Responses to the Second Reviewer's Comments

Pan Jiang, Zhengjing Ma*, Gang Mei**
School of Engineering and Technology
China University of Geosciences (Beijing)
gang.mei@cugb.edu.cn

Acknowledgement The authors would like to thank the editor and the reviewers for their comments.

Comment # 1:

1. The objective of this work is to review recent advances in the application of deep learning to landslide prediction and to highlight the challenges and opportunities in this field. The authors provide an extensive overview of existing model types, but the manuscript does not go into sufficient detail on the actual application of deep learning techniques to landslide prediction. I see potential in this review; however, it requires a thorough revision.

Response:

- Dear reviewer, we sincerely appreciate your recognition of the research objectives of this study and are deeply grateful for your valuable and insightful comments! We fully agree with your observation that the previous version of the manuscript did not adequately explore the practical application of deep learning techniques in landslide prediction. This observation prompted us to conduct a thorough reflection and a comprehensive reconstruction of the manuscript.
- Following your suggestion, we carried out a comprehensive and targeted revision of the manuscript. The central aim of this revision was to shift the focus from a general overview of model types to a detailed discussion of how these models are specifically applied to address practical problems in potential landslide identification.
- Specifically, we have made the following major revisions:
- (1) In **Section 2**, rather than merely introducing the principles and advantages or disadvantages of various data sources, we have incorporated numerous case studies demonstrating how these data sources are integrated with deep learning models (Please see **Comment #5**).
- (2) **Section 3** has been thoroughly reconstructed and substantially enriched; in each subsection, we have added detailed research cases and methodological descriptions. In accordance with your suggestion, we removed unnecessary descriptions of model architectures and instead focused on clarifying "which study employed which specific model architecture, addressed what type of landslide-related problem, utilized what kind of data, and achieved what key achievements. " (Please see **Comment #3**).
- (3) Similarly, in **Section 4**, we replaced general discussions with extensive examples illustrating

which types of landslides are best suited to specific deep learning models. (Please see **Section 4**).

Original Description in Section 4

4.1 Application of Deep Learning in the Identification of Rainfall-induced Landslides

……

Following the development of rainfall threshold models, real-time monitoring of historically rainfall-induced landslides is imperative. First, continuous surveillance enables early detection of subtle deformations and precursory anomalies (Guzzetti et al., 2020; Zhu et al., 2023b), facilitating timely reactivation warnings to mitigate secondary hazards to lives and infrastructure. Second, by continuously monitoring rainfall, soil moisture, and groundwater levels, we can support dynamic recalibration of threshold parameters. This data assimilation enhances model adaptability to evolving hydrogeological conditions, ensuring operational relevance across heterogeneous terrains.

……

4.3 Application of Deep Learning in the Identification of Human Activity-induced Landslides

……

Moreover, the triggers of human activity-induced landslides are not only related to natural conditions but also closely associated with dynamic human activities. Consequently, their analysis necessitates the integration of multimodal and cross-scale data to capture coupled environmental and behavioral drivers (see Fig. 6). In engineering operations such as mining or road construction, factors including proximity to potential landslide zones, excavation depth, and slope angles must be rigorously evaluated through geohazard risk assessments. During excavation phases, geotechnical investigations are imperative to identify weak lithological strata or fracture-dense zones predisposed to instability. Continuous slope stability monitoring requires deploying IoT-enabled sensors to track temporal variations in surface fissure dimensions and subsurface displacement vectors. Monitoring data from these sensors can be integrated into deep learning models for multimodal analytics, enabling dynamic risk prediction and adaptive mitigation planning.

To mitigate misclassification between anthropogenic signatures and natural terrain, integrating multispectral data with topographic elevation data enhances discriminative capacity (Meng et al., 2021; Selamat et al., 2023). For instance, in mountainous regions, DEMs revealing artificially excavated steep slopes combined with fractured geological strata from structural maps provide preliminary evidence of human influence on landslide susceptibility (Lian et al., 2024).

In fact, landslides triggered solely by human activities are relatively rare. Single human activities are typically insufficient to independently trigger landslides, with natural factors often acting in conjunction with human activities. Furthermore, the prohibitive cost of acquiring subsurface disturbance data results in sparse historical landslide samples for specific engineering scenarios, limiting data-driven model training.

4.4 Application of Deep Learning in the Identification of Human Activity-induced Landslides

Multi factor-induced landslides result from the synergistic interaction of multiple natural and anthropogenic factors (Hao et al., 2023). Their triggering mechanisms involve the dynamic spatiotemporal coupling of these factors, driving progressive destabilization of geomaterials through cumulative strength degradation. The formation of such landslides may involve various types of movements, including collapse, creep, and flow phenomena. They often exhibit characteristics such as complexity, nonlinearity, and suddenness. Therefore, their identification is markedly more complex compared to landslides triggered by singular factors.

Unlike simpler landslide types, identifying composite landslides necessitates multimodal data fusion to holistically assess predisposing conditions (Li, 2025; Yin et al., 2023). It further requires disentangling the nonlinear superposition effects of multiple factors and quantifying their relative contributions to failure initiation. For instance, Dou et al. (2019) analyzed how earthquake intensity and rainfall metrics jointly modulate landslide susceptibility, deriving failure probabilities under varying parameter combinations. In multi factor-induced landslides, earthquakes and rainfall often interact with other factors. For instance, during heavy rainfall, the rate of landslide formation after an earthquake may be higher, possibly driven by the removal of excessively steep slopes, changes in vegetation and groundwater, and alterations in the mechanical properties of the bedrock and weathered layers in the earthquake-induced landslides canopy. This necessitates systematic investigation of multi-hazard coupling effects to quantify emergent risks.

In addition to the approach of constructing physics-based models that account for multiple factors, GNNs can be employed. These models represent landslide-prone areas as graph nodes, dynamically updating node states through spatiotemporal edges (Lei et al., 2025). Furthermore, cross-attention mechanisms can be integrated into the model to capture spatiotemporal dependencies among contributing factors. Alternatively, gated fusion units may be incorporated to dynamically adjust the weights of multi-modal features (Yang et al., 2024a).

With the accumulation of new data and the dynamic variations in multi factor-induced landslides, regular model updates are critical to ensuring identification accuracy and adaptability. Existing studies predominantly apply these methods based on comprehensive historical landslide datasets and employ batch learning theory for identification. However, when new data becomes available, the model must be retrained from scratch. This approach is not only highly inefficient but also fails to account for the connections between newly observed and historical landslides. To address this limitation, incremental learning methods offer a promising solution. These methods enable gradual parameter optimization through new data without retraining the existing model (Huang et al., 2022). Compared to conventional deep learning models, those integrated with incremental learning can more effectively leverage historical landslide data and adaptively learn from newly incorporated data, thereby better accommodating the dynamic nature of landslides.

……

-

- Revised Description in Section 4

4.1 Application of Deep Learning in the Identification of Rainfall-induced Landslides

……

Following the development of rainfall threshold models, real-time monitoring of historically rainfall-induced landslides is imperative. First, continuous surveillance enables early detection of subtle deformations and precursory anomalies (Guzzetti et al., 2020; Zhu et al., 2023b), facilitating timely reactivation warnings to mitigate secondary hazards to lives and infrastructure. Second, by continuously monitoring rainfall, soil moisture, and groundwater levels, we can support dynamic recalibration of threshold parameters. This data assimilation enhances model adaptability to evolving hydrogeological conditions, ensuring operational relevance across heterogeneous terrains.

While the physical mechanisms governing rainfall-induced slope failures have been well studied (Arnone et al., 2011; Xiong et al., 2024), recent advances in deep learning have significantly improved our ability to automatically identify and predict such events using heterogeneous data.

In the context of rainfall-induced landslides, spatiotemporal data (e.g., rainfall intensity, cumulative precipitation, soil moisture, and slope displacement time series) are the primary inputs. Deep learning models are selected according to data characteristics and task objectives. For instance, CNNs are commonly used to extract spatial rainfall–topography features and delineate susceptible zones from remote sensing images (Peng and Wu 2024; Xu et al., 2022; Zhang et al., 2022). The encoder–decoder architecture, such as U-Net, enables pixel-level segmentation of rainfall-triggered landslides (Bhatta et al., 2025), with the inclusion of vegetation or soil moisture layers improving feature discrimination.

When temporal evolution is essential, RNNs and LSTM networks effectively model sequential dependencies between rainfall and slope deformation (Biniyaz et al., 2022; Liu et al., 2025). These models are capable of learning hysteretic responses and time lags between precipitation events and ground displacement, enabling early warning through time-series forecasting.

Furthermore, hybrid models combining CNN and LSTM components have been proposed to jointly learn spatial–temporal correlations (Chen and Fun 2022; Wu et al., 2025). By fusing rainfall distribution maps with displacement monitoring sequences, these architectures provide a more complete understanding of rainfall–landslide coupling mechanisms.

Deep learning also facilitates data-driven rainfall threshold estimation. Instead of relying solely on empirical or physically based thresholds, models such as fully connected neural networks (FNNs) and attention-based transformers can derive adaptive rainfall thresholds from multi-year rainfall–landslide records, capturing regional nonlinearities (Wu et al., 2023).

……

4.3 Application of Deep Learning in the Identification of Human Activity-induced Landslides

……

Moreover, the triggers of human activity-induced landslides are not only related to natural conditions but also closely associated with dynamic human activities (Tian et al., 2025). Consequently, their analysis necessitates the integration of multimodal and cross-scale data to capture coupled environmental and behavioral drivers (see Fig. 6). In engineering operations such as mining or road construction, factors including proximity to potential landslide zones, excavation depth, and slope angles must be rigorously evaluated through geohazard risk assessments. During excavation phases, geotechnical investigations are imperative to identify weak lithological strata or fracture-dense zones predisposed to instability. Continuous slope stability monitoring requires deploying IoT-enabled sensors to track temporal variations in surface fissure dimensions and subsurface displacement vectors. Monitoring data from these sensors can be integrated into deep learning models for multimodal analytics, enabling dynamic risk prediction and adaptive mitigation planning.

For spatial mapping and fissure extraction, CNNs and U-Net-based segmentation models have demonstrated strong capability in identifying artificial slope features from optical or SAR imagery. CNN-based models can capture high-level semantic information on excavation scars, road cuts, and spoil heaps that indicate anthropogenic disturbance. Tao et al. (2024) employed the DRs-UNet model to investigate the use of deep learning for UAV-based crack identification, the developmental patterns of fissures, and the feedback interactions between underground mining progress and corresponding surface conditions. Wu et al. (2021) proposed the PUNet model for detecting and mapping localized rapid subsidence induced by mining activities. Meng et al. (2025) introduced the GF-Former model to achieve precise segmentation of ground fissures in remote sensing imagery.

When surface deformation time series or micro-displacement data from GB-InSAR, InSAR, or IoT sensors are available, RNN-based models are applied to model the temporal evolution of slope deformation (Han et al., 2022; Li et al., 2025). These models are particularly effective in detecting precursory motion trends caused by progressive excavation or loading activities.

To mitigate misclassification between anthropogenic signatures and natural terrain, integrating multispectral data with topographic elevation data enhances discriminative capacity (Meng et al., 2021; Selamat et al., 2023). For instance, in mountainous regions, DEMs revealing artificially excavated steep slopes combined with fractured geological strata from structural maps provide preliminary evidence of human influence on landslide susceptibility (Lian et al., 2024).

In fact, landslides triggered solely by human activities are relatively rare. Single human activities are typically insufficient to independently trigger landslides, with natural factors often acting in conjunction with human activities. Furthermore, the prohibitive cost of acquiring subsurface disturbance data results in sparse historical landslide samples for specific engineering scenarios, limiting data-driven model training.

4.4 Application of Deep Learning in the Identification of Human Activity-induced Landslides

Multi factor-induced landslides result from the synergistic interaction of multiple natural and anthropogenic factors (Hao et al., 2023). Their triggering mechanisms involve the dynamic spatiotemporal coupling of these factors, driving progressive destabilization of geomaterials through cumulative strength degradation. The formation of such landslides may involve various types of movements, including collapse, creep, and flow phenomena. They often exhibit characteristics such as complexity, nonlinearity, and suddenness. Therefore, their identification is markedly more complex compared to landslides triggered by singular factors.

Unlike simpler landslide types, identifying composite landslides necessitates multimodal data fusion to holistically assess predisposing conditions (Li, 2025; Yin et al., 2023). It further requires disentangling the nonlinear superposition effects of multiple factors and quantifying their relative contributions to failure initiation. For instance, Dou et al. (2019) analyzed how earthquake intensity and rainfall metrics jointly modulate landslide susceptibility, deriving failure probabilities under varying parameter combinations. In multi factor-induced landslides, earthquakes and rainfall often interact with other factors. For instance, during heavy rainfall, the rate of landslide formation after an earthquake may be higher, possibly driven by the removal of excessively steep slopes, changes in vegetation and groundwater, and alterations in the mechanical properties of the bedrock and weathered layers in the earthquake-induced landslides canopy. This necessitates systematic investigation of multi-hazard coupling effects to quantify emergent risks.

In addition to constructing physics-based models that account for multiple factors and quantify their interactions through the solution of governing equations, GNNs can also be employed (Lei et al., 2025). These models are capable of capturing the spatiotemporal dependencies and nonlinear couplings among various triggering factors. By integrating cross-attention mechanisms, their ability to dynamically weight the relationships among stresses induced by rainfall, earthquakes, and human activities can be further enhanced. For example, Ren et al. (2025) employed a Graph Neural Network (GNN) to capture and model the complex spatiotemporal dependencies among multiple monitoring locations during landslide deformation. Zeng et al. (2022) used the graphical representation capability of the GNN model to analyze environmental relationships within a study region, where nodes were defined as geographic units delineated by terrain surface approximations, and edges captured the interactions between node pairs. Zhang et al. (2024) constructed a geographically constrained relational graph based on node features representing environmental similarity and employed a cosine similarity approach to associate landslides with their surrounding geographic environments.

Cross-attention mechanisms can also be integrated into the model to capture spatiotemporal dependencies among contributing factors. For instance, Hu et al. (2025) integrated global landslide feature vectors with local feature maps through a cross-attention mechanism to enhance the discriminative capability between landslides and background geomorphology. Alternatively, gated fusion units may be incorporated to dynamically adjust the weights of multi-modal features (Yang et al., 2024a). For instance, Liu et al. (2022) proposed a gated fusion unit (GFU) module for multimodal remote sensing image semantic classification, enabling early fusion of heterogeneous modality features.

With the accumulation of new data and the dynamic variations in multi factor-induced landslides, regular model updates are critical to ensuring identification accuracy and adaptability. Existing studies predominantly apply these methods based on comprehensive historical landslide datasets and employ batch learning theory for identification. However, when new data becomes available, the model must be retrained from scratch. This approach is not only highly inefficient but also fails to account for the connections between newly observed and historical landslides. To address this limitation, incremental learning methods offer a promising solution. These methods enable gradual parameter optimization through new data without retraining the existing model (Huang et al., 2022). Compared to conventional deep learning models, those integrated with incremental learning can more effectively leverage historical landslide data and adaptively learn from newly incorporated data, thereby better accommodating the dynamic nature of landslides.

……

Newly Added References

Arnone, E., Noto, L.V., Lepore, C. and Bras, R.L., 2011. Physically-based and distributed approach to analyze rainfall-triggered landslides at watershed scale. Geomorphology, 133(3-4), pp.121-131. doi:10.1016/j.geomorph.2011.03.019.

Bhatta, S., Roy, A. and Shahandashti, M., 2025. Land Cover Classification Using U-Net for Calibration of Rainfall-Induced Slope Susceptibility Maps. In International Conference on Transportation and Development 2025 (pp. 439-448). doi:10.1061/9780784486191.039.

Biniyaz, A., Azmoon, B., Sun, Y. and Liu, Z., 2022. Long short-term memory based subsurface drainage control for rainfall-induced landslide prevention. Geosciences, 12(2), p.64. doi:10.3390/geosciences12020064.

Chen, C. and Fan, L., 2022. CNN-LSTM-attention deep learning model for mapping landslide susceptibility in Kerala, India. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 10, pp.25-30. doi:10.5194/isprs-annals-X-3-W1-2022-25-2022.

Han, J., Yang, H., Liu, Y., Lu, Z., Zeng, K. and Jiao, R., 2022. A deep learning application for deformation prediction from ground-based insar. Remote Sensing, 14(20), p.5067. doi:10.3390/rs14205067.

Hu, W., Sun, G., Zeng, X., Tong, B., Wang, Z., Wu, X. and Song, P., 2025. Hierarchical cross attention achieves pixel precise landslide segmentation in submeter optical imagery. Scientific Reports, 15(1), p.21933. doi:10.1038/s41598-025-08695-8.

Li, J., Fan, C., Zhao, K., Zhang, Z. and Duan, P., 2025. Landslide displacement prediction using time series InSAR with combined LSTM and TCN: application to the Xiao Andong landslide, Yunnan Province, China. Natural Hazards, 121(4), pp.3857-3884. doi:10.1007/s11069-024-06937-y.

Liu, Q., Kampffmeyer, M., Jenssen, R. and Salberg, A.B., 2022. Multi-modal land cover mapping of remote sensing images using pyramid attention and gated fusion networks. International Journal of Remote Sensing, 43(9), pp.3509-3535. doi10.1080/01431161.2022.2098078.

Liu, Y., Brezzi, L., Liang, Z., Gabrieli, F., Zhou, Z. and Cola, S., 2025. Image analysis and LSTM methods for forecasting surficial displacements of a landslide triggered by snowfall and rainfall. Landslides, 22(3), pp.619-635. doi:10.1007/s10346-024-02328-3.

Meng, J., Xu, X., Li, P., Zhang, Z., Zhao, W., Ren, J. and Li, Y., 2025. Gf-former: an accurate UAV-based remote sensing image network for high-precision automatic segmentation of ground fissures in mining regions. International Journal of Machine Learning and Cybernetics, pp.1-22. doi:10.1007/s13042-025-02555-7.

Peng, B. and Wu, X., 2024. Optimizing rai.nfall-triggered landslide thresholds for daily landslide hazard warning in the Three Gorges Reservoir area. Natural Hazards and Earth System Sciences, 24(11), pp.3991-4013. doi:10.5194/nhess-24-3991-2024.

Ren X, Liu W, Yang W, Mao L, Li H. Landslide Deformation Uncertainty Quantification Using Conformalized Graph Neural Networks: A Case Study in Sichuan Province, China. IEEE Access. 2025 May 8. doi:10.1109/ACCESS.2025.3568273.

Tao, T., Han, K., Yao, X., Chen, X., Wu, Z., Yao, C., Tian, X., Zhou, Z. and Ren, K., 2024. Identification of ground fissure development in a semi-desert aeolian sand area induced from coal mining: Utilizing UAV images and deep learning techniques. Remote Sensing, 16(6), p.1046. doi:10.3390/rs16061046.

Tian, N., Lan, H., Li, L., Peng, J., Fu, B. and Clague, J.J., 2025. Human activities are intensifying the spatial variation of landslides in the Yellow River Basin. Science Bulletin, 70(2), pp.263-272. doi:10.1016/j.scib.2024.07.007.

Wu, H., Niu, J., Li, Y., Wang, Y. and Qiu, D., 2025. Landslide Susceptibility Prediction Based on a CNN–LSTM–SAM–Attention Hybrid Model. Applied Sciences, 15(13), p.7245. doi:10.3390/app15137245.

Wu, S., Li, X. and Chen, D., 2023, May. A Method of Rainfall-Runoff Prediction Based on Transformer. In Proceedings of the 15th International Conference on Digital Image Processing (pp. 1-6). doi:10.1145/3604078.3604095.

Wu, Z., Wang, T., Wang, Y., Wang, R. and Ge, D., 2021. Deep learning for the detection and phase unwrapping of mining-induced deformation in large-scale interferograms. IEEE Transactions on Geoscience and Remote Sensing, 60, pp.1-18. doi:10.1109/TGRS.2021.3121907.

Xiong, J., Pei, T. and Qiu, T., 2024. A Novel Framework for Spatiotemporal Susceptibility Prediction of Rainfall-Induced Landslides: A Case Study in Western Pennsylvania. Remote Sensing, 16(18), p.3526. doi: 10.3390/rs16183526.

Xu, G., Wang, Y., Wang, L., Soares, L.P. and Grohmann, C.H., 2022. Feature-based constraint deep CNN method for mapping rainfall-induced landslides in remote regions with mountainous terrain: An application to Brazil. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, pp.2644-2659. doi: 10.1109/JSTARS.2022.3161383.

Zeng, H., Zhu, Q., Ding, Y., Hu, H., Chen, L., Xie, X., Chen, M. and Yao, Y., 2022. Graph neural networks with constraints of environmental consistency for landslide susceptibility evaluation. International journal of geographical information science, 36(11), pp.2270-2295. doi:10.1080/13658816.2022.2103819.

Zhang, D., Wei, K., Yao, Y., Yang, J., Zheng, G. and Li, Q., 2022. Capture and prediction of rainfall-induced landslide warning signals using an attention-based temporal convolutional neural network and entropy weight methods. Sensors, 22(16), p.6240. doi: 10.3390/s22166240.

Zhang, Q., He, Y., Zhang, L., Lu, J., Gao, B., Yang, W., Chen, H. and Zhang, Y., 2024. A landslide susceptibility assessment method considering the similarity of geographic environments based on graph neural network. Gondwana Research, 132, pp.323-342. doi:10.1016/j.gr.2024.04.013.

Comment # 2:

2. Along the manuscript, I noticed several unsupported statements and a consistent lack of citations. There is also redundancy in the information presented, numerous grammar errors, and a confusing structure. For example, definitions and mechanisms of landslides appear scattered across different sections, rather than being organized logically. Since the manuscript focuses on landslides, I recommend a restructuring of the paper along the following lines:

1. Introduction
2. Landslide definition
    a. Landslide mechanisms
    b. Type of landslides
3. Deep learning for potential landslides
    a. Data sources and models
    b. Applications
    c. Challenges and Limitation
    d. Opportunities
4. Conclusions

Response:

- We sincerely thank the reviewer for the time and effort devoted to evaluating our manuscript and for providing such constructive and insightful comments! We carefully considered all suggestions and have made extensive revisions to improve the scientific rigor, clarity, and logical consistency of the paper. Below, we address each comment in detail and describe the corresponding revisions made to the manuscript.
- *** Regarding the issue of unsupported statements and a consistent lack of citations**
- Thank you for pointing out this important issue. We fully agree that some statements in the previous version lacked sufficient references. To address this, we carefully reviewed the entire manuscript and added numerous recent and authoritative citations to support key claims regarding deep

learning methods, data sources, and landslide mechanisms.

- *** Regarding the issues of redundant content and numerous grammatical errors**
- We sincerely apologize for the inadequate writing quality in the original submission and fully acknowledge that the grammar and expression did not meet the high standards required for journal publication. We appreciate your detailed remarks, which helped us identify the issues more clearly. We will also take this valuable feedback as an opportunity to strengthen our own academic writing competency in future research.
- To thoroughly address this concern, we have undertaken a comprehensive revision of the entire manuscript:
- (1) We have carefully checked and polished using advanced AI-powered writing assistance technology (specifically, Grammarly and ChatGPT) to correct grammatical errors, eliminate redundancy, and improve sentence clarity.
- (2) Following this, all authors have performed multiple rounds of manual proofreading to ensure the final text is coherent, focused, and meets the high standards expected for publication.
- *** Regarding the issue of revising the article structure**
- Thank you for this very insightful and constructive suggestion. We carefully considered the proposed restructuring and fully understand the motivation behind it. We agree that a clear definition of landslides, including their specific mechanisms and types, is a crucial aspect. We acknowledge that this was an important point missing from our original manuscript, and your detailed feedback has been instrumental in helping us refine our discussion.
- We carefully considered your recommendation to restructure the manuscript. However, since the primary objective of this review is to systematically summarize deep learning for active landslide identification, we chose to maintain the existing framework organized around data, models, applications, challenges, and opportunities. This structure better reflects the methodological logic and development trajectory of deep learning in geoscientific research. Furthermore, as our review indicates, the current mainstream application of deep learning for active landslides is heavily focused on identification and prediction. These approaches are powerful but often treat the problem primarily as one of pattern recognition from data (e.g., satellite or UAV imagery), and therefore typically do not incorporate the specific physical mechanisms or geological typologies of the landslides involved.
- Your comment provided us with significant inspiration by highlighting this gap. We strongly agree that moving beyond simple prediction is a critical and highly promising direction for the field. Integrating underlying physical mechanisms and specific landslide types is key to transforming deep learning models from "black boxes" into tools that offer deeper insights into why and how these disasters occur. This integration is also essential for ensuring the reliable and interpretable application of artificial intelligence in geoscience.

- Based on the above considerations, we decided to retain the original overall structure of the manuscript, as it aligns more closely with the thematic focus on deep learning for active landslide identification. However, we have supplemented and reorganized the relevant content concerning landslide-related concepts and classifications, which are now integrated cohesively within the **Introduction** to improve conceptual clarity and logical consistency (Please see **Comment # 4**).
- In addition, in the Section 6 "**Deep Learning for Potential Landslide Identification: Opportunities**", we have expanded our discussion of the field's current limitations and future perspectives. Specifically, we highlight the need to bridge the gap between data-driven prediction and physically interpretable understanding of landslide processes. Furthermore, in our outlook, we emphasize that developing new frameworks capable of incorporating landslide types and physical principles represents a vital avenue for future research. We believe these additions, inspired by your feedback, significantly strengthen the manuscript's scientific depth and forward-looking perspective.

Original Description in Subsection 6.3

……

Conventional knowledge-driven methods, grounded in physical mechanics, rely on precise prior knowledge of geological structures and hydrological conditions. However, landslides are influenced by complex, coupled multi-factor interactions, characterized by high parameter uncertainty, making it challenging to comprehensively address such scenarios (Roy and Saha, 2019). Purely data-driven approaches, though capable of extracting patterns from massive datasets, lack physical interpretability, are susceptible to noise interference, and struggle to establish causal relationships in prediction outcomes (Qi et al., 2024).

Building upon future disaster prevention concepts, such as "digital twin" and "smart Earth", we propose a knowledge-data dually driven paradigm for potential landslide identification (Chen et al., 2024b; Das et al., 2024; Huang et al., 2023a; Riahi et al., 2022; Sukor et al., 2019; Zhao et al., 2024c). The core concept involves leverage knowledge analysis to gain a deeper understanding of landslide triggering mechanisms and mechanical behaviors, while combine data-driven methods to extract potential landslide features and patterns from monitoring data and historical records. This synergy establishes a closed-loop "theory-practice" verification mechanism, thereby advancing the transformation of geological hazard mitigation from passive response to proactive prevention.

……

In the second stage, mechanistic constraints are integrated into the data-driven model to achieve knowledge-data dually driven fusion.

Before model construction, prior knowledge can be derived from external sources, including domain expertise, historical data, and physical principles. Alternatively, mechanistic models may be employed to preprocess raw monitoring data. The outputs of mechanistic models or prior knowledge serve as a foundation for initializing parameters in data-driven models (Cui et al., 2024; Liu et al., 2023a; Ma and Mei, 2025). This is because, in data-driven models, the selection of initial parameter values significantly impacts on both the training process and final model performance. Incorporating prior knowledge helps define more reasonable initial parameter ranges, enabling the model to converge toward near-optimal solutions earlier in the training phase.

Knowledge embedding involves translating landslide physics into model constraints to guide the training and optimization of data-driven models (Dahal and Lombardo, 2025; Liu et al., 2024). At the architectural level, layers derived from physical equations can be structurally integrated into the network design. These physical equations can even be directly encoded as network layers, forming differentiable physics-informed computational modules. Differentiability is essential to ensure that these physics-based layers function as effective computational modules within the network. This requirement stems from the fact that training relies on optimization algorithms, which adjust model parameters by computing gradients of the loss function with respect to those parameters. Only differentiable physics-encoded layers allow gradient computation during backpropagation, enabling the model to learn parameters consistent with physical laws. At the loss function level, physical equations can be directly embedded into the neural network's loss function to enforce predictions that adhere to physical principles. As the model seeks to minimize the loss function, it iteratively adjusts its parameters to align predictions with the constraints imposed by these physical equations.

……

Revised Description in Subsection 6.3

Conventional knowledge-driven methods, grounded in physical mechanics, rely on precise prior knowledge of geological structures and hydrological conditions. However, landslides are influenced by complex, coupled multi-factor interactions, characterized by high parameter uncertainty, making it challenging to comprehensively address such scenarios (Roy and Saha, 2019). Purely data-driven approaches, though capable of extracting patterns from massive datasets, lack physical interpretability, are susceptible to noise interference, and struggle to establish causal relationships in prediction outcomes (Qi et al., 2024). A critical challenge and opportunity, therefore, lies in bridging the gap between data-driven predictive capabilities and a physically interpretable understanding of landslide processes.

To bridge this critical gap, a fundamental shift towards a knowledge-data dually driven paradigm is imperative. This paradigm moves beyond simple combination to a deep integration, where physical principles actively constrain and inform the deep learning architecture. Future research should focus on developing novel frameworks capable of explicitly incorporating landslide typologies and physical laws. For instance, Physics-Informed Neural Networks (PINNs) can embed governing equations directly into the model's loss function, while knowledge graphs can structurally represent the complex relationships between predisposing factors and failure mechanisms.

This synergy, aligned with future concepts like "digital twin" and "smart Earth," establishes a closed-loop "theory-practice" verification mechanism (Chen et al., 2024b; Das et al., 2024; Huang et al., 2023a; Riahi et al., 2022; Sukor et al., 2019; Zhao et al., 2024c). The ultimate goal is to advance landslide identification from mere pattern recognition towards physically interpretable, causally-aware forecasting, thereby transforming geological hazard mitigation from passive response to proactive prevention.

……

In the second stage, mechanistic constraints are integrated into the data-driven model to achieve knowledge-data dually driven fusion.

Prior knowledge can be derived from external sources, including domain expertise, historical records, and physical principles, or mechanistic models can be employed to preprocess raw monitoring data. These outputs serve as a foundation for initializing parameters in data-driven models, which is crucial because the choice of initial values substantially affects both training efficiency and final performance (Cui et al., 2024; Liu et al., 2023a; Ma and Mei, 2025). Beyond initialization, knowledge embedding involves translating landslide physics into model constraints to guide learning and optimization (Dahal and Lombardo, 2025; Liu et al., 2024). At the architectural level, physical equations can be structurally encoded as differentiable network layers, enabling gradient-based optimization. At the loss function level, physical constraints can be directly incorporated into the training objective, ensuring that predictions remain consistent with established principles.

A representative example of this paradigm is the PINN framework (Raissi et al., 2019). PINNs embed governing equations (such as partial differential equations describing slope hydrology or stress-strain processes) into the neural network training objective, thereby constraining the learning process with domain knowledge. This approach not only reduces dependence on large annotated datasets but also enhances interpretability and cross-regional transferability (Karniadakis et al., 2021). Although applications of PINNs in landslide research remain limited (Moeineddin et al., 2023), they provide a promising avenue for bridging purely data-driven approaches with physically grounded mechanisms (Wu et al., 2022).

……

Newly Added References

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S. and Yang, L., 2021. Physics-informed machine learning. Nature Reviews Physics, 3(6), pp.422-440. doi:10.1038/s42254-021-00314-5.

Moeineddin, A., Seguí, C., Dueber, S. and Fuentes, R., 2023. Physics-informed neural networks applied to catastrophic creeping landslides. Landslides, 20(9), pp.1853-1863. doi:10.1007/s10346-023-02072-0.

Raissi, M., Perdikaris, P. and Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 378, pp.686-707. doi:10.1016/j.jcp.2018.10.045.

Wu, Y., Shao, K., Piccialli, F. and Mei, G., 2022. Numerical modeling of the propagation process of landslide surge using physics-informed deep learning. Advanced Modeling and Simulation in Engineering Sciences, 9(1), p.14. doi:10.1186/s40323-022-00228-6.

Comment # 3:

3. In addition, there is excessive discussion on the general use of deep learning, without providing sufficient concrete examples of its application to landslide prediction. I recommend focusing on the models currently used (3.a, 3.b) and on the models that could be used and how they will improve the landslide identification in 3.d. Below I provide some detailed comments (note that I did not highlight all grammar errors).

12. Line 300. Chapter 3.2. You talk a lot about each model but not the application to landslides. For example, give more details on the studies cited at line 336.

13. Line 364. Chapter 3.3. There is a lot of information but not related to landslides.

Response:

- We sincerely thank you for further highlighting the sections and passages in the manuscript where the discussion lacked specificity, which has greatly guided our precise revisions. We fully agree with your observation that the original manuscript presented an overly generalized discussion of deep learning models in Section 3, without closely linking them to practical applications in landslide prediction.
- In response to your comments (comments 3, 12, and 13), we have undertaken a focused and detailed revision of Section 3. Our revision strategy closely follows your recommendations: we concentrate on the currently utilized models, provide detailed explanations of their applications in landslide prediction, and outline specific prospective applications for models with future potential.
- Please see the new **Section 3** for details.

Original Description in Section 3

3.1 Models for Image Analysis and Processing in Potential Landslide Identification

Image data plays a critical role in potential landslide identification, especially through remote sensing, satellite, and UAV imagery. These images enable the acquisition of large-scale terrain data, encompassing complex geographical features, vegetation coverage, and ground fissures, which often serve as potential precursors to landslide occurrences. The adoption of deep learning has facilitated a shift from conventional manual visual interpretation to automated high-precision segmentation.

Convolutional neural networks (CNNs) represent the fundamental architecture in image processing. A CNN primarily comprises convolutional layers, pooling layers, and fully connected layers, each performing predefined functions on its input data (Kattenborn et al., 2021; Liu et al., 2022a).

The convolutional layer, as the core component of CNNs, contains multiple kernels that progressively extract more detailed feature representations (Hussain et al., 2019; Shi et al., 2020; Yao et al., 2021). Meanwhile, the shared-weight strategy inherent in convolutional layers allows for network training with fewer parameters than fully connected architectures. Convolutional kernels of different sizes facilitate multi-scale feature extraction. Small kernels focus on fine details, such as small cracks and the texture of localized soil loosening, while large kernels emphasize capturing overall shapes, such as the general outline of landslides and the macroscopic morphology of mountain bodies. Pooling layers, typically positioned after convolutional layers, serve to reduce the size of feature representations and enhancing the model's resistance to overfitting when handling diverse data. Common pooling methods include max pooling and average pooling, which enhance robustness to minor transformations such as translation and rotation, ensuring a degree of invariance in the features extracted by CNNs. Pooling operations downsample the convolved feature maps, reducing computational complexity while reinforcing feature robustness. Through the hierarchical stacking of multiple convolutional and pooling layers, CNNs incrementally extract more abstract and semantically rich features (Mao et al., 2024). The final fully connected layer flattens the pooled feature maps and performs classification, outputting results that distinguish potential landslide areas from non-landslide areas or enable further analysis of landslide types (Wu et al., 2024).

The layers of a CNN can be combined in various ways, forming distinct CNN architectures. These architectures are primarily determined by task requirements, which may include image classification, multi-class segmentation, or object localization within a scene.

Conventional CNNs typically consist of multiple stacked convolutional layers, pooling layers, and fully connected layers. However, increasing network depth introduces challenges such as vanishing gradients and degradation arise, resulting in model performance deterioration.

ResNet addresses these limitations by integrating residual blocks into the foundational CNN framework (Qi et al., 2020; Yang et al., 2022). These residual blocks utilize shortcut connections that preserve original feature information. This framework facilitates the construction of ultra-deep networks capable of extracting high-level semantic features for landslide detection, thereby enhancing adaptability to complex terrain classification tasks (Ullo et al., 2021). Models with higher parameter counts generally exhibit greater representational capacity but are prone to overfitting, while demanding higher computational resources and temporal costs for both training and inference. For instance, ResNet-152 contains orders of magnitude more parameters than ResNet-50, yet the latter is often preferable in computationally constrained environments due to its balanced efficiency and performance.

DenseNet is a further innovation of ResNet (Huang et al., 2017). Both of these neural networks are based on a similar idea, which is to establish a "shortcut" between different layers. However, the structure of DenseNet is simpler and more effective, with fewer parameters. In ResNet, each layer is only connected to the previous layer, while in DenseNet, each layer is directly connected to all previous layers, and each layer can obtain gradients from the loss function. This can optimize the information flow and gradients of the entire network, making it easier to train and performing better on small datasets. The structure of DenseNet can achieve better feature reuse and reduce the number of parameters. Moreover, the layers of DenseNet are narrower than those of other deep learning networks (Liu et al., 2021c), making it reduce redundancy by learning with fewer feature maps. This architecture is suitable for the extraction of multi-scale landslide features under complex terrains, even with limited landslide training samples.

With the rapid expansion of deep learning methods based on CNNs, semantic segmentation models have increasingly become the standard in landslide detection. Numerous advanced semantic segmentation networks have been proposed and validated for automatic landslide detection, significantly enhancing the efficiency and accuracy of large-scale detection. U-Net is a typical example (Ronneberger et al., 2015), which features a U-shaped architecture. U-Net employs an encoder-decoder structure, where the encoder is similar to conventional CNNs, progressively reducing image resolution and extracting features through convolution and pooling operations; the decoder then restores the image resolution through transposed convolution or upsampling operations (Dong et al., 2022; Nava et al., 2022). Skip connections bridge low-level detail features with deep semantic features, thereby refining segmentation precision.

When dealing with complex features in landslide-prone areas, DeepLab is a more suitable choice (Sandric et al., 2024). Built upon deep convolutional neural networks, DeepLab employs dilated convolutions to expand the receptive field and integrates an atrous spatial pyramid pooling (ASPP) module to capture multi-scale contextual information.

In contrast, the U-Net architecture is relatively simpler and better suited for small targets and high-resolution imagery, such as landslide crack segmentation or fine annotation of high-resolution UAV images. DeepLab, on the other hand, is more effective for large-scale landslide area detection and multispectral remote sensing image classification (see Fig. 2).

After achieving semantic segmentation to obtain the accurate extent of a landslide and the classification of ground objects, change detection is employed to monitor the changes in the landslide area over time. By comparing the segmentation results of multiple temporal phases or directly analyzing the feature differences, the dynamic evolution of potential hazards can be quantifie (Amankwah et al., 2022).

Wang (2023) demonstrates that 3D CNNs can directly process these 3D tensors. These models capture spatial and temporal features using convolutional kernels while transforming multi-temporal image sequences into change hotspot maps or temporal variation curves as output.

Some studies even have integrated attention mechanisms into conventional CNN architectures to enhance the analysis of multi-temporal remote sensing imagery, thereby enabling the identification of landslide hazard evolution over time.

3.2 Models for Time Series Analysis in Potential Landslide Identification

The occurrence of a landslide is a gradual accumulation process, usually influenced by a variety of factors. We refer to data that reflect the changing states of a landslide body over time as time series data. Time series data analysis aims to excavate the information hidden in the time series data to help identify potential landslides. Different from conventional time series data analysis methods, using deep learning models an automatically reveal the dynamic change trends and periodic patterns in the data, providing more accurate information for landslide prediction.

Recurrent neural networks (RNNs) are a class of deep learning models specialized in processing sequential data, capable of capturing temporal dependencies within input sequences (Ngo et al., 2021; Zaremba et al., 2014). Unlike conventional feedforward neural networks, in an RNN, each neuron not only receives the current input but also the output of the previous time step as additional input. This structure endows the RNN with a memory mechanism.

The architecture contains three primary components working in coordination:(1) The input layer means that one data point is input at each time step. (2) The hidden layer contains recurrent connections, which enable the information from the previous time step to be passed to the current time step, and the output serves as the input for the next time step simultaneously. (3) The output layer generates the output under the control of the state of the hidden layer (Cho et al., 2014; Zhao et al., 2021b).

During the training process, the RNN will process the data at each time step in sequence, continuously updating the hidden state. By combining the input of the current time step with the hidden state of the previous moment for calculation to gain an understanding of the data at the current moment, this structure enables the RNN to capture the temporal evolution patterns of landslide-related factors.

Due to conventional RNNs struggle to model long-term dependencies and limit their applicability to short-term temporal sequences, long short-term memory networks (LSTM) were developed (Wang et al., 2023b).

LSTM is an enhancement of RNNs, primarily processing long sequence data. Compared to standard RNNs, the hidden layer architecture of LSTM is much more complex. By incorporating memory cells and gating mechanisms, LSTM selectively propagates critical information across multiple time steps, thereby effectively capturing long-range temporal dependencies(Landi et al., 2021; Yu et al., 2019).

The basic unit of an LSTM consists of three primary gates: (1) the input gate, which determines what new information should be added to the cell state; (2) the forget gate, which decides what old information should be discarded; and (3) the output gate, which selects the information to be output from the cell state as the hidden state at the current time step (Sherstinsky, 2020; Smagulova and James, 2019; Staudemeyer and Morris, 2019). The output hidden state, after a nonlinear transformation, can be used for prediction or as the input for the next time step (Yang et al., 2019).

This structure allows the LSTM to retain key information over long sequences while selectively forgetting irrelevant information according to the requirements. Through learning from historical data, the LSTM can predict the likelihood of landslides occurring, as well as the possible scale and impact range of landslides under different future conditions.

Due to the ability to self-update weights and significantly improve network accuracy, LSTMs can also be used as a complex nonlinear component in the construction of larger deep neural networks. The model does not require separating trend and periodic components from the original deformation data, yet it can compensate for deformation trend predictions caused by unexpected interruptions in monitoring data. These properties make LSTMs particularly suited for high-accuracy research and analytical scenarios requiring large-scale datasets (Gidon et al., 2023; Xu and Niu, 2018).

Gated recurrent unit (GRU) is a simplified version of LSTM(Chung et al., 2014; Zhang et al., 2022b), which has fewer parameters. Due to their higher computational efficiency, GRU has potential advantages in real-time data processing scenarios in landslide monitoring.

GRU mainly consists of the update gate and reset gate. The update gate is used to control how much of the previous information should be preserved at the current time step, while the reset gate is used to determine whether to ignore the hidden state of the previous time step, enabling the model to adaptively learn information across different temporal scales. This dual-gate mechanism enables adaptive learning of multi-scale temporal patterns.

Compared with the LSTM, the GRU has fewer parameters and higher computational efficiency, giving it an advantage in some landslide monitoring scenarios where real-time performance is critical.

GRU is capable of effectively handling time series data with long-term dependencies, making it suitable for long-term prediction of landslide hazards. Moreover, by learning temporal patterns in historical data, GRU can identify critical conditions for landslide occurrence in advance. GRU particularly well-suited for applications involving real-time analysis of on-site monitoring data, where rapid detection of imminent landslide risks is essential and data volume is relatively limited.

Transformer was originally designed to handle sequential data in natural language processing, which was first introduced by Vaswani in 2017 (Vaswani et al., 2017). Unlike conventional recurrent and convolutional structures, the Transformer employs employs a self-attention mechanism to directly model the entire sequence.

Since the Transformer has the ability to adaptively learn latent features and patterns within the data, when it comes to processing landslide time series data, it can automatically tweak the model parameters to accommodate diverse landslide scenarios and temporal data variability (Wang et al., 2024a; Zerveas et al., 2021).

Transformer also can analyze positional relationships across the entire sequence, better capturing complex dependencies in long sequences, making it especially suitable for handling large-scale, long-term sequential datasets.

In contrast, RNN-based models exhibit a relatively simple architecture (Li et al., 2021a; Wang et al., 2020b). Their mechanisms are conceptually intuitive, making them more interpretable (see Fig. 3). On the other hand, Transformers are more complex in structure with numerous parameters, necessitating substantial computational resources during early training to process large-scale data, while being susceptible to overfitting on small datasets. Understanding how the model extracts features and makes decisions is not straightforward from large amounts of landslide data, posing challenges for its interpretability and practical deployment.

3.3 Models for Data Generation in Potential Landslide Identification

Data generation refers to modeling the underlying data distribution of data to generate entirely new samples independent of the original dataset (Kingma et al., 2014; Moreno-Barea et al., 2020; Shorten and Khoshgoftaar, 2019), thereby enriching the dataset. In potential landslide identification, data generation mitigates challenges of data scarcity and imbalanced class distributions, thereby enhancing the generalization capability of predictive models.

Deep generative models are the leading deep learning approach for synthetic data generation

(Alam et al., 2018; Karras et al., 2020; Ma et al., 2024; Xu et al., 2015). They operate on principles similar to those of deep learning, utilizing deep neural networks to learn data representations and optimizing the learning process through objective functions.

A fundamental characteristic of deep generative models lies in their probabilistic nature. These models learn an approximate probability distribution from observed samples and subsequently generate novel samples that maintain statistical consistency with the original dataset. Unlike conventional discriminative models, generative models not only classify data but also learn the underlying distribution and generate new data points. Commonly used deep generative models include generative adversarial networks (GANs), variational autoencoders (VAEs, a variant of autoencoders), and diffusion models.

GAN is a suitable choice to generate highly realistic and diverse new images (Goodfellow et al., 2014; Tran et al., 2021). Instead of explicitly modeling data distributions, GANs implicitly learn distributions through adversarial training between generator and discriminator networks.

During data generation, the generator network in a GAN synthesizes images or data resembling real samples by processing input noise vectors (Gui et al., 2021; Saxena and Cao, 2021). The discriminator, on the other hand, is used to distinguish between the generated data and the real data.These two components are continuously optimized through adversarial training. Eventually, the generator is able to produce high-quality synthetic data, which is highly similar to the real data in terms of features.

With this adversarial structure (Al-Najjar et al., 2021), GANs can generate high-quality data that closely matches the distribution of real data in an unsupervised learning context, making them well-suited for high-resolution image synthesis.

With the proposal and development of GANs, researchers have introduced various enhanced structures that are more effectively applied to potential landslide identification. For example, the conditional GAN (CGAN) (Kim and Lee, 2020; Loey et al., 2020), Pix2Pix (Qu et al., 2019), and Wasserstein GAN (WGAN) (Wang et al., 2019).

In the case of GANs, although the generated high-quality images may visually resemble real potential landslide regions, mode collapse can lead to a lack of diversity in the generated data, failing to cover all possible types of hazards (Fang et al., 2020). If certain types of potential landslides are underrepresented in the training dataset, GANs may struggle to generate those types effectively, thereby limiting the effectiveness of data augmentation. Given that the inherently unstable training process of the GANs may require more hyperparameter tuning and computational resources, this model will pose additional challenges in scenarios with limited data availability (Al-Najjar and Pradhan, 2021; Feng et al., 2024).

As a variant of the autoencoders (AEs), the variational autoencoder (VAE) introduces the idea of probabilistic generation (Kingma et al., 2013). VAE constrains the latent space through variational inference, thus enabling the generation, reconstruction, and transformation of sample data.

Compared to GANs, the samples generated by the VAE may have better diversity (Cai et al., 2024; Islam et al., 2021; Oliveira et al., 2022), because the structured constraints of its latent space are helpful for generating samples with continuous changes. This is beneficial for simulating potential landslides under different geological conditions.

The encoder of the VAE maps the input data to a low-dimensional latent space, where each vector represents the underlying features of the input. The decoder then reconstructs the original data based on the vectors in the latent space. Different from conventional AEs, the output of the VAE encoder includes two parameters: the mean value and the standard deviation. These two parameters define the probability distribution in the latent space, which is usually assumed to be a Gaussian distribution. The decoder samples a latent variable from this probability distribution and reconstructs it into output data, thus generating data with inherent randomness and diversity. Therefore, the VAE can extract latent features from landslide data and generate new landslide data based on these features.

By learning from extensive landslide datasets, VAEs capture critical geomorphological features and patterns, enabling the generation of novel samples that preserve these characteristics. This capability enables innovative applications in potential landslide analysis. This is crucial for exploring landslide scenarios under different feature combinations and identifying potential landslide patterns. Compared to GANs, VAEs exhibit superior sample diversity and training stability though the generated samples often lack the fine-grained details produced by GANs, particularly in high-resolution geospatial contexts. Moreover, VAEs may still face challenges in handling highly imbalanced data, as the generated samples tend to favor majority classes, which can limit its effectiveness in augmenting minority class data.

When computational resources and time are sufficient, and high-quality data generation with exceptional diversity is prioritized, diffusion models are the recommended choice (Croitoru et al., 2023; Yang et al., 2023a; Zhu et al., 2023a).

Diffusion models fundamentally learn the distribution of data. During training, the model applies a forward diffusion process that gradually adds noise to the original data until it approximates a Gaussian distribution. Then, in the reverse diffusion process, the model learns to iteratively refine its reconstruction of the original data distribution from the noisy data. After being fully trained, the model is able to capture the latent distribution patterns of the data, and thus can sample based on the learned distribution to generate new data (Ho et al., 2022). That is to say, by grasping the inherent laws and features of the data, the model has the ability to generate data that conforms to the distribution of the data.

Denoising diffusion probabilistic model (DDPM) is a classic implementation of the diffusion models, which lays the probabilistic framework for the diffusion models (Choi et al., 2021; Ho et al., 2020; Jing et al., 2023; Perera et al., 2023). The generation quality is optimized through variational inference and noise scheduling. Denoising diffusion implicit model (DDIM) has made improvements on the basis of DDPM (Song et al., 2020). It uses non-Markov chain reparameterization and deterministic sampling, and greatly improves the efficiency with almost no loss of quality.

Notably, DDIMs utilize the same training framework as DDPMs. If certain parameters of DDIMs are assigned particular values, its generation process becomes equivalent to DDPMs. Thus, DDIMs function as an accelerated sampling variant of DDPMs. The critical distinction lies in their sampling mechanisms. DDPMs employ stochastic and Markovian sampling, whereas DDIMs enhance efficiency through non-Markovian deterministic sampling, though this comes at the expense of reduced sample diversity.

Although diffusion models demonstrate strong capabilities in generating high-quality images and handling noise, they generate superior-quality data and ensure greater training stability compared to GANs and VAEs. However, diffusion models have not yet been widely applied directly to the identification of potential landslides and remain in the exploratory stage (see Fig. 4). We believe that as generative models advance in the field of geospatial remote sensing, they hold vast potential for application and could play a pivotal role in future landslide risk analysis and monitoring systems.

3.4 Models for Data Cleaning in Potential Landslide Identification

In potential landslide identification, data cleaning, particularly anomaly detection, is a critical issue (Deijns et al., 2020; Jiang et al., 2020). It can distinguish between normal fluctuations and true anomalies, identifying early signs such as subtle changes in the mountain's state or abnormal trends in surface displacement, thus enabling more accurate landslide hazard assessment. With the rapid development of deep learning, the applications in data cleaning have become increasingly widespread, enabling models to automatically learn latent data patterns and identify potential anomalies.

AEs and their variational counterparts are highly effective in unsupervised data cleaning. These models autonomously learn normal geomechanical patterns from data and flag deviations, achieving effective hazard identification even when labeled anomaly samples are scarce.

The AE is a typical unsupervised learning model consisting of an encoder and a decoder The encoder compresses the input data into low-dimensional features, and then the decoder reconstructs the input. During the training process, the autoencoder learns the intrinsic features and patterns of normal landslide data, so that for normal data, the reconstruction error is small. When abnormal landslide data is input, due to the difference between its features and the distribution of normal data, the reconstruction error will be large.

When performing anomaly detection, a suitable reconstruction error threshold is set. When the reconstruction error of the test data exceeds this threshold, it can be determined as abnormal data. In the anomaly detection of landslide displacement data monitored by sensors, if the error of the displacement data after being reconstructed by AEs during a certain period is significantly higher than the normal level, it may indicate that there is an abnormal situation of potential landslides during this period.

As previously introduced, VAE is an extension of AE. Compared to conventional autoencoders, VAE introduces randomness into the latent space, making it more effective in handling data uncertainty (Li et al., 2020; Park et al., 2018).

During training, VAEs learn the latent distribution of the data and can generate new samples resembling the training set. When input samples deviate significantly from this learned distribution, the VAE fails to reconstruct them accurately, thereby flagging anomalies through elevated reconstruction errors. For landslide monitoring, if a VAE is trained on imagery of stable slopes, it internalizes stable terrain features. When an image significantly differs from the stable region, the model will produce a high reconstruction error, indicating the presence of anomalous data.

In contrast, AEs are well-suited for univariate anomaly detection, particularly for landslide precursor detection, while VAEs capture latent space distributions and are more effective for multivariate anomaly detection.

GANs can also be utilized in data cleaning (Kang et al., 2024; Xia et al., 2022). In data cleaning, the discriminator is trained to distinguish between generated data and real data. When new test data is input, if the discriminator struggles to determine whether it is real or generated data, the test data may significantly deviate from the distribution of normal data, indicating a potential anomaly. In landslide monitoring, data may be influenced by various factors, GANs demonstrate robustness by filtering out such interference, thereby enhancing data cleaning accuracy (Radoi, 2022).

AnoGAN extends conventional GANs by directly incorporating data cleaning as one of its primary objectives (Lin et al., 2023; Thomine et al., 2023). It introduces an additional encoder during training, which maps input data to the latent space. The difference between this latent vector and the latent vector of normal samples generated by the generator serves as the basis for data cleaning.

RNNs, LSTMs, and GRUs are also effective for identifying anomalous patterns in sequential data (Zhang et al., 2022a). In potential landslide identification, these models process time series inputs to learn normal temporal dynamics and trends. When new data deviates significantly from the normal patterns learned by the model, such deviations can be flagged as anomalies. However, these models are primarily used for time series data, performing data cleaning by predicting future values of the sequence. For instance, if displacement measurements exhibit abrupt deviations while rainfall remains within historical norms, the model detects such discrepancies by comparing observed values with predictions based on learned temporal dependencies.

3.5 Models for Data Fusion in Potential Landslide Identification

In practical applications, the identification of potential landslide hazards is a complex task that influences by multiple factors (Zhang et al., 2018). These factors are often reflected through different data sources. We can roughly divide heterogeneous data into four categories: image data, time series data, structured data, and textual data. Data fusion is essential for the accurate identification of potential landslides. In order to better identify potential landslides, data fusion is essential.

Since the features, scales, and resolutions of heterogeneous data are all different, currently, the powerful feature learning ability of deep learning models is often utilized to automatically capture the nonlinear relationships and high-order interaction information among these heterogeneous data.

Due to the complex non-Euclidean structural characteristics of the geological environment, topographic data and their spatial relationships related to landslide hazards, conventional methods such as CNNs have difficulty in handling these relationships. As a neural network architecture for processing graph-structured data, graph neural networks (GNNs) can effectively model such spatial relationships (Ying et al., 2018; Zeng et al., 2022). They can treat the nodes in the geographical space (such as different geographical location points) and their connection relationships (such as the distance between adjacent nodes, terrain undulations, etc.) as the structure of a graph for processing.

When dealing with heterogeneous data, GNNs support feature interaction between different types of nodes through the message passing mechanism, thereby eliminating redundancy and mutual exclusivity among data sources and enabling dynamic fusion of multi-modal features (Zhang et al., 2024d; Zhao et al., 2024b). By passing and aggregating information across nodes, GNNs can also conduct a detailed analysis of various heterogeneous data in local areas. This capability allows GNNs to capture subtle geological structural changes and detect localized anomalies inmonitoring data, providing advantages for analyzing local features and early signs of potential landslide movements.

By learning a large amount of landslide potential cases, GNNs can discover the general patterns and rules of landslides, thus having good generalization ability. When facing new and unseen regions or data, GNNs can predict and assess the potential landslides in those regions based on the knowledge they have already learned.

According to the differences in message passing and aggregation methods, GNNs have derived various variants. For example, graph convolutional network (GCN) is generated by generalizing the convolutional operation to graph-structured data (Sharma et al., 2022; Wang et al., 2020a), and graph attention network (GAT) dynamically weights the importance of neighboring nodes by introducing the attention mechanism (Yuan et al., 2022; Zhou and Li, 2021). The emergence of these new architectures makes GNN variants more targeted than conventional GNNs and suitable for modeling heterogeneous relationships. Currently, they are often used for weighted analysis of the impacts of different geographical factors on landslides.

Transformer is also composed of stacked encoders and decoders (see Fig. 5). However, unlike other architectures, the Transformer architecture introduces the self-attention mechanism (Zhao et al., 2021a), which is a crucial innovation. This enables the Transformer to automatically calculate a weight vector for each position in the input sequence based on the relationship between this position and other positions, so as to represent the importance of this position in the entire sequence. Such a weight vector can be regarded as the "attention distribution" of each position in the input sequence, that is, the model determines which positions in the sequence to focus on. By considering all positions in the input sequence simultaneously, Transformer is able to calculate the correlations between each position and other positions in the sequence in parallel (Esser et al., 2021; Huang and Chen, 2023; Zerveas et al., 2021), rather than processing them step by step like CNNs or RNNs.

Transformer can also convert multimodal dFor different types of data, it transforms them into vector representations via different embedding layers.ata into a unified vector representation through different embedding layers. Subsequently, through the use of the self-attention mechanism and multilayer neural networks, these vectors are fused and feature representations are extracted, enabling the model to process and integrate data from various modalities within the same model framework (Lv et al., 2023; Tang et al., 2022).

Revised Description in Section 3

3.1 Models for Image Analysis and Processing in Potential Landslide Identification

Image data plays a critical role in potential landslide identification, especially through remote sensing, satellite, and UAV imagery. These images enable the acquisition of large-scale terrain data, encompassing complex geographical features, vegetation coverage, and ground fissures, which often serve as potential precursors to landslide occurrences. The adoption of deep learning has facilitated a shift from conventional manual visual interpretation to automated high-precision segmentation.

Convolutional neural networks (CNNs) represent the fundamental architecture in image processing. A CNN primarily comprises convolutional layers, pooling layers, and fully connected layers, each performing predefined functions on its input data (Kattenborn et al., 2021; Liu et al., 2022a).

Convolutional layers, the core of CNNs, use kernels of various sizes to extract multi-scale features from geospatial imagery, which is crucial for landslide identification (Hussain et al., 2019; Shi et al., 2020; Yao et al., 2021). Small kernels are effective in detecting fine-grained precursors such as ground fissures and localized soil texture changes. For instance, Hamaguchi et al. (2018) proposed a Local Feature Extraction (LFE) module to enhance the capability of CNNs in identifying small object instances in remote sensing imagery. Wang et al. (2024) demonstrated the exceptional capability of convolutional layers in extracting extremely small and subtle features by identifying cracks as narrow as 0.05 m width using a U-Net–based model. In contrast, larger kernels help in recognizing the overall morphology and boundaries of landslide bodies. From the perspective of general visual tasks, Ding et al. (2022) demonstrated that larger convolution kernels substantially improve the shape bias of CNNs, facilitating the recognition of large-scale structures and overall morphological patterns compared with using small kernels alone. Li et al. (2025) employed multiple large convolution kernels (kernel sizes = 5, 7, and 9) within the DLFFSKA module to fuse multi-scale features, thereby enhancing the global perception of landslide boundaries and morphology as well as the capture of contextual background information.

Pooling layers downsample feature maps, improving computational efficiency and model robustness. In landslide mapping, this translation invariance is particularly beneficial, as it allows the model to consistently identify landslide features regardless of their slight positional variations across different image patches (Mao et al., 2024).

The final fully connected layer flattens the pooled feature maps and performs classification, outputting results that distinguish potential landslide areas from non-landslide areas or enable further analysis of landslide types (Wu et al., 2024).

The layers of a CNN can be combined in various ways, forming distinct CNN architectures. These architectures are primarily determined by task requirements, which may include image classification, multi-class segmentation, or object localization within a scene.

Conventional CNNs typically consist of multiple stacked convolutional layers, pooling layers, and fully connected layers. However, increasing network depth introduces challenges such as vanishing gradients and degradation arise, resulting in model performance deterioration.

ResNet mitigates the vanishing gradient problem in very deep networks through residual connections (Qi et al., 2020; Yang et al., 2022). This architectural advancement has been successfully applied to landslide detection in complex terrains, such as the work by Ullo et al. (2021), who demonstrated that a ResNet-based classifier could achieve high accuracy in distinguishing landslide scars from surrounding vegetation and bare soil in satellite imagery by effectively learning hierarchical features.

Models with higher parameter counts generally exhibit greater representational capacity but are prone to overfitting, while demanding higher computational resources and temporal costs for both training and inference (Ebrahimi et al., 2021). For instance, He et al. (2016) introduced ResNet-152 and other deep residual network architectures, demonstrating that deeper structure achieve superior performance compared with shallower counterparts. Hasanah et al. (2023) explicitly highlighted the differences in layer depth and parameter count among various ResNet versions (ResNet-50, 101, and

152), noting that the increased number of parameters in deeper networks inevitably leads to longer training times.

DenseNet is a further innovation of ResNet (Huang et al., 2017). Both of these neural networks are based on a similar idea, which is to establish a "shortcut" between different layers. However, the structure of DenseNet is simpler and more effective, with fewer parameters. In ResNet, each layer is only connected to the previous layer, while in DenseNet, each layer is directly connected to all previous layers, and each layer can obtain gradients from the loss function. This can optimize the information flow and gradients of the entire network, making it easier to train and performing better on small datasets. The structure of DenseNet can achieve better feature reuse and reduce the number of parameters. Moreover, the layers of DenseNet are narrower than those of other deep learning networks (Liu et al., 2021c), making it reduce redundancy by learning with fewer feature maps. This architecture is suitable for the extraction of multi-scale landslide features under complex terrains, even with limited landslide training samples.

With the rapid expansion of deep learning methods based on CNNs, semantic segmentation models have increasingly become the standard in landslide detection. Numerous advanced semantic segmentation networks have been proposed and validated for automatic landslide detection, significantly enhancing the efficiency and accuracy of large-scale detection.

U-Net's encoder-decoder structure with skip connections has become a benchmark for landslide segmentation (Ronneberger et al., 2015). For example, Dong et al. (2022) proposed a new model, L-UNet, based on the U-Net architecture and successfully applied it to landslide extraction from remote sensing imagery. Nava et al. (2022) applied the Attention U-Net to Sentinel-1 SAR data for rapid mapping of earthquake-induced landslides, demonstrating the effectiveness of U-Net variants in pixel-level segmentation of landslide bodies under cloud-covered or topographically complex conditions.

When dealing with complex features in landslide-prone areas, DeepLab is a more suitable choice (Sandric et al., 2024). Built upon deep convolutional neural networks, DeepLab employs dilated convolutions to expand the receptive field and integrates an atrous spatial pyramid pooling (ASPP) module to capture multi-scale contextual information.

In contrast, the U-Net architecture is relatively simpler and better suited for small targets and high-resolution imagery, such as landslide crack segmentation or fine annotation of high-resolution UAV images. DeepLab, on the other hand, is more effective for large-scale landslide area detection and multispectral remote sensing image classification (see Fig. 2).

After achieving semantic segmentation to obtain the accurate extent of a landslide and the classification of ground objects, change detection is employed to monitor the changes in the landslide area over time. By comparing the segmentation results of multiple temporal phases or directly analyzing the feature differences, the dynamic evolution of potential hazards can be quantifie (Amankwah et al., 2022).

Wang (2023) demonstrates that 3D CNNs can directly process these 3D tensors. These models capture spatial and temporal features using convolutional kernels while transforming multi-temporal image sequences into change hotspot maps or temporal variation curves as output.

Some studies even have integrated attention mechanisms into conventional CNN architectures to enhance the analysis of multi-temporal remote sensing imagery, thereby enabling the identification of landslide hazard evolution over time.

3.2 Models for Time Series Analysis in Potential Landslide Identification

The occurrence of a landslide is a gradual accumulation process, usually influenced by a variety of factors. We refer to data that reflect the changing states of a landslide body over time as time series data. Time series data analysis aims to excavate the information hidden in the time series data to help identify potential landslides.

Different from conventional statistical or physical models, deep learning models can automatically reveal dynamic change trends and periodic patterns in the data, providing more accurate information for landslide prediction and early warning. Recently, deep learning–based temporal models, such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), and gated recurrent units (GRUs), have become key tools for extracting nonlinear dependencies and temporal evolution patterns in landslide-related time series.

Recurrent neural networks (RNNs) are a class of deep learning models specialized in processing sequential data, capable of capturing temporal dependencies within input sequences (Ngo et al., 2021; Zaremba et al., 2014). Unlike conventional feedforward neural networks, in an RNN, each neuron not only receives the current input but also the output of the previous time step as additional input. This structure endows the RNN with a memory mechanism.

In landslide prediction, RNNs have been employed to model displacement time series under rainfall or groundwater fluctuations, revealing short-term deformation patterns preceding slope failure (Chen et al., 2015; Zhang et al., 2022).

To overcome the vanishing gradient problem inherent in RNNs, LSTMs introduce memory cells and gating mechanisms that selectively retain relevant temporal information (Landi et al., 2021; Sherstinsky, 2020; Smagulova and James, 2019; Staudemeyer and Morris, 2019; Yu et al., 2019). This capability allows them to model the cumulative and delayed responses of slopes to prolonged rainfall or reservoir water level fluctuations.

LSTM models have been widely applied in landslide displacement prediction and early warning. Yang et al. (2019) analyzed the relationships among landslide deformation, rainfall, and reservoir water levels, and found that compared with static models, the LSTM approach more accurately captured the dynamic characteristics of landslides and effectively leveraged historical information. Xu and Niu (2018) used a LSTM model to predict the displacement evolution of the Baijiabao landslide using rainfall and hydrological level data, achieving a higher correlation compared with traditional regression models. In another study focused on shallow landslides, Xiao et al (2022) used a week-ahead LSTM model, which exhibited stable performance and improved prediction accuracy in short-term prediction scenarios. Additionally, Gidon et al (2023) constructed a Bi-LSTM model and achieved a detection accuracy of 93% in the Mawiongrim area. This effectively addresses the limitations of traditional methods and can provide a reliable technical solution for disaster early warning in this area as well as other similar landslide-prone areas.

The GRU is a simplified variant of the LSTM (Chung et al., 2014; Zhang et al., 2022b) that achieves similar accuracy with fewer parameters and reduced computational costs, making it well-suited for real-time landslide monitoring systems (Rawat et al., 2024).

Furthermore, GRU models effectively identify precursory displacement acceleration, allowing early detection of slope instability triggered by rainfall or seismic shaking (Chang et al., 2025; Yang et al., 2025).

Recently, Transformer architectures have been introduced for time series modeling due to their ability to capture global dependencies across long sequences through the self-attention mechanism (Vaswani et al., 2017).

Unlike RNNs or LSTMs that process data sequentially, Transformers analyze all time steps in parallel, offering better scalability and modeling of long-term deformation trends (Esser et al., 2021; Huang and Chen, 2023; Zerveas et al., 2021).

In landslide applications, Transformer-based approaches have shown promise in integrating multi-source time series—such as rainfall, soil moisture, and deformation—into a unified temporal framework. Zhao et al. (2024) combined the strengths of CNN and Transformer architectures, selecting and analyzing nine landslide-conditioning factors to successfully achieve accurate landslide localization and detailed feature capture. Ge et al. (2024) proposed the LiteTransNet model based on the Transformer framework, effectively capturing and interpreting the varying importance of historical information during the prediction process.

Although Transformer-based models require larger training datasets and higher computational resources, their capacity to model complex, long-range dependencies and integrate multi-factor information offers significant potential for the next generation of intelligent landslide early warning systems.

In summary, RNNs and their advanced variants (LSTM, GRU) have demonstrated strong capabilities in modeling landslide time series, enabling early detection of slope deformation acceleration and rainfall-induced instability (Li et al., 2021a; Wang et al., 2020b). Transformer architectures further extend this capability to capture cross-variable and long-term dependencies (Wang et al., 2024a; Zerveas et al., 2021; Zhuang et al., 2023), offering a new direction for multi-sensor, data-driven landslide prediction (see Fig. 3).

3.3 Models for Data Generation in Potential Landslide Identification

Data generation refers to modeling the underlying data distribution of data to generate entirely new samples independent of the original dataset (Kingma et al., 2014; Moreno-Barea et al., 2020; Shorten and Khoshgoftaar, 2019), thereby enriching the dataset. In potential landslide identification, data generation mitigates challenges related to data scarcity and class imbalance, which are particularly pronounced in geohazard mapping tasks where labeled landslide samples are limited. This process enhances the generalization capability of predictive models and enables the simulation of diverse landslide scenarios.

Deep generative models are the leading deep learning approach for synthetic data generation (Alam et al., 2018; Karras et al., 2020; Ma et al., 2024; Xu et al., 2015). They utilize deep neural networks to learn latent representations of data and optimize the learning process through specific objective functions. A key characteristic of deep generative models lies in their probabilistic nature—they not only classify or reconstruct data but also capture the underlying distribution of geospatial features, enabling the creation of new landslide samples that are statistically consistent with observed patterns. Commonly used deep generative models include generative adversarial networks (GANs), variational autoencoders (VAEs), and diffusion models.

GANs consist of a generator and a discriminator that compete in an adversarial process (Goodfellow et al., 2014). The generator synthesizes data resembling real samples, while the discriminator attempts to distinguish between generated and real data. Through iterative adversarial training, the generator learns to produce high-quality synthetic data that closely matches the distribution of real data (Gui et al., 2021; Saxena and Cao, 2021).

In the context of landslide studies, GANs have demonstrated strong capabilities in data augmentation and remote sensing image enhancement. For example, Al-Najjar and Pradhan (2021) proposed a novel approach that employs a GAN to generate synthetic inventory data. The results indicate that additional samples produced by the proposed GAN model can enhance the predictive performance of Decision Trees (DT), Random Forest (RF), Artificial Neural Network (ANN), and Bagging ensemble models. Feng et al. (2024) achieved the first implementation of using a GAN to generate synthetic high-quality landslide images, aiming to address the data scarcity issue that undermines the performance of landslide segmentation models.

Despite their advantages, GANs may suffer from mode collapse, leading to limited diversity in the generated data, especially when certain landslide types are underrepresented (Fang et al., 2020). Moreover, their unstable training process requires careful hyperparameter tuning and substantial computational resources, which may constrain their application in real-time hazard scenarios. Nevertheless, with improved architectures such as CGAN, Pix2Pix, and Wasserstein GAN (WGAN) (Kim and Lee, 2020; Loey et al., 2020; Qu et al., 2019; Wang et al., 2019), GANs are becoming increasingly viable tools for high-resolution landslide mapping and synthetic data generation in remote sensing-based susceptibility analyses.

As a probabilistic variant of autoencoders (AEs), VAEs introduce latent-space regularization through variational inference (Kingma et al., 2013). The encoder compresses input data into a latent representation characterized by a mean and a standard deviation, while the decoder reconstructs the data by sampling from this distribution. This enables the model to generate new data with inherent randomness and diversity (Islam et al., 2021; Oliveira et al., 2022).

In landslide research, VAEs have been successfully applied to learn and reconstruct geomorphological patterns of slope instability. For instance, Cai et al. (2024) proposed and demonstrated the superior capability of the VAE-GRU model in generating narrow predictive intervals while maintaining high coverage probabilities, representing a substantial improvement over the state-of-the-art methods for probabilistic landslide prediction.

Compared with GANs, VAEs produce more diverse but slightly less detailed samples, due to their structured latent space constraints. This characteristic is particularly beneficial for exploring a wide range of potential landslide morphologies and for augmenting training datasets used in susceptibility prediction. However, VAEs may still struggle with highly imbalanced datasets, as their probabilistic reconstruction tends to favor majority classes. Integrating VAEs with stratified sampling or cost-sensitive learning could help overcome this limitation and further enhance landslide prediction performance.

When computational resources and training time permit, diffusion models provide a powerful alternative for generating high-quality, diverse, and stable data (Ho et al., 2020; Croitoru et al., 2023; Yang et al., 2023a; Zhu et al., 2023a). These models learn the data distribution by gradually adding noise to real samples (forward diffusion) and then reconstructing clean data through a reverse denoising process. The resulting models can sample new, realistic data points that reflect complex terrain and geophysical variability.

Although diffusion models are still in the exploratory phase for landslide applications, recent geospatial AI research indicates their high potential for terrain simulation and deformation modeling. Lo et al. (2024) proposed a Terrain-Feature-Guided Diffusion Model (TFDM) to fill gaps in DEM data. Similarly, Zhao et al. (2024) employed a Denoising Diffusion Probabilistic Model (DDPM) conditioned on incomplete DEMs, which serves as a transitional kernel during diffusion reversal to progressively reconstruct sharp and accurate DEM surfaces.

Deep generative models provide a transformative solution for overcoming the challenges of limited and imbalanced landslide datasets. By synthesizing realistic, diverse, and statistically consistent samples, these models can improve the robustness and generalization of landslide prediction frameworks. GANs are effective for generating visually realistic imagery and data augmentation; VAEs capture probabilistic geomorphic transitions; and diffusion models ensure stability and fidelity in high-resolution terrain synthesis.

As generative AI continues to evolve, integrating these models with multi-source remote sensing inputs and physics-based constraints holds great promise for next-generation landslide hazard identification systems. Such integration is expected to enhance data diversity, reduce labeling dependency, and enable more precise, interpretable, and generalizable predictions for landslide risk assessment and early warning.

3.4 Models for Anomaly detection in Potential Landslide Identification

Anomaly detection plays a critical role in potential landslide identification, as it enables the distinction between normal environmental variations and genuine precursors of slope instability (Deijns et al., 2020; Jiang et al., 2020). In landslide monitoring, the goal of anomaly detection is to identify subtle yet significant deviations—such as abnormal surface displacements, changes in surface coherence, or irregular sensor signals—that may occur prior to failure events. With the advancement of deep learning, data filtering has evolved from rule-based threshold detection to automated feature learning, allowing models to capture complex spatiotemporal dependencies and identify anomalies within high-dimensional, multi-source datasets.

AEs are widely used for unsupervised anomaly detection due to their ability to reconstruct input data and highlight deviations from learned normal patterns. An AE consists of an encoder that compresses data into a low-dimensional latent representation and a decoder that reconstructs it.

During training, the AE learns the intrinsic features of normal landslide data—such as sensor-based displacement time series or radar backscatter from stable slopes. When abnormal data are input, such as sudden displacement spikes or incoherent radar signals, the reconstruction error increases significantly, serving as an indicator of potential instability. For instance, Shakeel et al. (2022) developed an InSAR deformation anomaly detector based on an AE–LSTM architecture (ALADDIn). Experimental analyses using synthetic deformation test scenarios achieved an overall performance accuracy of 91.25%.

By defining a reconstruction error threshold, anomalies can be quantitatively detected. When the reconstruction error of new sensor data exceeds this threshold, it may signal slope movement acceleration or surface disturbance associated with potential landslides. Thus, AEs provide a data-driven method to detect early-warning signs without requiring manually labeled failure data.

As previously introduced, VAE a probabilistic extension of AEs. VAEs introduce stochastic latent variables characterized by mean and variance, allowing them to model data uncertainty (Kingma et al., 2013; Li et al., 2020; Park et al., 2018). During training, VAEs learn the latent distribution of normal samples and reconstruct inputs accordingly. When new observation data deviate significantly from the learned distribution, the reconstruction error increases accordingly, and this phenomenon can be used as an indicator of potential anomalies.

In landslide applications, VAEs have been shown to outperform traditional AEs in handling complex, multivariate datasets that integrate topographic, meteorological, and geotechnical factors. For example, Han et al. (2025) proposed an unsupervised failure mode recognition algorithm based on a deep convolutional autoencoder, which integrates surface displacement, vertical displacement, and rainfall monitoring data from slopes to accurately identify the developmental stages of slope failure, achieving a recognition accuracy of 99.30%.

Another study by Yadav et al. (2024) proposed a novel unsupervised change detection (CD) model, termed CLVAE, designed to learn the spatiotemporal correlations within Sentinel-1 SAR time series. The model achieved a mean IoU of 70% and a mean F1-score of 81%, outperforming comparative models by at least 6% in F1-score and 8% in IoU.

Compared to AEs, VAEs are particularly advantageous for capturing uncertainty and latent correlations between environmental variables, making them ideal for data cleaning in integrated landslide early-warning systems. However, they require larger datasets for stable training, and their probabilistic outputs may demand post-processing for operational thresholding.

GANs can also be adapted for anomaly detection by exploiting their discriminator network's ability to distinguish between real and generated data (Kang et al., 2024; Xia et al., 2022). In landslide monitoring, GAN-based anomaly detection models learn the distribution of stable slope features, and deviations from this distribution can indicate abnormal conditions (Radoi, 2022).

AnoGAN extends conventional GANs by directly incorporating data cleaning as one of its primary objectives (Lin et al., 2023; Thomine et al., 2023). It introduces an additional encoder during training, which maps input data to the latent space. The difference between this latent vector and the latent vector of normal samples generated by the generator serves as the basis for data cleaning.

RNNs and their variants are particularly effective for time series–based anomaly detection, learning temporal dependencies and predicting future trends (Zamanzadeh et al., 2024; Zhang et al., 2022a). In landslide monitoring, these models can process continuous displacement or rainfall time series to identify deviations from expected temporal behavior. These temporal models complement image-based approaches by providing continuous surveillance and early detection capabilities (Wu et al., 2024).

When combined with AEs or GANs, RNN-type architectures can form hybrid frameworks capable of both spatial and temporal anomaly detection, enabling multi-source consistency checking in landslide early-warning systems. Geiger et al (2020) demonstrated a growing trend of utilizing LSTM networks as both the generator and discriminator within GAN frameworks for time-series anomaly detection. Similarly, Whitakeret al (2023) illustrated the application of LSTM–GAN architectures in identifying temporal anomalies.

3.5 Models for Data Fusion in Potential Landslide Identification

In practical applications, the identification of potential landslide hazards is a complex task that influences by multiple factors (Zhang et al., 2018). These factors are often reflected through different data sources. We can roughly divide heterogeneous data into four categories: image data, time series data, structured data, and textual data. Data fusion is essential for the accurate identification of potential landslides. In order to better identify potential landslides, data fusion is essential.

Since heterogeneous data differ in feature scale, spatial resolution, and data modality, deep learning models are increasingly utilized to automatically extract nonlinear and high-order feature interactions across data sources, offering significant advantages over traditional statistical fusion techniques. In landslide applications, deep learning-based data fusion can integrate multi-modal inputs such as Sentinel-1 InSAR deformation, rainfall time series, and terrain derivatives for regional-scale susceptibility mapping or real-time early warning.

Due to the non-Euclidean and topologically complex nature of landslide-related terrain, conventional CNN-based models are limited in representing irregular spatial dependencies. Graph Neural Networks (GNNs) have emerged as powerful architectures to model such relationships by representing spatial entities (e.g., slope units, grid cells, or sensor nodes) as graph nodes and their geospatial or topological interactions as edges (Ying et al., 2018; Zeng et al., 2022).

In landslide identification, GNNs enable explicit modeling of spatial connectivity and geological adjacency, allowing the propagation of geomorphic and hydrological information across neighboring units. For example, Kuang et al. (2022) proposed an innovative landslide forecasting model based on Graph Neural Networks, in which graph convolutions are employed to aggregate spatial correlations among different monitoring sites. Ren et al. (2025) introduced a novel Graph Neural Network framework with conformal prediction (GNN-CF) for landslide deformation interval forecasting, addressing the limitations of traditional models in handling predictive uncertainty.

According to the differences in message passing and aggregation methods, GNNs have derived various variants. For example, graph convolutional network (GCN) is generated by generalizing the convolutional operation to graph-structured data (Sharma et al., 2022; Wang et al., 2020a), and graph attention network (GAT) dynamically weights the importance of neighboring nodes by introducing the attention mechanism (Yuan et al., 2022; Zhou and Li, 2021). The emergence of these new architectures makes GNN variants more targeted than conventional GNNs and suitable for modeling heterogeneous relationships. Currently, they are often used for weighted analysis of the impacts of different geographical factors on landslides.

Transformer architectures, characterized by the self-attention mechanism, provide another promising avenue for landslide-related data fusion (Huang and Chen, 2023; Zhao et al., 2021a). Unlike CNNs or RNNs, which process spatial or temporal sequences sequentially, Transformers can jointly capture long-range dependencies across spatial and temporal dimensions, enabling unified processing of rainfall, InSAR time series, and topographic data (Esser et al., 2021; Lv et al., 2023).

Recent studies have begun adapting Transformer variants for landslide identification. Li et al. (2023) proposed a Transformer-based deep neural network capable of identifying landslides from hillshade maps and optical imagery. Piran et al. (2024) enhanced short-term precipitation forecasting by applying transfer learning with a pre-trained Transformer model. Zhang et al. (2024) incorporated Transformer modules to build a graph-Transformer model that integrates global contextual information for the generation and analysis of landslide susceptibility maps (LSMs).

Newly Added References

Cai, W., Lan, F., Huang, X., Hao, J., Xia, W., Tang, R., Feng, P. and Li, H., 2024. Generative probabilistic prediction of precipitation induced landslide deformation with variational autoencoder and gated recurrent unit. Frontiers in Earth Science, 12, p.1394129. doi:10.3389/feart.2024.1394129.

Chang, F., Dong, S., Yin, H., Ye, X., Wu, Z., Zhang, W. and Zhu, H., 2025. 3D displacement time series prediction of a north-facing reservoir landslide powered by InSAR and machine learning. Journal of Rock Mechanics and Geotechnical Engineering. doi:10.1016/j.jrmge.2024.10.033.

Chen, H., Zeng, Z. and Tang, H., 2015. Landslide deformation prediction based on recurrent neural network. Neural Processing Letters, 41(2), pp.169-178. doi:10.1007/s11063-013-9318-5.

Ding, X., Zhang, X., Han, J. and Ding, G., 2022. Scaling up your kernels to 31x31: Revisiting large kernel design in cnns. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11963-11975). doi:10.48550/arXiv.2203.06717.

Ebrahimi, M.S. and Abadi, H.K., 2021, July. Study of residual networks for image recognition. In Intelligent Computing: Proceedings of the 2021 Computing Conference, Volume 2 (pp. 754-763). Cham: Springer International Publishing. doi:10.1007/978-3-030-80126-7_53.

Ge, Q., Li, J., Wang, X., Deng, Y., Zhang, K. and Sun, H., 2024. LiteTransNet: An interpretable approach for landslide displacement prediction using transformer model with attention mechanism. Engineering Geology, 331, p.107446. doi:10.1016/j.enggeo.2024.107446.

Geiger, A., Liu, D., Alnegheimish, S., Cuesta-Infante, A. and Veeramachaneni, K., 2020, December. Tadgan: Time series anomaly detection using generative adversarial networks. In 2020 ieee international conference on big data (big data) (pp. 33-43). IEEE. doi:10.1109/BigData50022.2020.9378139.

Hamaguchi, R., Fujita, A., Nemoto, K., Imaizumi, T. and Hikosaka, S., 2018, March. Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery. In 2018 IEEE winter conference on applications of computer vision (WACV) (pp. 1442-1450). IEEE. doi:10.1109/WACV.2018.00162.

Han, N., Miao, W., Li, M., Mohamad Ismail, M.A., Hu, Q., Duan, L. and Tang, J., 2025. Integrating multi-source monitoring data and deep convolutional autoencoder technology for slope failure pattern recognition. Frontiers in Earth Science, 13, p.1531857. doi:10.3389/feart.2025.1531857.

Hasanah, S.A., Pravitasari, A.A., Abdullah, A.S., Yulita, I.N. and Asnawi, M.H., 2023. A deep learning review of resnet architecture for lung disease Identification in CXR Image. Applied sciences, 13(24), p.13111. doi:

He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). doi:10.3390/app132413111.

Kuang, P., Li, R., Huang, Y., Wu, J., Luo, X. and Zhou, F., 2022. Landslide displacement prediction via attentive graph neural network. Remote Sensing, 14(8), p.1919. doi:10.3390/rs14081919.

Li, J., Li, Q., Lu, J., Zheng, K., Wei, L. and Xiang, Q., 2025. A Transfer Learning Remote Sensing Landslide Image Segmentation Method Based on Nonlinear Modeling and Large Kernel Attention. Applied Sciences, 15(7), p.3855. doi:10.3390/app15073855.

Lo, K.S.H. and Peters, J., 2024. Diff-dem: A diffusion probabilistic approach to digital elevation model void filling. IEEE Geoscience and Remote Sensing Letters, 21, pp.1-5. doi:10.1109/LGRS.2024.3403835.

Piran, M.J., Wang, X., Kim, H.J. and Kwon, H.H., 2024. Precipitation nowcasting using transformer-based generative models and transfer learning for improved disaster preparedness. International Journal of Applied Earth Observation and Geoinformation, 132, p.103962. doi:10.1016/j.jag.2024.103962.

Rawat, P.S. and Barthwal, A., 2024. LANDSLIDE MONITOR: a real-time landslide monitoring system. Environmental Earth Sciences, 83(8), p.226. doi:10.1007/s12665-024-11526-0.

Ren, X., Liu, W., Yang, W., Mao, L. and Li, H., 2025. Landslide Deformation Uncertainty Quantification Using Conformalized Graph Neural Networks: A Case Study in Sichuan Province, China. IEEE Access. doi:10.1109/ACCESS.2025.3568273.

Shakeel, A., Walters, R.J., Ebmeier, S.K. and Al Moubayed, N., 2022. ALADDIn: Autoencoder-LSTM-based anomaly detector of deformation in InSAR. IEEE Transactions on Geoscience and Remote Sensing, 60, pp.1-12. doi:10.1109/TGRS.2022.3169455.

Xiao, Y., Ju, N., He, C., Xiao, Z., & Ma, Z. (2022). Week-ahead shallow landslide displacement prediction using chaotic models and robust LSTM. Frontiers in Earth Science, 10, 965071. doi:10.3389/feart.2022.965071.

Wang, K., Wei, B., Zhao, T., Wu, G., Zhang, J., Zhu, L. and Wang, L., 2024. An Automated Approach for Mapping Mining-Induced Fissures Using CNNs and UAS Photogrammetry. Remote Sensing, 16(12), p.2090. doi:10.3390/rs16122090.

Wang, Y., Fang, Z., Wang, M., Peng, L., Hong, H., 2020b. Comparative study of landslide susceptibility mapping with different recurrent neural networks. Computers and Geosciences 138, 1-18. doi:10.1016/j.cageo.2020.104445.

Whitaker, T., 2023. LSTM-GAN for Enhanced Anomaly Detection in Time Series Data. Journal of Computer Technology and Software, 2(2).

Wu, L., Zhou, J.T., Zhang, H., Wang, S.R., Ma, T., Yan, H. and Li, S.H., 2024. Time series analysis and gated recurrent neural network model for predicting landslide displacements. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 18(1), pp.172-185. doi:10.1080/17499518.2022.2138918.

Thi Ngo, P., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., Lee, S., 2021. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of iran. Geoscience Frontiers 12, 505-519. doi:10.1016/j.gsf.2020.06.013.

Yadav, R., Nascetti, A., Azizpour, H. and Ban, Y., 2024. Unsupervised flood detection on SAR time series using variational autoencoder. International Journal of Applied Earth Observation and Geoinformation, 126, p.103635. doi:10.1016/j.jag.2023.103635.

Yang, H., Liu, Y., Han, Q., Xu, L., Zhang, T., Wang, Z., Yan, A., Zhao, S., Han, J. and Wang, Y., 2025. Improved Landslide Deformation Prediction Using Convolutional Neural Network–Gated Recurrent Unit and Spatial–Temporal Data. Remote Sensing, 17(4), p.727. doi:10.3390/rs17040727.

Zamanzadeh Darban, Z., Webb, G.I., Pan, S., Aggarwal, C. and Salehi, M., 2024. Deep learning for time series anomaly detection: A survey. ACM Computing Surveys, 57(1), pp.1-42. doi:10.1145/3691338.

Zhang, D., Yang, J., Li, F., Han, S., Qin, L. and Li, Q., 2022. Landslide risk prediction model using an attention-based temporal convolutional network connected to a recurrent neural network. IEEE Access, 10, pp.37635-37645. doi:10.1109/ACCESS.2022.3165051.

Zhang, Q., He, Y., Zhang, Y., Lu, J., Zhang, L., Huo, T., Tang, J., Fang, Y. and Zhang, Y., 2024. A Graph-Transformer method for landslide susceptibility mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. doi:10.1109/JSTARS.2024.3437751.

Zhao, J., Yuan, Y., Dong, Y., Li, Y., Shao, C. and Yang, H., 2024. Void filling of digital elevation models based on terrain feature-guided diffusion model. Remote Sensing of Environment, 315, p.114432. doi:10.1016/j.rse.2024.114432.

Zhao, Z., Chen, T., Dou, J., Liu, G. and Plaza, A., 2024. Landslide susceptibility mapping considering landslide local-global features based on CNN and transformer. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, pp.7475-7489. doi:10.1109/JSTARS.2024.3379350.

Comment # 4:

4. Introduction: The section lacks sufficient citations, and the objectives are unclear. I recommend rephrasing them for clarity.

5. Line 28. What do you mean by relativity?

6. Line 33. What do you mean by potentials? Do you mean driving factors?

7. Line 57. What do you mean by remainder? Do you mean the structure of this paper? You don't need to mention it.

Response:

- We sincerely appreciate you for your valuable comments and constructive suggestions, which have greatly helped us improve the quality and clarity of the manuscript. We have carefully revised the paper according to each comment, and all modifications have been incorporated into the revised version. The following section provides our detailed, point-by-point responses to your remarks.
-
- *** Regarding the general comments on the Introduction**
- **4. Introduction: The section lacks sufficient citations, and the objectives are unclear. I recommend rephrasing them for clarity.**
- We fully agree with your comments. In the revised Introduction, we have implemented the following comprehensive revisions:

- (1) We have added essential references to key statements to provide a more robust academic background and stronger scholarly support.
- (2) We have thoroughly rewritten the section describing the research objectives to make them more specific and clearly defined. The revised list of objectives now explicitly outlines the four focal aspects addressed in this review—data, models, applications, and challenges with future directions—thereby eliminating the ambiguity present in the previous version.
-
- *** Regarding the comments on specific terminology and expressions**
- **5. Line 28. What do you mean by relativity?**
- We apologize for the lack of clarity in our original phrasing. The term "relativity" was intended to convey that the assessment of landslide potential is not absolute but is comparative and context-dependent. It refers to the relative likelihood, spatial probability, or comparative susceptibility of a landslide occurring in one area versus another, based on a set of conditioning factors (e.g., slope, geology, land use).
- We have revised the manuscript to improve precision. The phrase has been replaced with "the inherent uncertainty and dynamic nature" to better convey that landslide prediction is not absolute but is a probabilistic assessment fraught with challenges.
- **6. Line 33. What do you mean by potentials? Do you mean driving factors?**
- Thank you for pointing out the ambiguity here. We regret that the expression caused confusion.
- In Line 33, we used the term "potentials" to refer to the **landslide potential** or **the likelihood of a slope failure**. This potential is inherently dynamic, as it may vary over time due to external factors, and it does not denote the driving factors themselves.
- To clarify, we have revised the sentence to use the more precise term "**landslide potential**" (Line 33). This change better reflects our intended meaning that the probability of landslide occurrence is not static.
- **7. Line 57. What do you mean by remainder? Do you mean the structure of this paper? You don't need to mention i**t.
- Thank you for pointing this out. We agree that mentioning "the remainder" is unnecessary. We have removed the final sentence describing the structure of the paper.
-
- Once again, we would like to express our sincere appreciation to the reviewer for all the valuable comments and suggestions that have helped us to improve our manuscript.

Original Description in Introduction

1 Introduction

Landslides are geological hazards induced by either natural forces or human activities, typically involving the interplay of various factors such as geology, meteorology, hydrology, and topography. Every year, landslides cause significant global losses, particularly in regions with heavy rainfall, frequent earthquakes, and complex geological conditions, representing a major threat to human life, property, and infrastructure.

According to data released by the United Nations International Strategy for Disaster Reduction (UNISDR), more than 1,000 landslide-related disaster events occur annually, causing thousands of fatalities and substantial economic losses. As global climate change progresses, the frequency of extreme weather events increases, leading to a growing risk of landslides.

Potential landslides refer to slopes prone to instability that may fail and trigger disasters within a certain time frame. Potential landslides represent the precursor stage of landslide occurrence (Lin et al., 2024). If potential landslides are not identified and addressed promptly, the slope may eventually become unstable and develop into a landslide due to changes in internal stress conditions and external triggering factors.

Due to the relativity and dynamic nature of potential landslides, the identification work becomes extremely complicated. On the one hand, it is not possible to determine that a landslide will definitely occur just because there are signs of deformation on the slope. Multiple factors need to be comprehensively considered to assess the possibility of its instability. On the other hand, the uncertainty of external factors increases the difficulty of judgment. Sudden events such as heavy rainfall and earthquakes may instantly change the stress state of the slope and trigger signs of deformation. Given the dynamic characteristics of potentials, it is also essential to conduct long-term monitoring of the landslides with potential hazards after identification.

Conventional methods for landslide identification and monitoring, such as field surveys, geological analysis, and radar interferometry, can identify potential landslide areas to a certain extent. However, these methods often have problems such as high costs, significant time consumption, and difficulties in data collection, and their applications are limited in extensive areas. In addition, conventional machine learning requires tedious feature selection and lacks autonomy in feature extraction. As a result, it is difficult for these traditional methods to extract available information from big data and they are unable to represent complex monitoring processes (Sheng et al., 2023). For the above reasons, how to effectively identify and monitor areas with potential landslides has become an important topic in the current prevention and control of geological hazards.

Over the past few years, deep learning has stood out in the application of landslide hazards (Aslam et al., 2021; Nava et al., 2023; Wang et al., 2023a; Zhou et al., 2023). Deep learning is a branch of machine learning, consisting of consecutive operations (Janiesch et al., 2021). These operations gradually extract complex features by using the results of previous operations as inputs. Through the training of large-scale and multi-source data, deep learning models are able to automatically extract features, capture complex nonlinear relationships, and conduct pattern recognition in high-dimensional data, which shows great potential in the identification of potential landslides (Nava et al., 2021; Yang et al., 2024c).

In this review, we aim to summarize the applications of deep learning in the field of potential landslide identification, including data, models, applications, challenges, and future directions.

(1) We classify commonly used heterogeneous data into three categories for research. These data sources offer comprehensive data support for the application of deep learning in potential landslide identification.

(2) We introduce the roles of commonly used deep learning models in potential landslide identification, and compare the advantages and disadvantages among different models.

(3) We analyze the performance of deep learning models in different scenarios through case studies, discussing the adaptability of deep learning in potential landslide identification.

(4) We summarize the main challenges currently faced by the application of deep learning in potential landslide identification, and highlight new opportunities and promising future directions.

The remainder of this paper is organized as follows. Section 2 introduces seven main data sources. Section 3 summarizes five roles of deep learning models in potential landslide identification. Section 4 investigates the application of deep learning models in four typical landslides and provides a comprehensive summary. Section 5 analyzes the current challenges in potential landslide identification. Section 6 discusses future research directions. Section 7 provides the concluding remarks.

Revised Description in Introduction

1 Introduction

Landslides are complex geological hazards triggered by both natural processes and human activities, involving intricate interactions among geological, hydrological, topographic, and meteorological factors (Fidan et al., 2024). Globally, landslides cause significant loss of life and property each year, particularly in mountainous areas with intense rainfall, seismic activity, and fragile geological conditions (Askarinejad et al., 2018; Ehsan et al., 2025; Marín-Rodríguez et al., 2024). According to the United Nations Office for Disaster Risk Reduction (UNDRR, 2023), more than 1,000 landslide-related disasters occur annually, resulting in thousands of fatalities and substantial economic damage. With the intensification of climate change, extreme weather events are becoming more frequent, further increasing global landslide risks (Wang et al., 2023).

Potential landslides refer to slopes that exhibit early signs of instability and may evolve into landslides under external triggers such as rainfall or earthquakes. They represent the precursor stage of landslide development (Lin et al., 2024). Timely identification and monitoring of such slopes are crucial for disaster prevention and risk mitigation (Strząbała et al., 2024; Xie et al. 2020).

However, the **inherent uncertainty** and dynamic nature of potential landslides make their identification challenging. On the one hand, it is not possible to determine that a landslide will definitely occur just because there are signs of deformation on the slope (Peres and Cancelliere 2014; Zhang et al., 2019). Multiple factors need to be comprehensively considered to assess the possibility of its instability. On the other hand, the uncertainty of external factors increases the difficulty of judgment. Sudden events such as heavy rainfall and earthquakes may instantly change the stress state of the slope and trigger signs of deformation (Yang et al., 2024). Given the dynamic characteristics of **potential landslides**, it is also essential to conduct long-term monitoring of the landslides with potential hazards after identification (Lakhote et al., 2025).

Conventional approaches to potential landslide identification, including field surveys, geological analysis, and interferometric radar techniques, have contributed substantially to hazard assessment but remain costly, time-consuming, and limited in spatial coverage (Akosah et al., 2024; Zhao and Lu 2018). Machine learning has partially improved efficiency but still depends heavily on manual feature engineering, requiring expert knowledge to design relevant predictors (Sheng et al., 2023). These limitations restrict the scalability and adaptability of conventional approaches in complex geospatial environments.

In contrast, deep learning provides an effective data-driven alternative for landslide research. As a subfield of machine learning, deep learning performs hierarchical feature extraction through multiple nonlinear transformations (Janiesch et al., 2021; Nava et al., 2023). By leveraging large-scale, multi-source data, deep learning models can automatically extract representative features, capture nonlinear dependencies, and conduct pattern recognition in high-dimensional datasets (Aslam et al., 2021; Wang et al., 2023a; Zhou et al., 2023). These capabilities make DL particularly suitable for identifying and characterizing potential landslides across diverse spatial and temporal scales (Nava et al., 2021; Yang et al., 2024c).

In this review, we aim to summarize the applications of deep learning in the field of potential landslide identification, including data, models, applications, challenges, and future directions.

(1) We classify commonly used heterogeneous data into three categories for research. These data sources offer comprehensive data support for the application of deep learning in potential landslide identification.

(2) We introduce the roles of commonly used deep learning models in potential landslide identification, and compare the advantages and disadvantages among different models.

(3) We analyze the performance of deep learning models in different scenarios through case studies, discussing the adaptability of deep learning in potential landslide identification.

(4) We summarize the main challenges currently faced by the application of deep learning in potential landslide identification, and highlight new opportunities and promising future directions.

Newly Added References

Akosah, S., Gratchev, I., Kim, D.H. and Ohn, S.Y., 2024. Application of artificial intelligence and remote sensing for landslide detection and prediction: systematic review. Remote Sensing, 16(16), p.2947. doi:10.3390/rs16162947.

Askarinejad, A., Akca, D. and Springman, S.M., 2018. Precursors of instability in a natural slope due to rainfall: a full-scale experiment. Landslides, 15(9), pp.1745-1759. doi:10.1007/s10346-018-0994-0.

Ehsan, M., Anees, M.T., Bakar, A.F.B.A. and Ahmed, A., 2025. A review of geological and triggering factors influencing landslide susceptibility: artificial intelligence-based trends in mapping and prediction. International Journal of Environmental Science and Technology, pp.1-36. doi:10.1007/s13762-025-06741-6.

Fidan, S., Tanyaş, H., Akbaş, A., Lombardo, L., Petley, D.N. and Görüm, T., 2024. Understanding fatal landslides at global scales: a summary of topographic, climatic, and anthropogenic perspectives. Natural Hazards, 120(7), pp.6437-6455. doi:10.1007/s11069-024-06487-3.

Lakhote, A., Chan, Y.C., Lu, C.Y., Kumar, G. and Sun, C.W., 2025. Monitoring slow-moving deep-seated landslide using PSI technique: a case study of a potential sliding slope from southern Taiwan. Landslides, 22(5), pp.1677-1692. doi:10.1007/s10346-024-02453-z.

Marín-Rodríguez, N.J., Vega, J., Zanabria, O.B., González-Ruiz, J.D. and Botero, S., 2024. Towards an understanding of landslide risk assessment and its economic losses: a scientometric analysis. Landslides, 21(8), pp.1865-1881. doi:10.1007/s10346-024-02272-2.

Peres, D.J. and Cancelliere, A., 2014. Derivation and evaluation of landslide-triggering thresholds by a Monte Carlo approach. Hydrology and Earth System Sciences, 18(12), pp.4913-4931. doi:10.5194/hess-18-4913-2014.

Strząbała, K., Ćwiąkała, P. and Puniach, E., 2024. Identification of landslide precursors for early warning of hazards with remote sensing. Remote Sensing, 16(15), p.2781. doi:10.3390/rs16152781.

Wang, X., Wang, Y., Lin, Q. and Yang, X., 2023. Assessing global landslide casualty risk under moderate climate change based on multiple GCM projections. International Journal of Disaster Risk Science, 14(5), pp.751-767. doi:10.1007/s13753-023-00514-w.

Yang, W., Liu, L. and Shi, P., 2020. Detecting precursors of an imminent landslide along the Jinsha River. Natural Hazards and Earth System Sciences, 20(11), pp.3215-3224. doi:10.5194/nhess-20-3215-2020.

Yang, W., Zhang, Y., Zhang, L., Bai, G., Wan, B. and An, N., 2024. Comprehensive study on the stability and failure mechanism of landslides under rainfall and earthquake in northwest mountainous areas. Frontiers in Earth Science, 12, p.1470083. doi:10.3389/feart.2024.1470083.

Zhang, J., van Westen, C.J., Tanyas, H., Mavrouli, O., Ge, Y., Bajrachary, S., Gurung, D.R., Dhital, M.R. and Khanal, N.R., 2019. How size and trigger matter: analyzing rainfall-and earthquake-triggered landslide inventories and their causal relation in the Koshi River basin, central Himalaya. Natural hazards and earth system sciences, 19(8), pp.1789-1805. doi:10.5194/nhess-19-1789-2019.

Zhao, C. and Lu, Z., 2018. Remote sensing of landslides—A review. Remote Sensing, 10(2), p.279. doi:10.3390/rs10020279.

Comment #5:

8. Line 61. Chapter 2. It is unclear whether the data sources mentioned are actually used in deep learning for landslide detection, or whether they could be used. If they are used, please provide specific examples.

Response:

- We sincerely appreciate your time and effort in reviewing our manuscript! The comments are highly insightful and have been instrumental in refining our paper. We fully agree that the manuscript should explicitly clarify that these data sources have indeed been applied in existing studies, rather than merely describing their principles and characteristics. Following your suggestion, we have thoroughly revised **Section 2**.
- The major revisions are as follows:
- (1) **Structural adjustment:** We have closely linked each data category with its specific applications in deep learning, providing published literature to substantiate each type.
- (2) **Addition of specific cases:** For each data source (e.g., optical remote sensing, InSAR, and LiDAR), we now include real-world examples that have been successfully applied in deep learning-based landslide detection studies, with corresponding references. These cases clearly demonstrate how specific data types are integrated with particular deep learning models.

Original Description in Section 2

2 Deep Learning for Potential Landslide Identification: Data Source

Accurate identification of potential landslides is the primary step in effectively preventing and mitigating the impacts of landslide hazards. Data sources are the cornerstone of achieving this objective. Different types of data provide indispensable information for potential landslide identification from various perspectives, and drive ongoing advancements in related research and practices.

In potential landslide identification, the richness and reliability of data sources directly determine the accuracy and effectiveness of research. Data sources not only provide fundamental information to outline the landslide environments, but also enable dynamic monitoring and precise analysis. This section will comprehensively review the critical roles played by three main types of data sources: satellite observation data, airborne remote sensing data, and ground-based observation data (see Fig. 1)

2.1 Satellite Observation Data

Since the launch of Landsat-1, the first earth observation satellite for studying and monitoring the Earth's surface on July 23, 1972, satellite data has become widely accessible, extending beyond single-purpose analyses or results (Wulder et al., 2022). With the continuous development of satellite observation, its immense potential for application in landslide research has become evident (Liu et al., 2021d). Currently, satellite observation data primarily refers to data obtained through space-borne synthetic aperture radar (SAR) and optical remote sensing.

2.1.1 Space-borne SAR

SAR is an active microwave remote sensing system (Franceschetti and Lanari, 2018). It is not only capable of acquiring data on demand by actively emitting microwave signals but also facilitates partial penetration of vegetation cover through its longer wavelength bands (such as the L-band), thereby allowing the retrieval of surface deformation information beneath vegetated areas. The time series data provided by SAR can serve as input for deep learning models, allowing these models to be trained to identify long-term patterns of terrain change. Continuous monitoring of potential landslide areas is crucial, and SAR is widely employed in high-risk environments.

Interferometric synthetic aperture radar (InSAR) has been developed based on. It obtains surface elevation information by performing coherent processing on two sets of SAR images observed in the same area (Dai et al., 2022; Ma et al., 2023b; Zeng et al., 2024).

In contrast, SAR mainly provide backscatter information of ground objects. Although some features of ground objects can be identified according to the scattering characteristics, their ability to obtain topographic elevation information is relatively weak. InSAR, on the other hand, can directly generate topographic elevation data, which is of great significance for analyzing the topography and geomorphology in the identification of potential landslides, and determining key elements such as the topographic undulation and slope of potential landslide areas.

When screening for potential landslides over a large area, InSAR has higher efficiency (Dun et al., 2021; Tang et al., 2025; Zhang et al., 2021). When monitoring large potential landslide areas such as mountainous regions, InSAR can quickly obtain topographic deformation information over a large area, promptly detect potential areas with potential landslides, and reduce the workload and blind spots of manual inspections.

At present, InSAR is widely employed to generate ground deformation velocity maps and time-series data, which reveal the dynamic evolution of landslide-prone areas.

Differential interferometric synthetic aperture radar (D-InSAR) is an advancement of InSAR that eliminates topographic phase through differential processing, focusing specifically on deformation information extraction (Shen et al., 2022). The emergence of D-InSAR not only enables the transition from mixed deformation-topography signals to pure deformation signal extraction but also extends its applicability from detecting discrete deformation events to identifying slow-moving landslide processes, significantly enhancing the reliability of landslide monitoring.

2.1.2 Optical Remote Sensing

Optical remote sensing refers to the acquisition of surface information through sensors that measure reflected solar radiation. Its application in geological hazard investigations dates back to the 1970s.

Optical remote sensing offers high resolution, currently capable of achieving spatial resolutions as fine as 0.3 meters or better. In potential landslide identification, it not only facilitates the retrieval of detailed surface textures and color characteristics using rich spectral data but also enables the direct identification of morphological features and object contours via visual interpretation of imagery (Cheng and Han, 2016; Li et al., 2022b).

Landslide formation typically follows a progressive process from deformation to failure, accompanied by precursor indicators such as tensile cracks, stepped scarps, and localized collapses. These indicators exhibit distinct spectral signatures in optical imagery compared to their surroundings, enabling both manual interpretation and automated detection.

In vegetated mountainous regions, surface vegetation often undergoes detectable changes before a landslide event. Optical remote sensing leverages multispectral data, particularly red and near-infrared bands, to monitor vegetation health and identify potential landslide zones. Furthermore, the calculation of the normalized difference vegetation index (NDVI) facilitates the evaluation of vegetation health in potential landslide regions, providing critical insights into potential landslide precursors (Verrelst et al., 2015).

2.2 Airborne Remote Sensing Data

Airborne remote sensing data, typically acquired by manned aircrafts, provide high-resolution imagery of localized areas. Advanced airborne platforms equipped with oblique photogrammetry and, more recently, close-range photogrammetry technologies enable millimeter-level accuracy in 3D photogrammetry, facilitating the observation of subtle surface deformations, rock mass structures, and the construction of highly detailed 3D models of terrain and above-ground infrastructure (Macciotta and Hendry, 2021; Xu et al., 2023). Among these technologies, airborne photogrammetry and airborne radar are the most commonly used.

2.2.1 Airborne Light Detection and Ranging (LiDAR)

LiDAR has been used for landslide and other geological hazard investigations in many regions since the late 1990s. As an active remote sensing system, LiDAR can laterally scan a range of 60◦ and capture 400,000 points per second, enabling largescale 3D scanning of terrain, structures, and vegetation within a short period (Mallet and Bretar, 2009). It offers centimeter-level accuracy in both horizontal and vertical dimensions.

Airborne LiDAR is irreplaceable in capturing 3D details and penetrating vegetation, particularly in densely vegetated areas where conventional aerial photography faces significant limitations. Airborne LiDAR not only acquires high- resolution digital surface models (DSMs) from laser point cloud data but also generates high-accuracy DEMs by removing vegetation contributions (Fang et al., 2022; Jaboyedoff et al., 2012; Yan et al., 2023), thereby revealing concealed hazard features such as mountain fractures, loose deposits, and landslide masses under vegetation cover.

Point cloud data obtained from airborne LiDAR can monitor dynamic changes in mountainous terrain by detecting deformations such as subsidence, displacement, and uplift, while also facilitating the construction of 3D landslide models to simulate sliding directions and impact areas. Through intuitive visualization of slope morphology and structure from multiple perspectives, LiDAR enables researchers to conduct a comprehensive assessment of slope conditions and identify subtle hazard features that may not be easily discernible in 2D imagery.

2.2.2 Unmanned Aerial Vehicle (UAV)

UAV aerial photogrammetry provides outstanding maneuverability and high-precision measurements. Traversing over steep slopes and valleys, UAVs are able to monitor areas that are often inaccessible to satellites and manned aerial platforms (Niethammer et al., 2012), thus addressing critical observational limitations.

In large-scale and topographically complex regions, UAVs can perform efficient aerial inspections, overcoming the limitations of ground-based inspections in inaccessible or visually obstructed regions. By rapidly scanning mountain slopes, embankments, and gullies, UAVs provide a comprehensive understanding of the geological conditions and enable timely identification of macro-scale geomorphic anomalies. However, given cost-effectiveness constraints, UAVs are currently more commonly used for periodic and continuous monitoring in localized areas. They are particularly well-suited for rapid and dynamic monitoring of landslides in high-priority zones.

With the rapid advancement of UAVs, centimeter-level vertical and oblique aerial photogrammetry is now achievable (Fan et al., 2020). The high-definition cameras mounted on UAVs are able to capture the subtle cracks on the surface of the mountain.

These cracks may be early signs of a landslide (Sun et al., 2024a). By conducting a comparative analysis of the images taken at different times, the development and changes of the cracks can be monitored, including the increase in the length, width and depth of the cracks, as well as the changes in the crack orientation. In some mountainous areas or valleys, there may be a large number of loose accumulations. These accumulations may trigger landslides under specific conditions. Aerial photography by UAVs can clearly identify information such as the distribution range, accumulation quantity and accumulation shape of these loose accumulations, and assess their potential threats to the surrounding environment.

When equipped with LiDAR sensors, UAVs can effectively remove vegetation from the data. Then, assisting researchers to reveal landslide boundaries, crack patterns, and other deformation features hidden beneath vegetation cover. This integrated approach combines the strengths of photogrammetry and LiDAR, allowing for rapid deployment and targeted area monitoring while mitigating the challenges posed by vegetation cover in landslide detection and assessment.

After extreme weather events such as heavy rainstorms or geological events like earthquakes occur, the stability of the mountain may be affected, making it prone to triggering geological hazards. UAVs even can quickly conduct aerial monitoring of the relevant areas after extreme weather.

2.3 Ground-based Observation Data

Satellite observation and airborne remote sensing are mainly employed for identifying potential landslides based on surface morphology. However, due to the influence of various factors, the identification results may not always be fully accurate, leading to potential misjudgments. Therefore, the potential landslide points identified through remote sensing still necessitate field investigations by researchers for verification, differentiation, confirmation, or exclusion of hazards. In some cases, additional on-site observation and monitoring methods are needed for accurate assessment. Commonly used ground-based monitoring methods include ground-based SAR, 3D laser scanners and various sensor devices deployed or installed on the ground.

2.3.1 Ground-based Synthetic Aperture Radar (GB-SAR)

GB-SAR is an active ground-based microwave remote sensing system that has been developed over the past decade. Compared to spaceborne SAR, GB-SAR allows adjustment of radar wave incidence angles and azimuths, preventing phase decorrelation issues caused by terrain obstructions in satellite SAR, making it particularly suitable for monitoring steep slopes, canyons, and other areas with limited satellite line-of-sight (Noferini et al., 2007).

GB-SAR effectively integrates the principles of SAR imaging with electromagnetic wave interferometry. By leveraging precise measurements of sensor system parameters, attitude parameters, and geometric relationships between orbits, GB-SAR quantifies spatial positions and subtle changes at specific surface points, allowing for the measurement of surface deformations with millimeter or even sub-millimeter precision. 180

During landslide movement, the ground experiences noticeable subsidence, displacement, or cracking. GB-SAR can be configured for high-resolution, continuous observation to capture instantaneous deformations during the landslide creep phase and generate corresponding displacement maps (Liu et al., 2021a; Xiao et al., 2021). This capability facilitates the distinction between evolutionary stages of landslides and further analysis of the dynamics of landslide activity.

For small-scale regional monitoring, GB-SAR can establish customized geometric configurations specifically designed for target areas. Utilizing mobile rail systems or multi-antenna setups, GB-SAR reconstructs 3D deformation vector fields of landslide masses, identifying sliding directions and potential failure surfaces.

2.3.2 Terrestrial Laser Scanning (TLS)

TLS emerged in the mid-1990s. It plays a unique role in local refined monitoring by emitting laser pulses and measuring their reflection time (Stumvoll et al., 2021; Teza et al., 2007).

The landslide often manifests as a sharp change in the ground surface. TLS can provide data with sufficient accuracy, assisting researchers in identifying the features of these landslides. By combining topographic analysis, the location of the landslide surface can be accurately determined. TLS scanner can also help identify the landslide mass, that is, the flow path of the landslide materials. Through analyzing the point cloud data, the movement path of the landslide area, the soil accumulation area, and the accumulation location of the landslide materials can be extracted, providing detailed information for the analysis and assessment of potential landslides.

By quickly and massively collecting spatial point position information, TLS can automatically splice and rapidly obtain the appearance of the measured object. It can be used to construct high-precision surface models and appearance models of buildings and structures. The 3D model can display the shape and structure of the mountain and the detailed features of the ground surface from different angles and in all directions (Zhou et al., 2024), enabling geological experts and engineers to have a more intuitive understanding of the overall situation of the landslide area. For example, the cracks in the mountain, the loose accumulations, and the degree of weathering of the rocks can be clearly seen, providing richer information for the identification of potential landslide hazards.

Currently, TLS is commonly used in critical areas requiring localized precision. For historical landslide masses, it captures reactivation indicators such as rear tensile cracks and frontal bulging, with data input into anomaly detection models to identify reactivation signals.

2.3.3 Ground-based Sensor Devices

Compared to the aforementioned data sources, ground-based sensors offer key advantages, including high precision, realtime capabilities, and multi-parameter fusion (Dai et al., 2023). They can address the limitations of remote sensing and provide critical ground-based dynamic information for potential landslide identification.

Ground-based sensing devices are highly diverse, and the data they acquire directly reflect the state of landslide masses. These datasets provide foundational inputs for deep learning models, enabling multi-dimensional analysis and interpretation of potential landslide conditions. For example, ground sensors (e.g., GNSS receivers and crack meters) can collect parameters like displacement and tilt angle at frequencies ranging from minutes to seconds, capturing transient, anomalous signals just prior to landslide events, thereby filling the temporal resolution gap in remote sensing (Jiang et al., 2022). By integrating time series data with SAR imagery, deep learning models can be trained to uncover correlation patterns between surface deformations and subsurface parameters. Instruments such as piezometers and soil pressure gauges can directly monitor key parameters like pore water pressure and soil stress on the sliding surface. By combining the obtained subsurface data with geomechanical equations, the position of the sliding surface or geotechnical strength parameters can be inverted.

Revised Description in Section 2

2 Deep Learning for Potential Landslide Identification: Data Source

Accurate identification of potential landslides is the primary step in effectively preventing and mitigating the impacts of landslide hazards. Data sources are the cornerstone of achieving this objective. Different types of data provide indispensable information for potential landslide identification from various perspectives, and drive ongoing advancements in related research and practices.

In potential landslide identification, the richness and reliability of data sources directly determine the accuracy and effectiveness of research. Data sources not only provide fundamental information to outline the landslide environments, but also enable dynamic monitoring and precise analysis. This section will comprehensively review the critical roles played by three main types of data sources: satellite observation data, airborne remote sensing data, and ground-based observation data (see Fig. 1)

2.1 Satellite Observation Data

Since the launch of Landsat-1, the first earth observation satellite for studying and monitoring the Earth's surface on July 23, 1972, satellite data has become widely accessible, extending beyond single-purpose analyses or results (Wulder et al., 2022). With the continuous development of satellite observation, its immense potential for application in landslide research has become evident (Liu et al., 2021d). At present, satellite observation data mainly include space-borne synthetic aperture radar

(SAR) and optical remote sensing data, both of which are widely used as inputs for deep learning models in landslide identification.

2.1.1 Space-borne SAR

SAR is an active microwave remote sensing system (Franceschetti and Lanari, 2018). It is not only capable of acquiring data on demand by actively emitting microwave signals but also facilitates partial penetration of vegetation cover through its longer wavelength bands (such as the L-band), thereby allowing the retrieval of surface deformation information beneath vegetated areas. The time series data provided by SAR can serve as input for deep learning models, allowing these models to be trained to identify long-term patterns of terrain change. Continuous monitoring of potential landslide areas is crucial, and SAR is widely employed in high-risk environments.

Interferometric synthetic aperture radar (InSAR) has been developed based on. It obtains surface elevation information by performing coherent processing on two sets of SAR images observed in the same area (Dai et al., 2022; Ma et al., 2023b; Zeng et al., 2024).

In contrast, SAR mainly provide backscatter information of ground objects. Although some features of ground objects can be identified according to the scattering characteristics, their ability to obtain topographic elevation information is relatively weak. InSAR, on the other hand, can directly generate topographic elevation data, which is of great significance for analyzing the topography and geomorphology in the identification of potential landslides, and determining key elements such as the topographic undulation and slope of potential landslide areas.

When screening for potential landslides over a large area, InSAR has higher efficiency (Dun et al., 2021; Tang et al., 2025; Zhang et al., 2021). When monitoring large potential landslide areas such as mountainous regions, InSAR can quickly obtain topographic deformation information over a large area, promptly detect potential areas with potential landslides, and reduce the workload and blind spots of manual inspections. InSAR is widely employed to generate ground deformation velocity maps and time-series data, which reveal the dynamic evolution of landslide-prone areas.

Recent studies have integrated InSAR-derived deformation velocity fields with deep learning models to automatically detect slow-moving or latent landslides. For example, Liu et al. (2022) employed an InSAR-CNN framework to map active landslides in the Eastern Tibet Plateau area, achieving a detection accuracy of over 90%. Similarly, Zhang et al. (2022) proposed a two-stage detection deep learning network (InSARNet) for detecting anomalous deformation areas in Maoxian County, Sichuan Province, with a recognition accuracy of 93.88%. Targeting the complex deformation mechanisms of multi-type landslides in Zigui County, Three Gorges Reservoir Area, Hu et al (2025) used InSAR time-series displacement as the core data, develop a deep learning architecture based on the integrated framework of EMD and GRU, break through the limitations of traditional models such as single-type, single-target, and low-accuracy, and achieve dual-accurate prediction of displacement and failure time for multi-type landslides.

Differential interferometric synthetic aperture radar (D-InSAR) is an advancement of InSAR that eliminates topographic phase through differential processing, focusing specifically on deformation information extraction (Shen et al., 2022). The emergence of D-InSAR not only enables the transition from mixed deformation-topography signals to pure deformation signal extraction but also extends its applicability from detecting discrete deformation events to identifying slow-moving landslide processes, significantly enhancing the reliability of landslide monitoring (Zhong et al., 2024).

2.1.2 Optical Remote Sensing

Optical remote sensing refers to the acquisition of surface information through sensors that measure reflected solar radiation. Its application in geological hazard investigations dates back to the 1970s.

Optical remote sensing offers high resolution, currently capable of achieving spatial resolutions as fine as 0.3 meters or better. In potential landslide identification, it not only facilitates the retrieval of detailed surface textures and color characteristics using rich spectral data but also enables the direct identification of morphological features and object contours via visual interpretation of imagery (Cheng and Han, 2016; Li et al., 2022b).

Landslide formation typically follows a progressive process from deformation to failure, accompanied by precursor indicators such as tensile cracks, stepped scarps, and localized collapses. These indicators exhibit distinct spectral signatures in optical imagery compared to their surroundings, enabling both manual interpretation and automated detection. In deep learning applications, multispectral optical images have been widely used to train CNN-based models for landslide identification. Lu et al. (2023) developed a method for achieving accurate landslide mapping using medium-resolution remote sensing images and DEM data, which has the potential for deployment in large-scale landslide detection. Jiang et al. (2022) proposed a TL-Mask R-CNN for identifying a small number of old landslide samples in the area along the Sichuan-Tibet Transportation Corridor. The results show that the pixel accuracy of segmentation for new landslides and old landslides can reach 87.71% and 75.86% respectively.

In vegetated mountainous regions, surface vegetation often undergoes detectable changes before a landslide event. Optical remote sensing leverages multispectral data, particularly red and near-infrared bands, to monitor vegetation health and identify potential landslide zones. Furthermore, the calculation of the normalized difference vegetation index (NDVI) facilitates the evaluation of vegetation health in potential landslide regions, providing critical insights into potential landslide precursors (Verrelst et al., 2015).

While both space-borne SAR and optical remote sensing are pivotal for large-area landslide screening, they offer complementary capabilities and have distinct limitations. Optical remote sensing provides intuitive visual interpretation of geomorphological features but is rendered useless by cloud cover and darkness. In contrast, space-borne SAR, with its all-weather, day-and-night imaging capability, excels in detecting millimeter-to-centimeter-scale surface deformation through InSAR techniques, which is a direct precursor to landslide failure. However, InSAR performance can be degraded in heavily vegetated areas due to temporal decorrelation and in steep terrain due to geometric distortions (Lin et al., 2022; Yan et al., 2024), areas where optical stereo imaging for DEM generation might be less affected. Therefore, the integration of SAR-derived deformation maps and optical-based geomorphological maps is considered a best practice for regional-scale landslide inventory mapping and preliminary hazard assessment (Xun et al., 2022).

2.2 Airborne Remote Sensing Data

Airborne remote sensing data, typically acquired by manned aircrafts, provide high-resolution imagery of localized areas. Advanced airborne platforms equipped with oblique photogrammetry and, more recently, close-range photogrammetry technologies enable millimeter-level accuracy in 3D photogrammetry, facilitating the observation of subtle surface deformations, rock mass structures, and the construction of highly detailed 3D models of terrain and above-ground infrastructure (Macciotta and Hendry, 2021; Xu et al., 2023). Among these technologies, airborne photogrammetry and airborne radar are the most commonly used.

2.2.1 Airborne Light Detection and Ranging (LiDAR)

LiDAR has been used for landslide and other geological hazard investigations in many regions since the late 1990s. As an active remote sensing system, LiDAR can laterally scan a range of 60◦ and capture 400,000 points per second, enabling largescale 3D scanning of terrain, structures, and vegetation within a short period (Mallet and Bretar, 2009). It offers centimeter-level accuracy in both horizontal and vertical dimensions.

Airborne LiDAR is irreplaceable in capturing 3D details and penetrating vegetation, particularly in densely vegetated areas where conventional aerial photography faces significant limitations. Airborne LiDAR not only acquires high- resolution digital surface models (DSMs) from laser point cloud data but also generates high-accuracy DEMs by removing vegetation contributions (Fang et al., 2022; Jaboyedoff et al., 2012; Yan et al., 2023), thereby revealing concealed hazard features such as mountain fractures, loose deposits, and landslide masses under vegetation cover.

Point cloud data obtained from airborne LiDAR can monitor dynamic changes in mountainous terrain by detecting deformations such as subsidence, displacement, and uplift, while also facilitating the construction of 3D landslide models to simulate sliding directions and impact areas. Through intuitive visualization of slope morphology and structure from multiple perspectives, LiDAR enables researchers to conduct a comprehensive assessment of slope conditions and identify subtle hazard features that may not be easily discernible in 2D imagery.

These high-precision DEMs and point clouds serve as critical inputs for deep learning models. For instance, Wei et al. (2023) proposed the DAG-Net model to construct dynamic edge features for enhancing point cloud representations, achieving the highest mean Intersection over Union (mIoU) of 0.743 and an F1-score of 0.786. Similarly, Farmakis et al. (2022) Based on the advanced PointNet and PointNet++ architectures, we developed deep neural networks for 3D point cloud learning. The best-performing model achieved accuracies of approximately 89% and 84% during the final and shortest monitoring campaigns, respectively. These examples demonstrate that airborne LiDAR data are not only suitable but have been effectively applied in deep learning-based landslide analysis.

2.2.2 Unmanned Aerial Vehicle (UAV)

UAV aerial photogrammetry provides outstanding maneuverability and high-precision measurements. Traversing over steep slopes and valleys, UAVs are able to monitor areas that are often inaccessible to satellites and manned aerial platforms (Niethammer et al., 2012), thus addressing critical observational limitations.

In large-scale and topographically complex regions, UAVs can perform efficient aerial inspections, overcoming the limitations of ground-based inspections in inaccessible or visually obstructed regions. By rapidly scanning mountain slopes, embankments, and gullies, UAVs provide a comprehensive understanding of the geological conditions and enable timely identification of macro-scale geomorphic anomalies. However, given cost-effectiveness constraints, UAVs are currently more commonly used for periodic and continuous monitoring in localized areas. They are particularly well-suited for rapid and dynamic monitoring of landslides in high-priority zones.

With the rapid advancement of UAVs, centimeter-level vertical and oblique aerial photogrammetry is now achievable (Fan et al., 2020). The high-definition cameras mounted on UAVs are able to capture the subtle cracks on the surface of the mountain.

These cracks may be early signs of a landslide (Sun et al., 2024a). By conducting a comparative analysis of the images taken at different times, the development and changes of the cracks can be monitored, including the increase in the length, width and depth of the cracks, as well as the changes in the crack orientation. In some mountainous areas or valleys, there may be a large number of loose accumulations. These accumulations may trigger landslides under specific conditions. Aerial photography by UAVs can clearly identify information such as the distribution range, accumulation quantity and accumulation shape of these loose accumulations, and assess their potential threats to the surrounding environment. This capability is leveraged in deep learning applications, where time-series UAV imagery is processed using RNNs or 3D CNNs to monitor the spatiotemporal evolution of these cracks, providing a data-driven approach for early warning (Xu et al., 2025; Sandric et al., 2024).

When equipped with LiDAR sensors, UAVs can effectively remove vegetation from the data. Then, assisting researchers to reveal landslide boundaries, crack patterns, and other deformation features hidden beneath vegetation cover. This integrated approach combines the strengths of photogrammetry and LiDAR (Mandlburger et al., 2020), allowing for rapid deployment and targeted area monitoring while mitigating the challenges posed by vegetation cover in landslide detection and assessment. In addition, Wallace et al (2012) demonstrated that integrating LiDAR with UAVs can maintain high accuracy while reducing costs to a certain extent.

After extreme weather events such as heavy rainstorms or geological events like earthquakes occur, the stability of the mountain may be affected, making it prone to triggering geological hazards. UAVs even can quickly conduct aerial monitoring of the relevant areas after extreme weather.

2.3 Ground-based Observation Data

Satellite- and airborne-based observations primarily identify potential landslides through large-scale surface morphological analysis. However, these approaches are often affected by vegetation cover, viewing geometry, and atmospheric noise, which may lead to misclassification or omission. Therefore, ground-based observation techniques play a critical complementary role, offering higher temporal resolution, accuracy, and localized verification for potential landslide identification. In recent years, data collected from ground-based monitoring instruments have not only been used for field validation but also increasingly incorporated into deep learning frameworks to improve temporal continuity and physical interpretability in landslide detection and forecasting.

2.3.1 Ground-based Synthetic Aperture Radar (GB-SAR)

GB-SAR is an active ground-based microwave remote sensing system that has been developed over the past decade. Compared to spaceborne SAR, GB-SAR allows adjustment of radar wave incidence angles and azimuths, preventing phase decorrelation issues caused by terrain obstructions in satellite SAR, making it particularly suitable for monitoring steep slopes, canyons, and other areas with limited satellite line-of-sight (Noferini et al., 2007).

GB-SAR effectively integrates the principles of SAR imaging with electromagnetic wave interferometry. By leveraging precise measurements of sensor system parameters, attitude parameters, and geometric relationships between orbits, GB-SAR quantifies spatial positions and subtle changes at specific surface points, allowing for the measurement of surface deformations with millimeter or even sub-millimeter precision.

During landslide movement, the ground experiences noticeable subsidence, displacement, or cracking. GB-SAR can be configured for high-resolution, continuous observation to capture instantaneous deformations during the landslide creep phase and generate corresponding displacement maps (Liu et al., 2021a; Xiao et al., 2021). For example, Long et al. (2018) proposed a GBSAR persistent scatterer (PS) point selection method based on the mean coherence coefficient, amplitude dispersion index, estimated signal-to-noise ratio, and displacement accuracy index. Han et al. (2022) proposed an LSTM (long short-term memory)-based approach for processing GB-InSAR time series data. Kačan et al. (2022) employed two deep learning methods to investigate the potential and advantages of processing raw GBSAR data for automatic radar classification.

For small-scale regional monitoring, GB-SAR can establish customized geometric configurations specifically designed for target areas. Utilizing mobile rail systems or multi-antenna setups, GB-SAR reconstructs 3D deformation vector fields of landslide masses (Shi et al., 2025), identifying sliding directions and potential failure surfaces.

2.3.2 Terrestrial Laser Scanning (TLS)

TLS emerged in the mid-1990s. It plays a unique role in local refined monitoring by emitting laser pulses and measuring their reflection time (Stumvoll et al., 2021; Teza et al., 2007).

The landslide often manifests as a sharp change in the ground surface. TLS can provide data with sufficient accuracy, assisting researchers in identifying the features of these landslides. By combining topographic analysis, the location of the landslide surface can be accurately determined. TLS scanner can also help identify the landslide mass, that is, the flow path of the landslide materials. Through analyzing the point cloud data, the movement path of the landslide area, the soil accumulation area, and the accumulation location of the landslide materials can be extracted, providing detailed information for the analysis and assessment of potential landslides.

By quickly and massively collecting spatial point position information, TLS can automatically splice and rapidly obtain the appearance of the measured object. It can be used to construct high-precision surface models and appearance models of buildings and structures. The 3D model can display the shape and structure of the mountain and the detailed features of the ground surface from different angles and in all directions (Zhou et al., 2024), enabling geological experts and engineers to have a more intuitive understanding of the overall situation of the landslide area. For example, the cracks in the mountain, the loose accumulations, and the degree of weathering of the rocks can be clearly seen, providing richer information for the identification of potential landslides.

In the context of deep learning, TLS-derived 3D point clouds have become critical inputs for morphological feature extraction and automatic landslide identification. For example, Senogles et al. (2022) integrated TLS point cloud data to assess surface displacements induced by landslide movements. Wang et al. (2025) provided a practical and adaptable solution for landslide monitoring by integrating TLS point clouds with embedded RGB imagery.

These examples confirm that TLS data are not only suitable but already actively used in deep learning-based landslide recognition, providing precise geometric constraints for multi-source fusion frameworks that combine DEM, optical, and InSAR information.

2.3.3 Ground-based Sensor Devices

Compared to the above methods, ground-based sensors offer key advantages, including high precision, realtime capabilities, and multi-parameter fusion (Dai et al., 2023). They can address the limitations of remote sensing and provide critical ground-based dynamic information for potential landslide identification.

Ground-based sensing devices are highly diverse, and the data they acquire directly reflect the state of landslide masses. These datasets provide foundational inputs for deep learning models, enabling multi-dimensional analysis and interpretation of potential landslide conditions. For example, ground sensors (e.g., GNSS receivers and crack meters) can collect parameters like displacement and tilt angle at frequencies ranging from minutes to seconds, capturing transient, anomalous signals just prior to landslide events, thereby filling the temporal resolution gap in remote sensing (Jiang et al., 2022). These data are often used as input sources for RNN models and their variants (Bai et al., 2022; Wang et al.,2021). By integrating time series data with SAR imagery, deep learning models can be trained to uncover correlation patterns between surface deformations and subsurface parameters.

Instruments such as piezometers and soil pressure gauges can directly monitor key parameters like pore water pressure and soil stress on the sliding surface. By combining the obtained subsurface data with geomechanical equations, the position of the sliding surface or geotechnical strength parameters can be inverted.

Therefore, GB-SAR, TLS, and ground-based sensors are not only auxiliary observation techniques but are increasingly serving as key data sources for deep learning-driven landslide identification. Their integration into CNN, LSTM, and GAN frameworks enables high-resolution spatial-temporal modeling of slope behavior, bridging the gap between field-scale monitoring and large-scale hazard prediction.

Newly Added References

Bai, D., Lu, G., Zhu, Z., Tang, J., Fang, J. and Wen, A., 2022. Using time series analysis and dual-stage attention-based recurrent neural network to predict landslide displacement. Environmental Earth Sciences, 81(21), p.509. doi:10.1007/s12665-022-10637-w.

Farmakis, I., DiFrancesco, P.M., Hutchinson, D.J. and Vlachopoulos, N., 2022. Rockfall detection using LiDAR and deep learning. Engineering Geology, 309, p.106836. doi:10.1016/j.enggeo.2022.106836.

Han, J., Yang, H., Liu, Y., Lu, Z., Zeng, K. and Jiao, R., 2022. A deep learning application for deformation prediction from ground-based insar. Remote Sensing, 14(20), p.5067. doi:10.3390/rs14205067.

Hu, X., Sun, Z., Wang, Z., Huang, X., Zhou, M., He, S., ... & Xu, W. (2025). InSAR-based deep learning prediction model for multi-type landslides displacement and failure time in Zigui, Three Gorges Area, China. Landslides, 1-15. doi:10.1007/s10346-025-02613-9.

Jiang, W., Xi, J., Li, Z., Zang, M., Chen, B., Zhang, C., ... & Zhu, W. (2022). Deep learning for landslide detection and segmentation in high-resolution optical images along the Sichuan-Tibet transportation corridor. Remote Sensing, 14(21), 5490.doi:10.3390/rs14215490.

Kačan, M., Turčinović, F., Bojanjac, D. and Bosiljevac, M., 2022. Deep learning approach for object classification on raw and reconstructed gbsar data. Remote sensing, 14(22), p.5673. doi:10.3390/rs14225673.

Lin, Y.N., Chen, Y.C., Kuo, Y.T. and Chao, W.A., 2022. Performance study of landslide detection using multi-temporal SAR images. Remote Sensing, 14(10), p.2444. doi:10.3390/rs14102444.

Liu, Y., Yao, X., Gu, Z., Zhou, Z., Liu, X., Chen, X., & Wei, S. (2022). Study of the automatic recognition of landslides by using InSAR images and the improved mask R-CNN model in the Eastern Tibet Plateau. Remote Sensing, 14(14), 3362. doi:10.3390/rs14143362.

Long, S., Tong, A., Yuan, Y., Li, Z., Wu, W. and Zhu, C., 2018. New approaches to processing ground-based SAR (GBSAR) data for deformation monitoring. Remote sensing, 10(12), p.1936. doi:10.3390/rs10121936.

Lu, W., Hu, Y., Zhang, Z., & Cao, W. (2023). A dual-encoder U-Net for landslide detection using Sentinel-2 and DEM data. Landslides, 20(9), 1975-1987. doi:10.1007/s10346-023-02089-5.

Mandlburger, G., Pfennigbauer, M., Schwarz, R., Flöry, S. and Nussbaumer, L., 2020. Concept and performance evaluation of a novel UAV-borne topo-bathymetric LiDAR sensor. Remote Sensing, 12(6), p.986. doi:10.3390/rs12060986.

Sandric, I., Chitu, Z., Ilinca, V. and Irimia, R., 2024. Using high-resolution UAV imagery and artificial intelligence to detect and map landslide cracks automatically. Landslides, 21(10), pp.2535-2543. doi:10.1007/s10346-024-02295-9.

Senogles, A., Olsen, M.J. and Leshchinsky, B., 2022. SlideSim: 3D landslide displacement monitoring through a physics-based simulation approach to self-supervised learning. Remote Sensing, 14(11), p.2644. doi:10.3390/rs14112644.

Shi, X., Zhao, Z., Dai, Y., Dai, K. and Ju, A., 2025. Post-Disaster High-Frequency Ground-Based InSAR Monitoring and 3D Deformation Reconstruction of Large Landslides Using MIMO Radar. Remote Sensing, 17(18), p.3183. doi:10.3390/rs17183183.

Wallace, L., Lucieer, A., Watson, C. and Turner, D., 2012. Development of a UAV-LiDAR system with application to forest inventory. Remote sensing, 4(6), pp.1519-1543. doi:10.3390/rs4061519.

Wang, J., Nie, G., Gao, S., Wu, S., Li, H. and Ren, X., 2021. Landslide deformation prediction based on a GNSS time series analysis and recurrent neural network model. Remote Sensing, 13(6), p.1055. doi:10.3390/rs13061055.

Wang, Z., Butt, J.A., Huang, S., Medic, T. and Wieser, A., 2025. Dense 3D Displacement Estimation for Landslide Monitoring via Fusion of TLS Point Clouds and Embedded RGB Images. arXiv preprint arXiv:2506.16265. doi:10.48550/arXiv.2506.16265.

Wei, R., Ye, C., Ge, Y., Li, Y. and Li, J., 2023. Dynamic graph attention networks for point cloud landslide segmentation. International Journal of Applied Earth Observation and Geoinformation, 124, p.103542. doi:10.1016/j.jag.2023.103542.

Xu, H., Wang, L., Shu, B., Zhang, Q. and Li, X., 2025. Automatic Detection of Landslide Surface Cracks from UAV Images Using Improved U-Network. Remote Sensing, 17(13), p.2150. doi:10.3390/rs17132150.

Xun, Z., Zhao, C., Kang, Y., Liu, X., Liu, Y. and Du, C., 2022. Automatic extraction of potential landslides by integrating an optical remote sensing image with an InSAR-derived deformation map. Remote Sensing, 14(11), p.2669. doi:10.3390/rs14112669.

Yan, L., Xiong, Q., Li, D., Cheon, E., She, X. and Yang, S., 2024. InSAR-Driven Dynamic Landslide Hazard Mapping in Highly Vegetated Area. Remote Sensing, 16(17), p.3229. doi:10.3390/rs16173229.

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

Zhong, J., Li, Q., Zhang, J., Luo, P., & Zhu, W. (2024). Risk assessment of geological landslide hazards using D-InSAR and remote sensing. Remote Sensing, 16(2), 345. doi:10.3390/rs16020345.

Comment #6:

9. Line 84. The phrase stops in the middle of the sentence.

Response:

- Thank you for catching this incomplete sentence! We have revised it to be grammatically correct and clearer.
- Please see revised **Subsection 2.1.1** for details.

Original Description in Subsection 2.1.1

……

Interferometric synthetic aperture radar (InSAR) has been developed based on. It obtains surface elevation information by performing coherent processing on two sets of SAR images observed in the same area (Dai et al., 2022; Ma et al., 2023b; Zeng et al., 2024).

……

Revised Description in Subsection 2.1.1

……

Interferometric synthetic aperture radar (InSAR) has been developed based on the principle of measuring phase differences between two or more SAR images of the same area (Dai et al., 2022; Ma et al., 2023b; Zeng et al., 2024). By coherently processing these images, InSAR obtains high-precision surface elevation information and can be further applied to detect ground deformation.

……

Comment #7:

8. Line 219-227. Theis read as introduction.

Response:

- Thank you for this astute observation! We have revised it to be grammatically correct and clearer.
- To address this, we have restructured and rewritten the opening paragraph of Section 3 to make it more concise and focused on the roles and mechanisms of deep learning models in potential landslide identification. The revised text now removes redundant background discussion (e.g., traditional methods and feature extraction challenges) and directly introduces the types of deep learning models and their applications to landslide analysis. (Please see revised **Section 3** for details).
- The modification improves the logical flow and ensures that Section 3 begins with a clear technical overview consistent with the reviewer's suggestion.

Original Description in Section 3

3 Deep Learning for Potential Landslide Identification: Models

Potential landslide identification relies heavily on extensive data analysis, and the key is how to efficiently and accurately extract features that are helpful for identifying landslide occurrences. Conventional landslide identification methods often rely on human expertise or rules, often necessitating expert knowledge for identifying relevant features. With continuous exploration, deep learning, through its powerful feature learning capabilities, enables the automatic extraction of meaningful features from raw data, significantly reducing manual intervention. Especially when dealing with high-dimensional and complex landslide data, deep learning models can extract deep features related to landslides from raw data in a data-driven manner, without the need for manual feature design.

The choice of deep learning models typically depends on the type of data and the task requirements. Although each model typically has multiple effects, its internal architecture results in different focal points when it comes to automated feature extraction. This section analyzes several commonly used deep learning models from five perspectives: image analysis and processing, time series analysis, data generation, data cleaning, and data fusion.

……

Revised Description in Section 3

3 Deep Learning for Potential Landslide Identification: Models

The effectiveness of deep learning in potential landslide identification largely depends on selecting an appropriate model architecture suited to the data type and specific task. While all deep learning models excel at automated feature extraction, their internal architectures predispose them to excel in different aspects of the overall workflow. Therefore, this section does not merely list models, but organizes them based on their primary function in the landslide identification pipeline. We analyze several commonly used deep learning models by categorizing them into five functional roles: image analysis and processing, time series analysis, data generation, anomaly detection, and data fusion.

……

Comment #8:

11. Line 280. The citation refers to medical research. While cross-disciplinary examples can be useful, this seems out of scope in the current context.

Response:

- Thank you for your valuable comment!
- We agree that citing only the original U-Net paper from the medical imaging domain may appear out of scope. Our intention was to acknowledge the seminal work of Ronneberger et al. (2015), which first introduced the U-shaped encoder–decoder architecture. To improve the relevance, we have revised the text to emphasize the subsequent adoption of U-Net in geoscience and remote sensing applications. In particular, we have added domain-specific references that demonstrate the application of U-Net in landslide detection and related remote sensing tasks. (Please see revised **Subsection 3.1** for details).

Original Description in Subsection 3.1

……

U-Net is a typical example (Ronneberger et al., 2015), which features a U-shaped architecture. U-Net employs an encoder-decoder structure, where the encoder is similar to conventional CNNs, progressively reducing image resolution and extracting features through convolution and pooling operations; the decoder then restores the image resolution through transposed convolution or upsampling operations (Dong et al., 2022; Nava et al., 2022). Skip connections bridge low-level detail features with deep semantic features, thereby refining segmentation precision.

……

Revised Description in Subsection 3.1

……

U-Net is a typical example, which features a U-shaped architecture (Ronneberger et al., 2015). U-Net's encoder-decoder structure with skip connections has become a benchmark for landslide segmentation (Chandra et al., 2023; Chen et al., 2022; Meena et al., 2022). For example, Dong et al. (2022) proposed a new model, L-UNet, based on the U-Net architecture and successfully applied it to landslide extraction from remote sensing imagery. Nava et al. (2022) applied the Attention U-Net to Sentinel-1 SAR data for rapid mapping of earthquake-induced landslides, demonstrating the effectiveness of U-Net variants in pixel-level segmentation of landslide bodies under cloud-covered or topographically complex conditions.

……

Newly Added References

Chandra, N., Sawant, S. and Vaidya, H., 2023. An efficient u-net model for improved landslide detection from satellite images. PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 91(1), pp.13-28. doi:10.1007/s41064-023-00232-4.

Chen, X., Yao, X., Zhou, Z., Liu, Y., Yao, C. and Ren, K., 2022. DRs-UNet: A deep semantic segmentation network for the recognition of active landslides from InSAR imagery in the three rivers region of the Qinghai–Tibet Plateau. Remote Sensing, 14(8), p.1848. doi:10.3390/rs14081848.

Meena, S.R., Soares, L.P., Grohmann, C.H., Van Westen, C., Bhuyan, K., Singh, R.P., Floris, M. and Catani, F., 2022. Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides, 19(5), pp.1209-1229. doi:10.1007/s10346-022-01861-3.

Comment #9:

14. Line 434. Although: this expects something negative after.

15. Line 436: widely applied: you need to give reference on what they are applied to.

Response:

- Thank you for these critical observations! We agree that the original use of "Although" created an illogical sentence structure, and the claim that diffusion models have not been "widely applied" lacked necessary references, weakening our argument.
- We have completely rewritten the **Subsection 3.3** to simultaneously resolve both issues. The revision achieves the following:
- (1) We replaced "Although" with "Diffusion models" as the sentence subject and restructured the passage. The revised version clearly presents the advantages of diffusion models first, followed by

their current limitations and potential for future application.

- (2) We have added recent studies demonstrating the use of diffusion models in related geospatial and remote sensing fields, such as high-resolution satellite image synthesis, cloud removal, and topographic data reconstruction. These examples illustrate that, while diffusion models have shown promising performance in image generation and enhancement, their direct application to landslide identification is still in the exploratory phase.

Original Description in Subsection 3.3

……

Although diffusion models demonstrate strong capabilities in generating high-quality images and handling noise, they generate superior-quality data and ensure greater training stability compared to GANs and VAEs. However, diffusion models have not yet been widely applied directly to the identification of potential landslides and remain in the exploratory stage (see Fig. 4). We believe that as generative models advance in the field of geospatial remote sensing, they hold vast potential for application and could play a pivotal role in future landslide risk analysis and monitoring systems.

……

Revised Description in Subsection 3.3

……

Diffusion models demonstrate strong capabilities in generating high-quality images and handling noise (Liu et al., 2024). They produce superior-quality data and ensure greater training stability compared to GANs and VAEs. In recent years, diffusion models have been successfully applied to a variety of geospatial tasks, including remote sensing image super-resolution (Sui et al., 2024; Xiao et al., 2023), cloud removal and denoising (Leher et al., 2025; Zou et al., 2024), and terrain surface reconstruction from sparse LiDAR data (Zou et al., 2024). However, diffusion models have not yet been widely applied directly to the identification of potential landslides and remain in the exploratory stage (see Fig. 4). We believe that as generative models continue to advance in the field of geospatial remote sensing, they hold vast potential for application and could play a pivotal role in future landslide risk analysis and monitoring systems.

……

Newly Added References

Leher, Q. O., Bezerra, E. S., Paixão, T., Palomino-Quispe, F., & Alvarez, A. B. (2025). Denoising Diffusion Probabilistic Models for Cloud Removal and Land Surface Temperature Retrieval From a Single Sample. IEEE Access. doi:10.1109/ACCESS.2025.3542014.

Liu, Y., Yue, J., Xia, S., Ghamisi, P., Xie, W., & Fang, L. (2024). Diffusion models meet remote sensing: Principles, methods, and perspectives. IEEE Transactions on Geoscience and Remote Sensing. doi:10.1109/TGRS.2024.3464685.

Sui, J., Ma, X., Zhang, X., Pun, M. O., & Wu, H. (2024). Adaptive semantic-enhanced denoising diffusion probabilistic model for remote sensing image super-resolution. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. doi:10.1109/JSTARS.2024.3504569.

Wang, Z., Li, D., Wu, Y., He, T., Bian, J., & Jiang, R. (2024). Diffusion models in 3d vision: A survey. arXiv preprint arXiv:2410.04738. doi:10.48550/arXiv.2410.04738.

Xiao, Y., Yuan, Q., Jiang, K., He, J., Jin, X., & Zhang, L. (2023). EDiffSR: An efficient diffusion probabilistic model for remote sensing image super-resolution. IEEE Transactions on Geoscience and Remote Sensing, 62, 1-14. doi:10.1109/TGRS.2023.3341437.

Zou, X., Li, K., Xing, J., Zhang, Y., Wang, S., Jin, L., & Tao, P. (2024). Diffcr: A fast conditional diffusion framework for cloud removal from optical satellite images. IEEE Transactions on Geoscience and Remote Sensing, 62, 1-14. doi:10.1109/TGRS.2024.3365806.

Comment #10:

16. Line 510: missing reference

Response:

- Thank you for the valuable suggestion! We agree that this sentence requires supporting references. In the revised manuscript, we have added recent studies that demonstrate the use of GNNs, particularly GCNs and GATs, for analyzing the spatial dependencies and weighted contributions of different geo-environmental factors in landslide detection. These works provide empirical evidence for the statement.

Original Description in Subsection 3.5

……

According to the differences in message passing and aggregation methods, GNNs have derived various variants. For example, graph convolutional network (GCN) is generated by generalizing the convolutional operation to graph-structured data (Sharma et al., 2022; Wang et al., 2020a), and graph attention network (GAT) dynamically weights the importance of neighboring nodes by introducing the attention mechanism (Yuan et al., 2022; Zhou and Li, 2021). The emergence of these new architectures makes GNN variants more targeted than conventional GNNs and suitable for modeling heterogeneous relationships. Currently, they are often used for weighted analysis of the impacts of different geographical factors on landslides.

……

Revised Description in Subsection 3.5

……

According to the differences in message passing and aggregation methods, GNNs have derived various variants. For example, graph convolutional network (GCN) is generated by generalizing the convolutional operation to graph-structured data (Sharma et al., 2022; Wang et al., 2020a), and graph attention network (GAT) dynamically weights the importance of neighboring nodes by introducing the attention mechanism (Yuan et al., 2022; Zhou and Li, 2021). The emergence of these new architectures makes GNN variants more targeted than conventional GNNs and suitable for modeling heterogeneous relationships. Currently, they are often used for weighted analysis of the impacts of different geographical factors on landslides (Kuang et al., 2022; Li et al., 2025; Zhang et al., 2024).

……

Newly Added References

Kuang, P., Li, R., Huang, Y., Wu, J., Luo, X., & Zhou, F. 2022. Landslide displacement prediction via attentive graph neural network. Remote Sensing, 14(8), 1919. doi:10.3390/rs14081919.

Li, Y., Chen, T., Lv, L., Niu, R., & Plaza, A. 2025. IED-GCN: An Internal and External Decoupled Graph Convolutional Network for Landslide Susceptibility Assessment. IEEE Transactions on Geoscience and Remote Sensing. doi: 10.1109/TGRS.2025.3595205.

Zhang, Q., He, Y., Zhang, Y., Lu, J., Zhang, L., Huo, T., ... & Zhang, Y. 2024. A Graph-Transformer method for landslide susceptibility mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. doi:10.1109/JSTARS.2024.3437751.

Comment #11:

17. Line 579. Thus?

Response:

- Thank you very much for raising this concern! The term "Thus" was indeed imprecise in indicating the logical relationship between the preceding discussion on rainfall thresholds and the subsequent sentence about monitoring systems. To improve clarity and logical coherence, we have revised this sentence to better reflect the causal connection. Please see the revised **Subsection 4.1** for details.

Original Description in Subsection 4.1

……

What determines the critical threshold for rainfall-induced landslides? First, it is essential to define the critical threshold as the minimum amount of rainfall required to trigger a landslide under specific geological and topographic conditions (Naidu et al., 2018; Segoni et al., 2018b). This threshold is typically classified into two types: empirical thresholds, which are derived from statistical relationships between historical landslide events and rainfall data, and physically based thresholds, which incorporate hydromechanical models. Both approaches assume rainfall as the primary destabilizing driver. Monitoring systems thus integrate rain gauge and remote sensing to assess proximity to critical saturation thresholds (Li et al., 2023a; Piciullo et al., 2018). Moreover, the relationship between rainfall and landslides is often nonlinear and influenced by multiple factors. Deep learning models enable data-driven determination of context-specific critical rainfall values across diverse geological and topographical settings (Sala et al., 2021; Segoni et al., 2018a). For example, Badakhshan et al. (2025) incorporated the role of soil strength. Soares et al. (2022) utilized the U-Net model, reveals that the inclusion of a normalized vegetation index layer enhances model balance and significantly improves segmentation accuracy.

……

Revised Description in Subsection 4.1

……

What determines the critical threshold for rainfall-induced landslides? First, it is essential to define the critical threshold as the minimum amount of rainfall required to trigger a landslide under specific geological and topographic conditions (Naidu et al., 2018; Segoni et al., 2018b). This threshold is typically classified into two types: empirical thresholds, which are derived from statistical relationships between historical landslide events and rainfall data, and physically based thresholds, which incorporate hydromechanical models. Both approaches assume rainfall as the primary destabilizing driver. To operationalize these thresholds for landslide prediction, monitoring systems integrate rain gauge and remote sensing to assess proximity to critical saturation levels (Li et al., 2023a; Piciullo et al., 2018). Moreover, the relationship between rainfall and landslides is often nonlinear and influenced by multiple factors. Deep learning models enable data-driven determination of context-specific critical rainfall values across diverse geological and topographical settings (Sala et al., 2021; Segoni et al., 2018a). For example, Badakhshan et al. (2025) incorporated the role of soil strength. Soares et al. (2022) utilized the U-Net model, reveals that the inclusion of a normalized vegetation index layer enhances model balance and significantly improves segmentation accuracy.

......

Comment #12:

18. Line 612. Missing reference in the first phrase: "The Newmark model is….."

Response:

- Thank you for pointing out that this part of our content lacks authoritative citations! In the revised version, we have added citations to Newmark (1965) and Jibson (2007). Newmark first proposed this model, while Jibson further extended it and applied it to the assessment of earthquake-induced landslides. These references provide a solid theoretical and methodological foundation for the related statements.

Original Description in Subsection 4.2

……

The Newmark model is a commonly used basic model in the research of earthquake-induced landslides. Based on a simplified assumption, it regards the rock and soil masses on the slope as rigid blocks. When these rigid blocks are affected by seismic vibrations, they slide on the slope surface. By calculating the cumulative downhill displacement of the rigid blocks caused by the continuous increase of seismic vibrations, the stability of the slope under the action of an earthquake is measured. In other words, the greater the cumulative downslope displacement, the more unstable the slope is during the earthquake, and the higher the likelihood of a landslide occurring. However, Newmark's model exhibits critical limitations: (1) Dependence on oversimplified soil or rock strength assumptions and (2) Inadequate integration of high-resolution seismic motion data. Deep learning models address these gaps by processing massive real-time datasets, filtering noise from obscured remote sensing imagery (Wang et al., 2024b), and fusing seismic parameters with multispectral satellite data through cross-modal architectures (Dahal et al., 2024).

……

Revised Description in Subsection 4.2

……

The Newmark model is a commonly used basic model in the research of earthquake-induced landslides (Jibson 2007; Newmark 1965). Based on a simplified assumption, it regards the rock and soil masses on the slope as rigid blocks. When these rigid blocks are affected by seismic vibrations, they slide on the slope surface. By calculating the cumulative downhill displacement of the rigid blocks caused by the continuous increase of seismic vibrations, the stability of the slope under the action of an earthquake is measured. In other words, the greater the cumulative downslope displacement, the more unstable the slope is during the earthquake, and the higher the likelihood of a landslide occurring. However, Newmark's model exhibits critical limitations: (1) Dependence on oversimplified soil or rock strength assumptions and (2) Inadequate integration of high-resolution seismic motion data. Deep learning models address these gaps by processing massive real-time datasets, filtering noise from obscured remote sensing imagery (Wang et al., 2024b), and fusing seismic parameters with multispectral satellite data through cross-modal architectures (Dahal et al., 2024).

……

Newly Added References

Jibson, R. W. 2007. Regression models for estimating coseismic landslide displacement. Engineering geology, 91(2-4), 209-218. doi: 10.1016/j.enggeo.2007.01.013.

Newmark, N. M. 1965. Effects of earthquakes on dams and embankments. Geotechnique, 15(2), 139-160. doi:10.1680/geot.1965.15.2.139.

Comment #13:

19. Line 788. Although deep leaning model. Needs reference

Response:

- Thank you for pointing this out! We have added appropriate references to support the statement that deep learning models have achieved success in landslide identification. Specifically, we now cite several representative studies demonstrating the successful application of CNNs, U-Net, and Transformer-based models for landslide mapping and detection. These studies provide empirical evidence that deep learning has significantly improved the accuracy and efficiency of landslide identification.

Original Description in Subsection 5.2

……

Although deep learning models have achieved success in landslide identification, they also have certain problems of their own. The most critical challenge is interpretability (Li et al., 2025). This means that it is difficult to explain how these models achieve these results.

……

Revised Description in Subsection 5.2

……

Although deep learning models have achieved success in landslide identification (Meena et al., 2022; Su et al., 2021; Yi et al., 2020; Zhao et al., 2024), they also have certain problems of their own. The most critical challenge is interpretability (Li et al., 2025). This means that it is difficult to explain how these models achieve these results.

……

Newly Added References

Meena, S. R., Soares, L. P., Grohmann, C. H., Van Westen, C., Bhuyan, K., Singh, R. P., ... & Catani, F. (2022). Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides, 19(5), 1209-1229. doi: 10.1007/s10346-022-01861-3.

Su, Z., Chow, J. K., Tan, P. S., Wu, J., Ho, Y. K., & Wang, Y. H. 2021. Deep convolutional neural network-based pixel-wise landslide inventory mapping. Landslides, 18(4), 1421-1443. doi:10.1007/s10346-020-01557-6.

Yi, Y., & Zhang, W. 2020. A new deep-learning-based approach for earthquake-triggered landslide detection from single-temporal RapidEye satellite imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 6166-6176. doi:10.1109/JSTARS.2020.3028855.

Zhao, Z., Chen, T., Dou, J., Liu, G., & Plaza, A. (2024). Landslide susceptibility mapping considering landslide local-global features based on CNN and transformer. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 7475-7489. doi:10.1109/JSTARS.2024.3379350.

- With our clarifications and revisions, we hope that we have addressed your concerns. Thank you so much for your kind consideration!
- Have a nice day!
- Pan Jiang & Zhengjing Ma & Gang Mei