^* = argmax_\pi \mathbb{E}_\pi \left[\sum_{t=0}^{T} \eta^t r_t\right] \tag{7}$$

where η is the discount factor used to balance immediate and long-term rewards. Through the bidirectional exploration mechanism, the agent maintains stable exploration under reward-sparse conditions, converges rapidly to landslide candidate areas via backward prior experience, and improves detection accuracy for blurred-texture and small-scale landslide targets.

D. Agent Network Architecture (SAC-ACTION)

The SAC-ACTION agent implements the bidirectional exploration policy using soft actor–critic (SAC), an off-policy and maximum-entropy algorithm. The entropy regularization mechanism enables sufficient exploration of the candidate window action space, which effectively avoids premature convergence to suboptimal window solutions under the reward-sparse scenario of potential landslide detection tasks.

State encoding: The current candidate window is characterized by a normalized coordinate vector to represent its position and size on the image plane. Meanwhile, the CS stereoscopic image is fed into a shared lightweight convolutional feature extractor to obtain compact image feature representations. The normalized window coordinate information and extracted image features are concatenated to form a unified state embedding, which is subsequently used for subsequent network inference.

Actor network: The integrated state embedding is processed by stacked fully connected hidden layers with non-linear activation functions. The output layer parameterizes a diagonal Gaussian distribution over the action space, generating the statistical parameters required for action sampling. All action components corresponding to window translation and scaling are sampled via the reparameterization trick, constrained to valid bounded ranges, and finally rescaled to update the candidate window parameters.

Value network: Each critic network takes the concatenation of state embedding and sampled action vector as input. The network adopts stacked fully connected hidden layers to extract high-level feature representations and outputs a soft Q-value for state–action pair evaluation. The dual critic networks share the convolutional feature extraction module with the actor network while retaining independent fully connected layers to ensure adequate evaluation diversity.

Experience replay buffers: Two independent replay buffers are constructed to store transition samples generated from forward exploration and backward exploration. In each gradient update iteration, mini training batches are sampled from both buffers at a fixed proportion. This training strategy combines reliable prior knowledge from backward exploration and real interactive experience from forward exploration, which prevents the policy network from overfitting to single backward prior trajectories and improves the generalization ability of the model.

3.2.2. Stage-Confluent GAN

To address the problems of limited training samples, blurred textures, and complex negative samples in landslide identification, this paper introduces the stage-confluent GAN. This network combines a generator with a dual-discriminator structure, enabling fine segmentation of candidate areas within virtual windows while simultaneously evaluating the rationality of detected areas, thus achieving joint optimization of detection and segmentation.

A. Network Structure

Stage-confluent GANs consist of three components:

The generator G receives the virtual window G_t and input terrain image I, and generates the segmentation mask M of the potential landslide area through a convolutional–deconvolutional network.

The segmentation discriminator D_S determines the consistency of the generated segmentation result I, M with the manually annotated I, y.

The region discriminator D_d determines whether the candidate window $I,\ G_t$ matches the true landslide area window I, B. The dual discrimination mechanism enables the network to simultaneously optimize segmentation accuracy and detection area rationality.

Specifically, the generator G adopts the U-Net encoder–decoder architecture from the original pix2pix framework. The encoder comprises successive down-sampling convolutional blocks equipped with standard normalization and leaky activation functions, and the number of feature channels gradually increases layer by layer to compress the resized input window into compact bottleneck feature representations. The decoder is symmetrically designed to match the encoder, consisting of up-sampling deconvolutional blocks with standard activation functions, and dropout regularization is applied to shallow decoder layers during training. Skip connections are utilized to concatenate feature maps from encoder layers and corresponding decoder layers with identical spatial resolution, which retains delicate boundary features that tend to disappear at the bottleneck stage. The final decoder layer adopts tanh activation to output a single-channel segmentation mask M.

Both the segmentation discriminator D_S and region discriminator D_d employ the PatchGAN architecture proposed in pix2pix. Instead of performing global discrimination on the whole image, this structure judges local overlapping patches, so it can effectively capture high-frequency texture features closely related to landslide boundaries. Each discriminator stacks multiple convolutional layers with standard normalization and leaky activation, and the last convolution layer generates a single-channel probability map, where each position corresponds to the true/fake judgment result of a local image patch. D_S takes the concatenation of CS image patches and segmentation masks (generated or ground-truth) as input, while D_d receives the combined features of CS images and candidate window regions. The two discriminators share identical network layouts but are trained with separate weight parameters, as they are responsible for evaluating different parts of the overall detection and segmentation workflow.

B. Score Feedback Mechanism

Stage-confluent GANs synthesize the segmentation and region evaluation results into a reward for the bidirectional exploration DRL:

$$Score = (D_d(|C, I|) + D_s(|G(C \odot I)|, I)) \times \alpha \tag{8}$$

where D_d guides the detection window toward potential landslide areas; D_s guides the detection in selecting window perspectives favorable for segmentation; α is a weighting coefficient empirically set to three; and $\odot$ denotes the element-wise product of the window and the original CS stereoscopic map.

C. Loss Function Design

The generator and discriminators employ different loss functions for pixel-level labeled data (positive samples) and weakly labeled data (negative samples with blank masks), realizing weakly supervised learning. Generator loss is calculated as follows:

$$L_G = \begin{cases} L_T(I_l, C_t, y_l) + L_{BCE}(I_l, C_t, 1), & \text{for labeled data} \\ L_{BCE}(I_u, C_t, 1), & \text{for unlabeled data} \end{cases} \tag{9}$$

where L_T is the Tversky loss function, used to balance the positive–negative sample imbalance in landslide identification:

$$L_T(I, C, y) = Tversky((C \odot I), y) \tag{10}$$

Discriminator losses are calculated as:

$$L_{D_s} = L_{BCE}(D_s(I_l, C_t), 0) + L_{BCE}(D_s(I_l, y_1), 1) \tag{11}$$

$$L_{D_d} = L_{BCE}(D_d(I_l, C_t), 0) + L_{BCE}(D_d(I_l, B), 1) \tag{12}$$

where L_{D_s} is the segmentation discriminator loss, and L_{D_d} is the region discriminator loss.

Stage-confluent GANs prevent segmentation networks from being disturbed by irrelevant backgrounds (non-potential landslide areas); offer dual-constraint advantages of simultaneously optimizing "detected regions" and "segmentation results" to reduce false detections; effectively utilize unannotated CS stereoscopic map data to reduce annotation costs; and leverage Tversky loss to improve identification of blurred landslide textures and small-scale landslides.

3.2.3. Cooperative Feedback Mechanism

In potential landslide identification, relying solely on detection or segmentation networks often leads to information asymmetry—the detection stage may produce deviations, while the segmentation stage cannot effectively constrain the detected regions. Therefore, this paper introduces a cooperative feedback mechanism between the bidirectional exploration DRL and the stage-confluent GAN, realizing bidirectional optimization of detection and segmentation.

A. Basic Principles

Detection-driven segmentation: The bidirectional exploration DRL generates virtual windows of candidate areas through iterative actions, guiding the stage-confluent GAN to perform fine segmentation in specific areas, avoiding background interference from global search.

Segmentation feeding back into detection: The stage-confluent GAN evaluates the quality of segmentation results (provided by dual discriminators) and converts them into reward signals to be fed back to the RL agent, correcting deviations during exploration.

This mutually beneficial cyclic optimization continuously improves the accuracy of both detection and segmentation while reducing error accumulation.

B. Mathematical Formulation

At each time step t with state s_t, a_t, the GAN evaluates the segmentation result of candidate region C_t and gives an evaluation score calculated by:

$$r_t = Score(I, C_t) \tag{13}$$

where $Score$ integrates the outputs of the region discriminator D_d and segmentation discriminator D_s:

$$Score = (D_d([C_t, I]) + D_s([G(C_t \odot I), I])) \times \alpha \tag{14}$$

This score serves both as a supervision signal for the GAN and as a reward function for the DRL, realizing cooperative optimization of detection and segmentation.

C. Cooperative Feedback Workflow

Input data: CS stereoscopic images are generated.

Forward–backward exploration: DRL initializes virtual windows and progressively adjusts window position and scale.

Local segmentation: The stage-confluent GAN generates segmentation results in the window region.

Discrimination evaluation: The segmentation discriminator D_s and region discriminator D_d simultaneously evaluate detection and segmentation performance.

Reward feedback: The comprehensive discrimination result forms score r_t as the DRL reward signal.

Policy update: DRL updates the policy based on feedback signals, progressively approaching the optimal detection window for the landslide area.

The cooperative feedback mechanism reduces the negative impact of detection errors on subsequent segmentation; conversely, the dynamic adaptability of rewards varying with segmentation results enables DRL convergence even in reward-sparse environments, enabling effective identification of potential landslide areas under complex terrain conditions (vegetation cover, noise interference, blurred textures).

4. Experiments

4.1. Sample Data

Considering the convenience of data acquisition and the consistency of the study area, the training and test data in this paper cover the Wushan–Fengjie–Yunyang area of

Chongqing, China, within the coordinate range of (30.8000° N, 109.2000° E) to (31.2000° N, 109.6000° E). Downloaded data consist of 10 m resolution DEMs as the primary data source. To avoid the influence of vegetation information on model training and identification, vegetation information was removed during data download to obtain bare-terrain data. The 10 m resolution can effectively capture the micro-terrain features of slope surfaces, meeting the accuracy requirements for potential landslide identification. Compared with lower-resolution DEM data, the 10 m resolution achieves a good balance between computational efficiency and terrain detail expression; can finely represent topographic signals of potential landslides, such as slope-angle mutations and curvature anomalies; and is applicable for large-scale automated potential landslide identification research. After obtaining the 10 m resolution DEM data, a CS stereoscopic map is constructed following the procedure described in Section 2.4, coupling slope and curvature information to form high-contrast terrain visualization images as the basic input for subsequent sample annotation and model training.

4.1.1. Potential Landslide Area Annotation

Potential landslide areas refers to slope units that, without yet exhibiting obvious sliding, already show typical landslide geomorphic features and have an instability tendency. Unlike occurred landslides, potential landslide areas are characterized by morphological readiness without yet having become unstable, and may become unstable under triggering factors.

(1) Multi-constraint annotation criteria: This paper constructs a “morphological–structural” dual-constraint annotation system, combining the landslide susceptibility assessment method proposed by Uchida et al. [34], systematically integrating rainfall-induced factors, geological conditions, topographic features, and other controlling elements. It should be noted that the core objective of this paper is solely to identify potential landslide areas based on topographic and geomorphic features; the next phase of research plans to incorporate key geological factors, such as rainfall, earthquakes, and soil water content, for further landslide susceptibility assessment of potential landslide areas.

Based on the above research context and the geomorphic characteristics of Southwestern China’s mountainous areas, a “morphological–structural” dual-constraint framework is established for potential landslide area identification. Morphological criteria require that the target area possesses at least one of three feature types: (i) ancient landslide remnants, including amphitheater/horseshoe-shaped depressions, landslide terraces, and frontal bulging; (ii) long-term gravitational deformation landforms, such as rock creep slopes, linear depressions, reverse scarps, and gentle ridge tops; and (iii) specific terrain combinations, i.e., slope angles in the 20–40° range, large catchment-area steep slope units, and convex slope-angle transitions. Structural criteria require that the target area satisfies at least one of three features: rear-edge tension trends identifiable through curvature anomaly zones and crack terrain; middle-zone sliding deformation manifested as surface micro-undulations and secondary sliding surfaces; and front-edge accumulation trends reflected in slope-toe bulging and gully blocking.

Since environmental criteria are outside the scope of this study, detailed elaboration is omitted here. In the specific annotation process, areas simultaneously satisfying both morphological and structural criteria are classified as potential landslide areas (positive samples); stable areas satisfying no constraint conditions are classified as non-landslide areas (negative samples); and other uncertain areas that cannot be clearly identified are not included as negative samples for model training.

(2) Sample annotation and quality control: The selected Wushan–Fengjie–Yunyang area covers approximately 1800 km^2. After vegetation removal from the raw terrain data, a

CS stereoscopic map was constructed in accordance with the method proposed in Reference [33], and each area is assessed according to the above criteria. After combining expert experience to eliminate samples with ambiguous boundaries and misclassifications, a total of 811 independent landslide outlines are extracted as positive samples (potential landslide areas), as shown in Figure 8. In stable areas satisfying neither criterion, 120 negative sample areas are extracted through grid sampling, forming the original sample library.

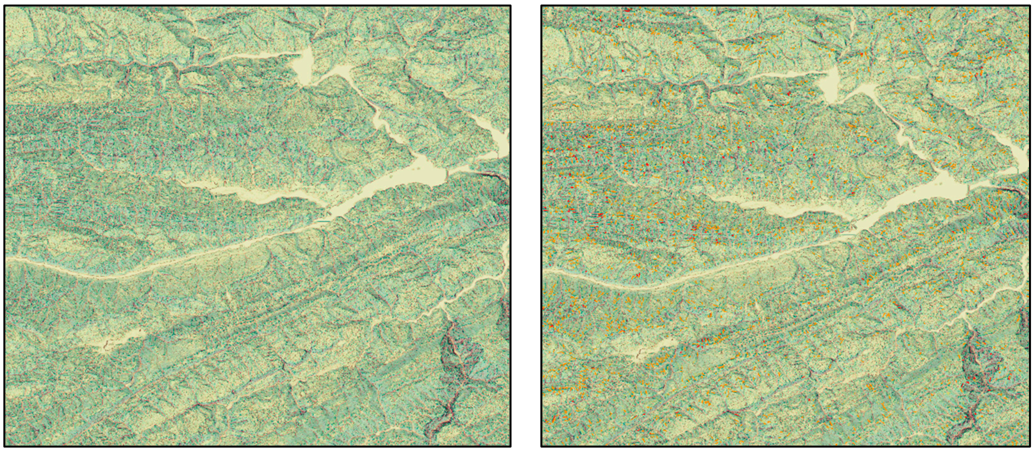

Figure 8. Annotation of potential landslide areas. (**Left**): Original CS image. (**Right**): Artificially marked potential landslide areas.

4.1.2. Sample Data Preprocessing

(1) Positive samples:

Local cropping is performed on CS maps with marked potential landslide areas from the above annotation, containing landslide feature textures and internal textures.

(2) Negative samples:

In the above annotation map, negative samples come from stable slopes, river valley flatlands, or other non-landslide areas, avoiding confusion with landslide morphology. All samples are uniformly cropped to 512 $\times$ 512 pixel image patches (size selection rationale—this size can simultaneously accommodate complete landslide body outlines and necessary terrain contextual backgrounds, and is compatible with the pix2pix generative network input specifications).

After augmentation with 35 data enhancement strategies, 5000 positive samples (with pixel-level mask images) and 10,000 negative samples (with blank mask images) are obtained (Figure 9). The 35 enhancement strategies are divided into four categories by function as geometric transformations (16—angle rotation, horizontal/vertical/bidirectional flipping, random cropping and scaling), photometric transformations (five—level brightness adjustment), texture enhancements (sharpening with three intensity levels, Gaussian noise with four variance levels), and combined enhancements (seven—combinations of the above strategies), forming a total of 15,000 training pairs. All 5000 positive samples are equipped with corresponding binary mask annotations (white for potential landslide areas, black for stable areas), and 10,000 negative samples are equipped with all-black masks (blank labels indicating absence of landslide), forming a weakly supervised training setup in which positive samples provide pixel-level mask supervision and negative samples provide weak image-level supervision.

Dataset splits: The 15,000 augmented image–mask pairs generated above were randomly split into training, validation, and test sets at the augmented-pair level (i.e., after augmentation, not at the level of the original 811 positive and 120 negative source samples). Consequently, different augmented variants (e.g., different rotations or crops) derived from the same original landslide outline or the same original negative-sample region could be assigned to different subsets. Because most of the applied augmentations (rotation, flipping, brightness adjustment, minor cropping/scaling) preserve the core landslide morphology and texture statistics of the source patch with only limited spatial perturbation, this split strategy is likely to be optimistic and may inflate the reported test set's performance relative to a strictly sample-disjoint split. This is acknowledged explicitly as a limitation of the present evaluation protocol in Section 5, and a source-disjoint re-split (splitting at the level of the original 811 + 120 samples before augmentation) is planned as a necessary validation step in future work.

(1) Training set: 80%—4000 positive + 8000 negative = 12,000 pairs;
(2) Validation set: 10%—500 positive + 1000 negative = 1500 pairs;
(3) Test set: 10%—500 positive + 1000 negative = 1500 pairs.

4.2. Model Training and Validation

4.2.1. Algorithm Flow

The algorithm is shown in Algorithm 1, and the workflow is illustrated in Figure 10.

Algorithm 1: DGARL-LS for Potential Landslide Recognition

Input: CS images with labeled set {(I_l,y_l)} (5000 positive pairs with pixel-level masks) and negative set {I_u} (10,000 stable region patches).
Output: Detected and segmented potential landslide regions.
1. Initialize experience replay buffers for forward and backward exploration.
2. Initialize the DRL agent with policy and value networks.
3. Initialize the generator (G) and discriminators Ds (segmentation) and Dd (region).
4. For each training epoch **do:**
  4.1 For each input image (I) **do:**
    (a) Forward exploration:
Start from initial central window.
At each step, the agent selects an action to update the window.
The generator produces a segmentation mask from the cropped region.
The discriminators evaluate region consistency and mask quality.
The reward is computed from discriminator scores and stored with the transition in the forward buffer.
    (b) Backward exploration (for labeled data):
Starting from the final state, sample inverse actions to trace back possible trajectories.
Store the transitions and rewards in the backward buffer.
    (c) GAN update:
Update the generator and discriminators using Tversky + BCE loss for labeled data, and adversarial BCE loss for unlabeled data.
    (d) Policy update:
Sample a minibatch of experiences from both buffers with mixing ratio.
Update the policy and value networks using the SAC objective with entropy regularization.
  4.2 End For
5. End For
6. Return the trained DRL policy and generator for potential landslide detection and segmentation.

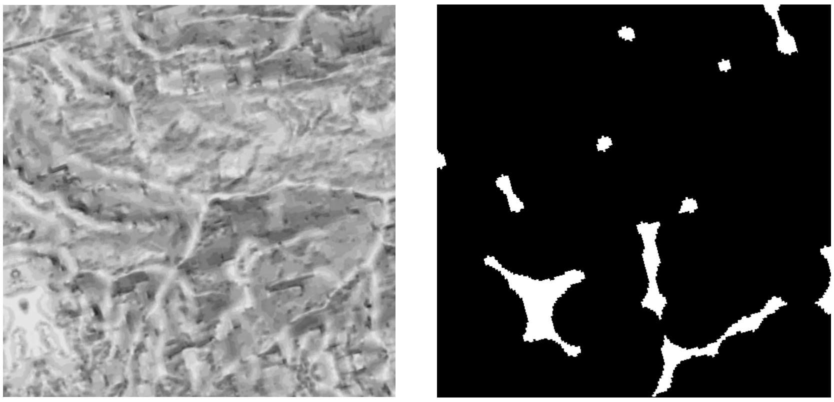

Figure 9. Data sample examples. (**Left**): CS image (grayscale). (**Right**): Mask annotation.

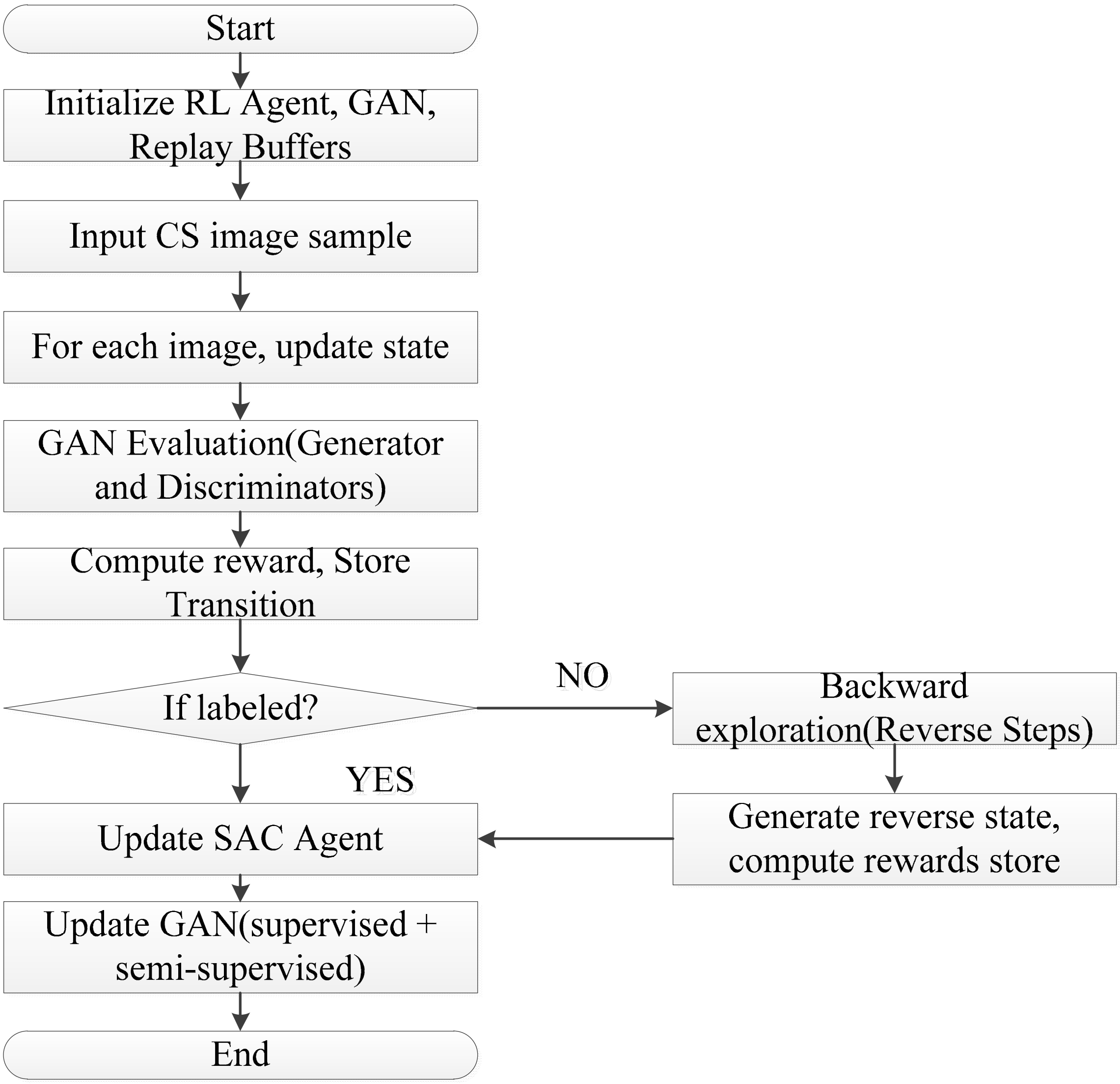

Figure 10. Algorithm workflow.

4.2.2. Parameter Settings

In the potential landslide identification experiments, all hyperparameters are validated and fine-tuned on the training set to ensure optimal model performance under complex terrain data. The hyperparameters of DGARL-LS are divided into basic and key categories (Table 2).

Table 2. Basic hyperparameter settings.

Parameter	Value/Setting
DRL initial learning rate	1×10^{-4}
GAN generator learning rate	1×10^{-4}
GAN discriminator learning rate	1×10^{-5}
Learning rate decay coefficient	0.95
Input image size	512×512
RL method	Soft actor–critic (SAC)
Backward exploration step length	30
Backward experience usage ratio	50%
Discount factor (gamma)	0.99
Batch size	32

(1) Basic hyperparameters: The initial learning rate for both the DRL and GAN generator is set to 1×10^{-4}, and the GAN discriminator learning rate is set to 1×10^{-5}, with gradual decay according to a coefficient of 0.95.

(2) Key hyperparameters: SAC (soft actor–critic) is selected as the core DRL method; its entropy regularization property helps avoid premature convergence to suboptimal solutions in complex environments. The backward exploration step length is set to 30 to ensure that exploration in landslide texture regions covers sufficient terrain context without adding excessive redundant information. The backward exploration experience usage ratio is set to 50%, achieving balanced utilization of forward and backward exploration experience, which helps alleviate training instability under sparse landslide sample conditions.

(3) Experimental environment: All input CS images are uniformly processed as 512×512 pixel sliced images. Experiments are implemented based on PyTorch 1.8.0. Data are generated from CS images created from DEM data, covering a terrain map area of approximately 1800 km^2 within the sample library, which are annotated using the dual-constraint annotation system and then sliced and augmented to generate training data. Through the above hyperparameter and environment settings, the DGARL-LS model can stably converge in CS maps of complex terrain environments, realizing robust interpretation of potential landslide areas.

4.3. Experimental Results

To verify the effectiveness and superiority of the proposed deep generative adversarial reinforcement learning model (DGARL-LS) in the task of identifying potential landslide areas, this paper selects four deep learning segmentation models as comparative baselines: U-Net, pix2pix, Trans U-Net, and Swin U-Net. Among them, Trans U-Net combines convolutional feature extraction with Transformer encoders, achieving significant results in the field of medical image segmentation [35,36]; conversely, Swin U-Net adopts a pure Transformer U-shaped encoding and decoding structure, models long-distance spatial dependence with a sliding window attention mechanism, and has theoretical advantages in global perception of terrain texture. All baseline models are trained on the same dataset with identical hyperparameter strategies to ensure fair comparison. Datasets are divided as follows: 80% training set (4000 positive + 8000 negative = 12,000 pairs), 10% validation set (1500 pairs), and 10% test set (1500 pairs).

4.3.1. Comparative Experiments

To validate the effectiveness and superiority of the proposed DGARL-LS model in potential landslide area identification, four deep learning segmentation models are selected as comparison baselines: U-Net, pix2pix, Trans U-Net, and Swin U-Net. All models are trained with the previously-mentioned training, validation and test datasets with identical split ratios to guarantee fair comparisons.

(1) Experimental design: Input resolution is uniformly 512 × 512; training epochs and batch sizes are kept consistent across all models—training epochs are set to 200, while batch sizes are set to 32. All baseline models (U-Net, pix2pix, Trans U-Net, Swin U-Net) are trained using the Adam optimizer with an initial learning rate of 2×10^{-4}, and then Cosine Annealing LR is applied after 100 epochs, following their original settings. DGARL-LS uses separate learning rates for its components, as specified in Table 2 (GAN generator and DRL: 1×10^{-4}; GAN discriminator: 1×10^{-5}), with the same Cosine Annealing schedule applied after 100 epochs. U-Net and pix2pix both adopt a combined Binary Cross-Entropy (BCE) and Dice loss, while the proposed DGARL-LS additionally introduces adversarial loss and RL reward functions to enhance dynamic feedback and structural consistency.

(2) Evaluation metrics: Intersection over Union (IoU), Dice coefficient, precision, and Recall are used as evaluation metrics; training convergence speed and model stability under complex terrain conditions are also compared.

(3) Experimental results: The results are shown in Table 3 and Figure 11.

Table 3. Comparative experimental results.

Method	IoU (%)	Dice (%)	Precision (%)	Recall (%)	Epoch
U-Net	72.4	78.6	81.0	75.2	160
Pix2pix	74.8	80.9	83.1	77.3	150
Trans U-Net	74.1	80.3	82.8	76.5	145
Swin U-Net	75.3	81.7	84.0	78.1	138
Proposed DGARL-LS	76.8	83.5	85.6	80.2	120

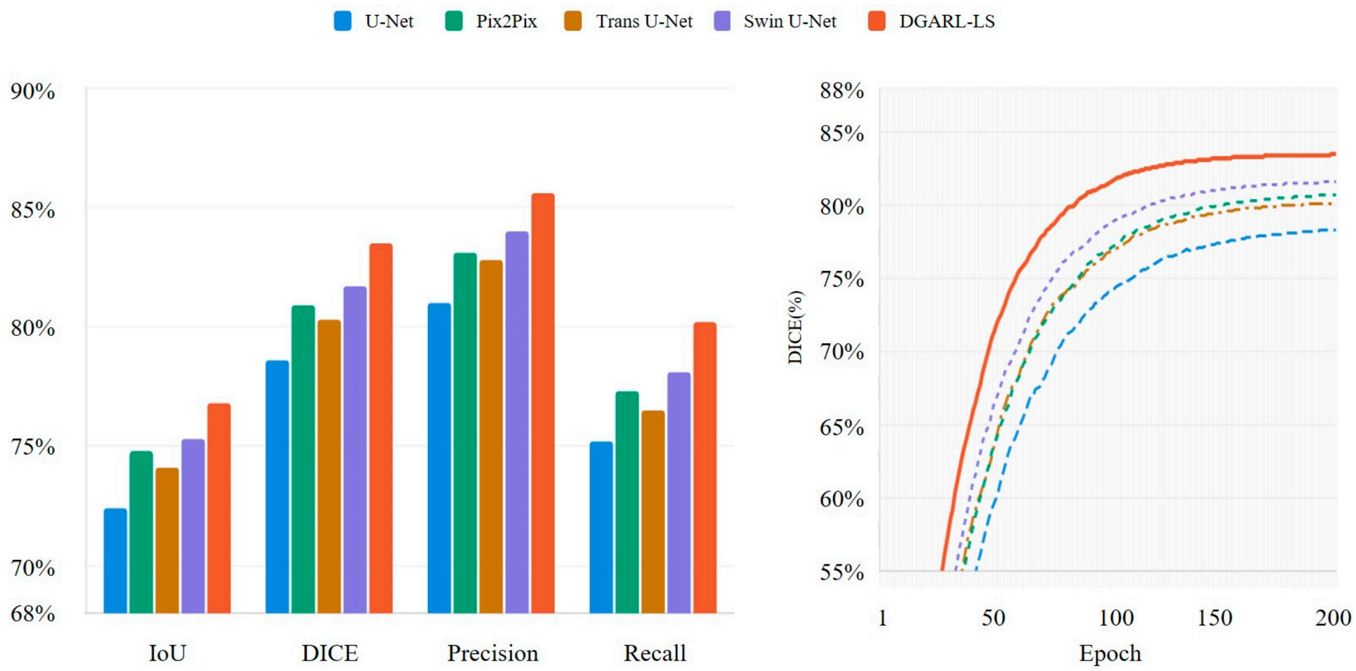

Figure 11. Performance comparison of different models in potential landslide identification task. DGARL-LS outperforms U-Net, pix2pix, Trans U-Net, and Swin U-Net in all four metrics of IoU, Dice, precision, and Recall. The improvements in Recall and Dice are particularly significant, indicating that this method excels in texture recognition and completeness preservation for small-scale and low-contrast landslide areas. (**Left**): Performance comparison of various indicators. (**Right**): Dice convergence curve on the validation set.

(4) Results analysis: As shown in the table, the proposed DGARL-LS model outperforms all comparison models across all metrics. Compared to convolutional baselines, IoU and Dice are improved by approximately 4.4% and 4.9%, respectively, compared to U-Net, and by approximately 2.0% and 2.6% compared to pix2pix, validating the effectiveness of introducing bidirectional exploratory reinforcement learning and stage-coupled generation of discriminative structures.

Compared to the Transformer series baselines, DGARL-LS also maintains its advantage, as Trans U-Net is slightly lower than pix2pix (74.8%) in IoU (74.1%), which is related to its limited transfer capability on small-sample terrain datasets due to its reliance on large-scale pretrained weights; conversely, Swin U-Net, with its sliding window attention mechanism for global modeling of spatial structure, outperforms Trans U-Net and U-Net in IoU (75.3%) and Dice (81.7%), but is still lower than DGARL-LS. This indicates that although the Transformer architecture possesses stronger global awareness, it still struggles to compensate for its insufficient local response to small-scale, low-contrast terrain textures in this task. DGARL-LS, through its reinforcement learning-based dynamic window search mechanism, can adaptively focus on landslide candidate regions, effectively overcoming this limitation.

Furthermore, DGARL-LS's convergence speed (by the 120th epoch) is significantly faster than all comparative models, further validating the role of the collaborative feedback mechanism in accelerating model training.

4.3.2. Ablation Experiments

(1) Experimental design: To ensure comparability with comparative experiment results, the ablation experiments use exactly the same dataset splits and training settings—training set 80% (12,000 pairs), validation set 10% (1500 pairs), test set 10% (1500 pairs), input resolution 512 $\times$ 512, batch size 32, 200 training epochs, and Cosine Annealing LR decay. Ablation configurations are shown in Table 4.

Table 4. Ablation experiment configurations.

No.	Configuration Name	Module Adjustment	Specific Content
I	W/o Backward	Remove reverse exploration	Only forward exploration is retained; backward trajectory backtracking and prior experience generation are disabled.
II	W/o Forward	Remove forward exploration	Retaining only backward exploration, the agent starts from a known high-confidence region and initializes the search without a central point.
III	W/o Bidirectional	Remove all RL exploration	It degenerates into static pix2pix splitting with no dynamic window search mechanism.
IV	W/o Stage-GAN	Remove the stage coupling discriminator	Retains the single segment discriminator and remove the region discriminator Dd.
V	W/o Feedback	Remove collaborative feedback mechanism	Using a fixed constant reward instead of discriminator scores, GAN and RL are trained independently.
VI	DGARL-LS (full)	Full modules	The complete model in this paper.

(2) Experimental results: Ablation results are shown in Table 5 and Figure 12. The contributions of each module are analyzed as follows.

Contribution of the bidirectional exploration mechanism (I vs. II vs. III): When only backward exploration is removed (configuration I), Recall decreases from 80.2% to 75.0%, and IoU decreases by 3.5 percentage points, indicating that backward trajectory backtracking is crucial for capturing low-contrast and small-scale potential landslide areas—in reward-sparse scenarios, the lack of prior experience makes it difficult for the agent to effectively lock onto the target area. When only forward exploration is removed (configuration II), all indicators further decrease, with IoU dropping to 72.1%, indicating that the progressive search starting from the center of the CS graph is the foundation for the agent to establish initial terrain perception; without a center initialization, the prior experience of backward exploration cannot be effectively anchored, and the exploration efficiency decreases significantly. Completely removing bidirectional RL exploration (configuration III) degenerates the model into a static pix2pix model, with performance consistent with the pix2pix results in the comparative experiment (IoU 74.8%), verifying the independent gain of the bidirectional exploration mechanism relative to pure GAN segmentation from the perspective of control variables.

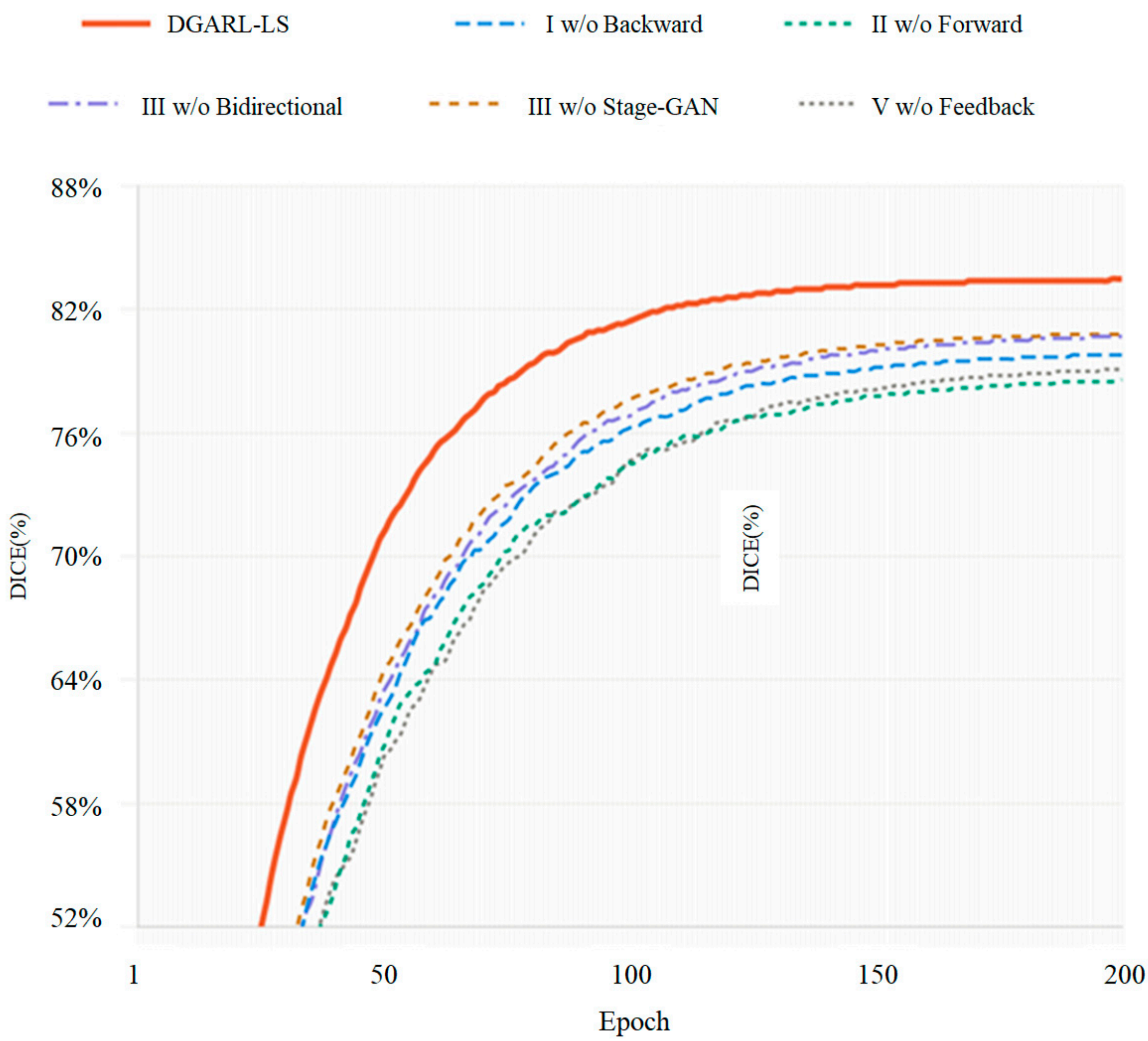

Figure 12. Comparison of validation set Dice curves for different ablation configurations of the DGARL-LS model.

Contribution of the stage-coupled discriminator (IV): After removing the region discriminator Dd, the model's constraint on the rationality of candidate window regions disappeared, leading to increased edge blurring and background misclassification in the segmentation results. Precision decreased from 85.6% to 83.4%, validating the necessity of the dual-discriminator structure in jointly optimizing detection regions and segmentation accuracy.

Table 5. Ablation experiment results.

No.	IoU (%)	Dice (%)	Precision (%)	Recall (%)	Epoch
I	73.3	80.0	82.5	75.0	148
II	72.1	78.8	81.3	73.6	155
III	74.8	80.9	83.1	77.3	150
IV	74.5	81	83.4	77.2	143
V	72.8	79.4	81.9	74.1	162
VI(DGARL-LS)	76.8	83.5	85.6	80.2	120

Contribution of the collaborative feedback mechanism (V): After removing collaborative feedback and replacing the dynamic discriminator score with a fixed reward, training stability significantly decreased, the number of convergence epochs increased to 162 (35% slower than the complete model), and IoU and Recall were among the lowest of all configurations. This indicates that the dynamic reward signal is the core link for effective collaboration between GAN and RL—without real-time feedback from the discriminator score, RL cannot perceive the current segmentation quality, and policy optimization falls into blind exploration.

In summary, the three modules (bidirectional exploration mechanism, stage-coupled discriminator, and collaborative feedback mechanism) each have their own functions and are interdependent in the DGARL-LS framework: bidirectional exploration is responsible for the dynamic localization of candidate regions, stage-coupled GAN is responsible for fine segmentation and quality evaluation, and collaborative feedback mechanism realizes the information loop between the two. The absence of any one of the three will lead to a significant degradation in recognition accuracy and training stability.

4.4. Cross-Regional Validation

4.4.1. Purposes of Cross-Regional Experiment

To examine whether the proposed method retains its relative advantage when applied to a geological and topographic setting different from the primary study area, an independent cross-regional validation experiment was conducted in the Dazhou area of Sichuan Province (107.20° E–108.80° E, 31.20° N–32.80° N), located outside the Wushan–Fengjie–Yunyang section of the Three Gorges Reservoir area used for model development in Sections 4.1–4.3. Dazhou was selected because it lies within the broader Sichuan–Yunnan mountainous belt but is governed by a distinct structural and lithological setting from the original Three Gorges study area, providing a meaningful test of whether the terrain–morphology relationships learned by DGARL-LS transfer beyond the specific tectonic unit on which it was trained.

4.4.2. Experimental Design

A 10 m resolution DEM covering the Dazhou study area was acquired and processed into CS image (Figure 13) following the identical workflow described in Section 2.4. Positive and negative samples were independently annotated for this region using the same morphological–structural dual-constraint criteria described in Section 4.1.1 (with reference to Uchida et al. [34]). The Dazhou dataset comprises 99 valid CS tiles in total: 49 positive tiles and 50 negative tiles. This positive-tile count is substantially smaller than the number of original landslide outlines identified in the Three Gorges area (811 positive samples)—corresponding to roughly 6.0% of the Three Gorges original positive sample count, and only about 1.0% of the 5000 augmented positive pairs used to train the primary-area models. Because of this pronounced data scarcity, this experiment deliberately did not apply the data augmentation strategy; all five models were instead trained directly on the original, unaugmented Dazhou tiles. This design choice serves two purposes: first, it avoids confounding the cross-regional comparison with the data-leakage caveat already noted for the augmented-pair-level split described in Section 4.1.2; and second, since all five comparison models (DGARL-LS, U-Net, pix2pix, Trans U-Net, and Swin U-Net) were retrained from random initialization under this identical, deliberately data-scarce condition—with the same 80%/10%/10% training/validation/test split ratio used in Section 4.1.2—the comparison remains fully controlled, even though it is not informative about zero-shot transfer of the original Three-Gorges-trained weights. We therefore frame this experiment as a test of methodological transferability (i.e., whether the DGARL-LS architecture retains its relative advantage when retrained under realistic data scarce conditions in a new region) rather than a test of direct cross-regional generalization without retraining; the latter remains an open question for future work, as discussed in Section 5.

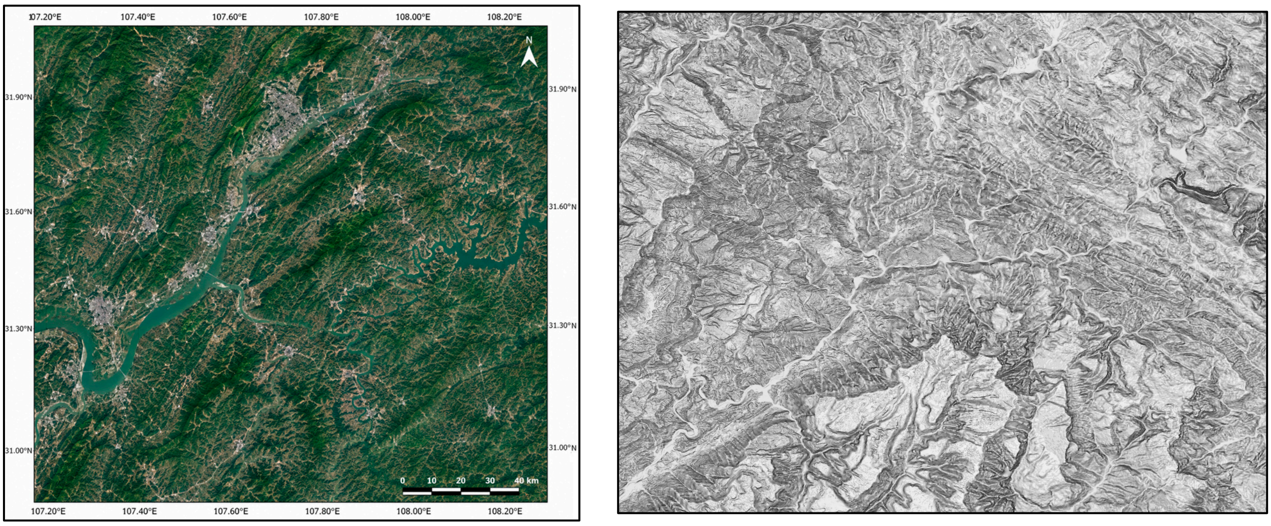

Figure 13. CS images covering Dazhou area. (**Left**): Satellite remote sensing image. (**Right**): CS image (grayscale).

4.4.3. Results

Cross-regional validation results on the Dazhou test set are listed in Table 6.

Table 6. Cross-regional validation results.

Method	IoU (%)	Dice (%)	Precision (%)	Recall (%)	Epoch
U-Net	58.51	73.83	78.57	69.62	24
Pix2pix	59.80	74.85	76.25	73.49	20
Trans U-Net	35.64	52.55	50.70	54.55	17
Swin U-Net	20	33.33	33.90	32.79	1
DGARL-LS	70.83	82.93	87.18	79.07	24

As shown in Table 6, DGARL-LS achieves the highest precision (87.18%), Recall (79.07%), Dice (82.93%), and IoU (70.83%) among all five models on the Dazhou test set, outperforming the second-best model pix2pix by 11.03, 5.58, 8.08, and 11.03 percentage points, respectively. DGARL-LS converges at 24 epochs, while pix2pix converges faster at 20 epochs. It is worth noting that Swin U-Net terminated abnormally after only one training epoch, and both Swin U-Net and Trans U-Net exhibit obvious limitations when trained on small-sample datasets. This indicates that the relative advantage of the bidirectional exploration mechanism and stage-confluent GAN observed in the primary study area is preserved when the method is retrained under a new geological and topographic setting.

At the same time, all five models show a clear decline relative to their performance on the primary Three Gorges test set (Table 3); for DGARL-LS specifically, precision increases slightly from 85.6% to 87.18% (+1.58 points), Recall decreases from 80.2% to 79.07% (−0.93 points), Dice decreases from 83.5% to 82.93% (−0.57 points), and IoU decreases from 76.8% to 70.83% (−5.97 points). The performance degradation can be attributed to the inherent discrepancies in geological structure, lithology and erosion characteristics between the Three Gorges Reservoir area and Dazhou. Such regional differences weaken the transferability of terrain–morphology correlation rules learned by the model. Notably, precision slightly rises rather than dropping, indicating that the model's positive prediction results stay relatively credible under cross-regional conditions. By contrast, Recall, Dice and IoU suffer moderate performance losses, which reveals that the recognition of full landslide coverage and accurate boundary segmentation rather than false positive suppression, become the main challenges when migrating the model to geologically distinct regions.

It should be emphasized that this experiment retrained all models on Dazhou data rather than applying the Three-Gorges-trained weights directly (zero-shot transfer). Notably, Swin U-Net aborted after merely one training epoch, and both Swin U-Net and Trans U-Net suffer severe performance limitations under small-sample conditions. Therefore, the results only prove that the DGARL-LS architecture, especially its bidirectional exploration and cooperative feedback design, maintains stable superiority over U-Net, pix2pix, Trans U-Net, and Swin U-Net in new geologic regions with limited data rather than verifying that a pretrained model can realize cross-regional generalization without retraining.

The cross-regional validation experiment reveals an obvious limitation: the model's ability to accurately segment landslide boundaries across regions degrades far more severely than its capacity to distinguish landslide categories. Such performance loss mainly stems from the inherent geological differences in lithology, structure, and erosion patterns between the two study areas, which weaken the transferability of terrain-morphology features learned by the model.

In subsequent research on higher-resolution, fine-grained terrain interpretation, the negative impact brought about by cross-regional geological domain shift will be alleviated via a multi-region sample fusion training strategy. Instead of training the model using data from only one single area, CS stereoscopic samples covering geologically diverse zones including the Three Gorges Reservoir, Dazhou, and other mountainous areas in the Sichuan–Yunnan region will be integrated into a unified training set. This strategy enables the model to learn abundant terrain–morphology rules controlled by various lithology and structural conditions during training, instead of transferring knowledge learned from a single region to unfamiliar geological environments. Accordingly, the cross-regional generalization ability and landslide boundary segmentation precision of the model can be simultaneously improved.

5. Conclusions

This paper addresses the engineering challenge of large-scale automatic identification of potential landslides in complex mountainous areas. It proposes a terrain interpretation method based on deep generative adversarial reinforcement learning (DGARL-LS), using 10 m resolution DEMs acquired by Earth observation satellites as the sole data source. By constructing slope–curvature (CS) stereo maps, it captures prior geometric signals during the landslide incubation stage. The method integrates pix2pix conditional generative adversarial networks with bidirectional exploratory deep reinforcement learning into a closed-loop framework of "regional localization–fine-grained segmentation–feedback optimization." The main conclusions are as follows.

First, DEMs, as a core derivative product of Earth observation methods, possess significant advantages. such as being unaffected by vegetation cover and illumination conditions, being globally available and freely accessible, and having derived terrain factors directly reflecting the slope's mechanical response. Experiments demonstrate that CS stereo maps constructed based on DEMs, by simultaneously encoding complementary terrain information of slope and curvature, significantly outperform single elevation input in representing potential landslide areas anomalies, providing a high-contrast, learnable spatial feature representation for subsequent deep learning models.

Secondly, regarding model performance, DGARL-LS achieved IoU, Dice, precision, and Recall of 76.8%, 83.5%, 85.6%, and 80.2%, respectively, on a test set covering an area of 1800 km^2 in the Wushan–Fengjie–Yunyang region of Chongqing. These figures outperformed the four baseline models: U-Net, pix2pix, Trans U-Net, and Swin U-Net. Specifically, compared to Swin U-Net (IoU 75.3%), which possesses global long-range modeling capabilities, DGARL-LS adaptively focuses on landslide candidate areas through a dynamic window search mechanism using reinforcement learning, effectively compensating for the Transformer architecture's insufficient response to small-scale, low-contrast terrain textures. Its convergence speed is approximately 25% faster than U-Net, validating the contribution of the collaborative feedback mechanism to training stability and convergence acceleration.

Third, the ablation experiment system verified the independent contributions of the three core modules, as Recall decreased by 5.2 percentage points and IoU decreased by 3.5 percentage points after removing backward exploration, indicating that the trajectory backtracking mechanism is crucial for capturing small-scale targets in reward-sparse scenarios. All indicators further decreased after removing forward exploration (IoU dropping to 72.1%), verifying that center initialization search is the foundation for the agent to establish terrain perception. After completely removing bidirectional RL and degenerating to static pix2pix, the performance was consistent with the pix2pix baseline (IoU 74.8%), confirming the independent gain of the bidirectional exploration mechanism from the perspective of control variables; after removing collaborative feedback, the number of convergence epochs increased to 162 (35% slower than the full model), quantitatively proving the role of dynamic reward signals as a collaborative link between GAN and RL.

Fourth, cross-regional validation on the Dazhou dataset further demonstrates the robust generalization superiority of the DGARL-LS architecture under data-scarce and heterogeneous geological conditions. Retrained from scratch on the unenhanced Dazhou dataset with distinct lithology, structural settings, and erosion regimes from the Three Gorges area, DGARL-LS still achieves optimal IoU (70.83%), Dice (82.93%), precision (87.18%), and Recall (79.07%) values among all comparative models, surpassing the second-best pix2pix model by substantial margins. In contrast, Trans U-Net and Swin U-Net exhibit severe performance degradation under small-sample cross-regional conditions, where Swin U-Net even fails to complete full training and terminates after only one epoch. The performance decline of all models in cross-regional tests is mainly attributed to inherent regional geological discrepancies, which weaken the transferability of terrain–morphology correlation rules learned from the Three Gorges dataset. In particular, the model's boundary segmentation accuracy is more sensitive to cross-regional geological domain shifts than overall category discrimination capability, revealing the key limitation of fine-grained landslide boundary interpretation across different geological units.

This study has the following limitations: First, the current framework only uses terrain geometric features as the basis for interpretation and has not yet incorporated triggering factors, such as rainfall, earthquakes, and soil saturation, resulting in the identification conclusion of "potential unstable areas at the terrain level" rather than "predictions of impending landslides". Second, while the cross-regional validation experiment in Dazhou

(Section 4.4) shows that DGARL-LS retains its relative advantage over the four baseline models after being retrained on a new, data-scarce region, this experiment retrained the model from scratch rather than testing zero-shot transfer of the Three-Gorges-trained weights, and used an unaugmented, substantially smaller training set; direct zero-shot cross-regional generalization, together with validation across a wider range of tectonic units using fully augmented region-specific datasets, therefore still needs further verification. Third, when the terrain background is highly complex or the slope texture is extremely blurry, the model still has certain false positives and false negatives. Fourth, the training/validation/test split was performed at the level of augmented image–mask pairs rather than at the level of the original source samples, so different augmented variants of the same original landslide outline or negative-sample region could fall into different subsets. Fifth, the cross-regional decline observed in Section 4.4—particularly the disproportionately large drop in IoU—reflects a compounding effect between training-data scarcity in the new region and genuine differences in lithology, structural setting, and terrain texture relative to the Three Gorges area, and the present experimental design cannot disentangle the relative contribution of these two factors.

Future research will proceed in the following directions: first, multimodal fusion of triggering factors such as rainfall intensity, peak ground acceleration, and soil moisture content with CS topographic features will be implemented to upgrade from "potential area identification" to "dynamic prediction of occurrence probability"; second, multi-temporal DEM differential information will be introduced to construct a spatiotemporal joint interpretation framework in combination with temporal topographic evolution features; third, the method will be extended to higher resolution (1–5 m) airborne LiDAR DEM data to achieve refined interpretation at the individual slope level, further supporting engineering risk management of highway slopes and reservoir bank slopes; fourth, to address the compounding effect between data scarcity and cross-regional domain shift identified in Section 4.4, a multi-region sample fusion training strategy will be adopted in subsequent fine-grained terrain interpretation research—pooling CS stereoscopic samples from multiple geologically distinct regions (e.g., the Three Gorges Reservoir area, Dazhou, and additional Sichuan–Yunnan mountainous sites) into a unified training set rather than training on a single region before testing elsewhere—so that the model can learn a wider range of lithology- and structure-dependent terrain–morphology associations during training itself, thereby improving both cross-regional generalization and boundary-level accuracy; and fifth, a systematic sensitivity analysis of the layer-weighting (blending opacity) and color-mapping scheme used to construct the CS stereoscopic map will be conducted.

Author Contributions: X.-N.L. conceived the overall research framework, developed the DGARL-LS methodology, implemented the software, conducted the investigation, prepared the original draft, and was responsible for visualization. B.-C.Y. contributed to the data curation, sample investigation, and visualization. J.-B.H. performed the validation and formal analysis and secured the funding. B.-Q.L. contributed to the conceptualization, methodology refinement, resource provision, and manuscript review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Key Research and Development Program of Yunnan Province Science and Technology Department [grant number: 202503AA080030].

Data Availability Statement: The DEM data used in this study were obtained from a third-party commercial provider and are not publicly available due to licensing restrictions. The annotated CS stereoscopic map sample dataset constructed for the Three Gorges Reservoir area, including the 15,000 training pairs generated through data augmentation, is also not publicly available at this time due to ongoing related research. Researchers who wish to access the derived data or methodology details for non-commercial academic purposes may contact the corresponding author upon reasonable request, and to support the reproducibility of the proposed method, the core code can also be provided to researchers upon reasonable request by contacting the corresponding author.

Acknowledgments: The authors would like to thank the data provider for supplying the DEM data used in this study. During the preparation of this manuscript, the authors used AI for the purposes of English-language refinement and grammatical revision. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

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

References

1. Liu, G. Landslide disasters and prevention along railway lines in China. *Roadbed Eng.* **1989**, *8*, 11–18.
2. Tian, Y.; Zhang, J.; Zhao, W.; Cheng, C.; Deng, Y.; Ma, R.; Huang, R. Displacement prediction for newly monitored slopes based on a comprehensive dataset with lithological classification. *Peking Univ. J. Platf.* **2025**, *4*, 709–718. [CrossRef]
3. Cao, P.; Li, Y.; Li, Z.; Zhang, S. Geological Structure Characteristics and Genetic Mechanism of Baige Landslide Slope in Changdu, Tibet. *Earth Sci.* **2021**, *46*, 3397–3409. [CrossRef]
4. Yin, Z.; Cheng, G.; Li, X.; Wang, L.; Hu, G.; Wei, G. Relationships Between Landslide Forming Mechanisms and Climate Changes Since the Early-Middle Pleistocene in the Upper Reaches of the Yellow River and the Three Gorges Reservoir Area. *Quat. Sci.* **2010**, *30*, 37–45.
5. Yang, C.; Lin, G.; Zhang, M.; Zhang, R.; Sun, X. Soil Landslide Susceptibility Assessment Based on DEM. *J. Geo-Inf. Sci.* **2016**, *12*, 1624–1633. [CrossRef]
6. Qiu, Z.; Guo, C.; Zhang, Y.; Yang, Z.; Wu, R. Spatial structural characteristics of the Deda ancient landslide in the eastern Tibetan Plateau: Insights from Audio-frequency Magnetotellurics and the Microtremor Survey Method. *China Geol.* **2024**, *7*, 188–202. [CrossRef]
7. Ehsan, M.; Anees, M.T.; Bakar, A.F.B.A.; Ahmed, A. A review of geological and triggering factors influencing landslide susceptibility: Artificial intelligence-based trends in mapping and prediction. *Int. J. Environ. Sci. Technol.* **2025**, *22*, 17347–17382. [CrossRef]
8. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Alsdorf, D. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. [CrossRef]
9. Tadono, T.; Ishida, H.; Oda, F.; Naito, S.; Minakawa, K.; Iwamoto, H. Precise Global DEM Generation by ALOS PRISM. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *II-4*, 71–76. [CrossRef]
10. Xu, Q.; Dong, X.; Li, W. Integrated Space-Air-Ground Early Detection, Monitoring and Warning System for Potential Catastrophic Geohazards. *Geomat. Inf. Sci. Wuhan Univ.* **2019**, *44*, 957–966. [CrossRef]
11. Guzzetti, F.; Carrara, A.; Cardinali, M.; Reichenbach, P. Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* **1999**, *31*, 181–216. [CrossRef]
12. Bhuyan, K.; Rana, K.; Ozturk, U.; Nava, L.; Rosi, A.; Meena, S.R.; Fan, X.; Floris, M.; van Westen, C.; Catani, F. Towards automatic delineation of landslide source and runout. *Eng. Geol.* **2025**, *245*, 107866. [CrossRef]
13. Li, W.; Zhan, W.; Lu, H.; Xu, Q.; Pei, X.; Wang, D.; Ge, D. Precursors to large rockslides visible on optical remote-sensing images and their implications for landslide early detection. *Landslides* **2022**, *20*, 1–12. [CrossRef]
14. Zou, F.; Leng, Y.; Tao, X.; He, S. Landslide hazard identification based on slope unit: A case study of shallow soil slope in Wanshan, Guizhou Province. *Chin. J. Geol. Hazard Control* **2022**, *3*, 114–122. [CrossRef]
15. Bakhsoliani, D.; Magalashvili, A.; Gaprindashvili, G. Identification of Landslide Boundaries and Key Morphological Features Using UAV LiDAR Data: A Case Study in Surami, Georgia. *GeoHazards* **2025**, *6*, 73. [CrossRef]
16. Yang, X.; Chen, D.; Dong, Y.; Xue, Y.; Qin, K. Identification of potential landslide in Jianzha county based on InSAR and deep learning. *Sci. Rep.* **2024**, *14*, 21346. [CrossRef] [PubMed]
17. He, Y.; Zhao, Z.; Zhu, Q.; Liu, T.; Zhang, Q.; Yang, W.; Wang, Q. An integrated neural network method for landslide susceptibility assessment based on time-series InSAR deformation dynamic features. *Int. J. Digit. Earth* **2023**, *17*, 2295408. [CrossRef]
18. Li, N.; Feng, G.; Zhao, Y.; Xiong, Z.; He, L.; Wang, X.; Wang, W.; An, Q. A Deep-Learning-Based Algorithm for Landslide Detection over Wide Areas Using InSAR Images Considering Topographic Features. *Sensors* **2024**, *24*, 4583. [CrossRef] [PubMed]
19. Ma, S.; Li, S.; Bi, X.; Qiao, H.; Duan, Z.; Sun, Y.; Li, X. Automatic landslide identification by Dual Graph Convolutional Network and GoogLeNet model-a case study for Xinjiang province, China. *Front. Earth Sci.* **2023**, *11*, 1248340. [CrossRef]
20. Feng, L.; Zhang, M.; Mao, Y.; Liu, H.; Yang, C.; Dong, Y.; Nanehkaran, Y.A. Convolutional neural network-based deep learning for landslide susceptibility mapping in the Bakhtegan watershed. *Sci. Rep.* **2025**, *15*, 13250. [CrossRef] [PubMed]
21. Tong, Y.; Luo, H.; Qin, Z.; Xia, H.; Zhou, X. Enhanced Landslide Susceptibility Assessment in Western Sichuan Utilizing DCGAN-Generated Samples. *Land* **2024**, *14*, 34. [CrossRef]

22. Wang, B.K.; Su, J.L.; Xi, J.B.; Chen, Y.Y.; Cheng, H.Y.; Li, H.L.; Chen, C.; Shang, H.X.; Yang, Y. Landslide Detection with MSTA-YOLO in Remote Sensing Images. *Remote Sens.* **2025**, *17*, 2795. [CrossRef]
23. Bartsch, A. Deep learning for potential landslide identification: Data, models, applications, challenges, and opportunities. *Nat. Hazards Earth Syst. Sci.* **2026**, *26*, 487. [CrossRef]
24. Hirokazu, K. Extraction of Landslide Mass by Deep Learning Using 3D Geographical Information: Significance of Topographic Map Representation. *J. Jpn. Landslide Soc.* **2021**, *58*, 65–72. [CrossRef]
25. Feng, X.; Du, J.; Wu, M.; Chai, B.; Miao, F.; Wang, Y. Potential of synthetic images in landslide segmentation in data-poor scenario: A framework combining GAN and transformer models. *Landslides* **2024**, *21*, 2211–2226. [CrossRef]
26. Al-Najjar, H.A.H.; Pradhan, B. Spatial landslide susceptibility assessment using machine learning techniques assisted by additional data created with generative adversarial networks. *Geosci. Front.* **2021**, *12*, 625–637. [CrossRef]
27. Highland, L.M.; Bobrowsky, P. *The Landslide Handbook—A Guide to Understanding Landslides*; Circular; U.S. Geological Survey: Reston, VA, USA, 2008. [CrossRef]
28. Ma, S.; Qiu, H.; Zhu, Y.; Yang, D.; Tang, B.; Wang, D.; Cao, M. Topographic Changes, Surface Deformation and Movement Process before, during and after a Rotational Landslide. *Remote Sens.* **2023**, *15*, 662. [CrossRef]
29. Alkhasawneh, M.S.; Ngah, U.K.; Tay, L.T.; Mat Isa, N.A.; Al-batah, M.S. Determination of Important Topographic Factors for Landslide Mapping Analysis Using MLP Network. *Sci. World J.* **2013**, *12*, 415023. [CrossRef] [PubMed]
30. Deng, N.; Li, Y.; Ma, J.; Shahabi, H.; Hashim, M.; de Oliveira, G.; Chaeikar, S.S. A comparative study for landslide susceptibility assessment using machine learning algorithms based on grid unit and slope unit. *Front. Environ. Sci.* **2022**, *10*, 1009433. [CrossRef]
31. Xu, C.; Xu, X.; Yao, X.; Dai, F. Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis. *Landslides* **2013**, *11*, 441–461. [CrossRef]
32. Liu, M.; Xu, B.; Li, Z.; Mao, W.; Zhu, Y.; Hou, J.; Liu, W. Landslide Susceptibility Zoning in Yunnan Province Based on SBAS-InSAR Technology and a Random Forest Model a New Method for Assessing Deep Catastrophic Landslide Susceptibility. *Remote Sens.* **2023**, *15*, 2864. [CrossRef]
33. Toda, K. Topographic interpretation using curved-slope stereoscopic mapping (CS stereo maps). *For. Site* **2014**, *56*, 75–79.
34. Uchida, T.; Yokoyama, O.; Suzuki, R.; Tamura, K.; Ishizuka, T. A New Method for Assessing Deep Catastrophic Landslide Susceptibility. *Int. J. Eros. Control Eng.* **2011**, *4*, 32–42. [CrossRef]
35. Chen, J.; Mei, J.; Li, X.; Lu, Y.; Yu, Q.; Wei, Q.; Zhou, Y. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. *Med. Image Anal.* **2024**, *97*, 103280. [CrossRef] [PubMed]
36. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. In *Computer Vision—ECCV 2022 Workshops; ECCV 2022. Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2023; Volume 13803. [CrossRef]