Technical Note

Landslides
DOI 10.1007/s10346-026-02789-8
Received: 26 November 2025
Accepted: 18 May 2026

Chengyong Fang · Xuanmei Fan · Xin Wang · Kushanav Bhuyan · Xiangyang Dou · Hao Zhong · Mingyao Xia · Filippo Catani

Rapid and robust landslide mapping from optical EO imagery using a mamba-based deep learning framework

Abstract Rapid and accurate landslide mapping following major rainfall or earthquake events is essential for emergency response, hazard mitigation, and post-event management. Existing deep learning-based landslide mapping models often achieve high accuracy at the cost of increasingly complex architectures and substantial computational demands, which limit their scalability. Rapid-response mapping therefore requires models that maintain strong accuracy while enabling more computationally efficient and scalable deployment without reliance on large models or resource-intensive hardware. To address this challenge, the Landslide Mapping Network (LDM-Net) is introduced as a computationally efficient Mamba-based network enhanced with a Global Vision State Space module. Designed for very-high-resolution (VHR) optical imagery, LDM-Net captures multi-scale contextual information while maintaining favourable efficiency for large-area deployment. The model was evaluated using the newly developed Unmanned Aerial Vehicle Landslide Mapping Dataset (ULMD), derived from 0.2 m UAV imagery, together with several public benchmarks. Compared with Transformer-based baselines, LDM-Net required only ~ 25–50% of the GPU memory and nearly an order of magnitude fewer floating-point operations (FLOPs), while achieving modest but consistent gains in segmentation accuracy (approximately 2%). Additional experiments on independent rainfall- and earthquake-triggered landslide cases further demonstrated its cross-event applicability under fixed post-event deployment conditions. Overall, the results indicate that LDM-Net enables accurate post-event landslide mapping with favourable computational efficiency and scalability, supporting timely inventory generation for emergency response and subsequent hazard analysis.

Keywords Landslide mapping · Very-high-resolution landslide dataset · State space model · Inventory-based susceptibility modelling

Introduction

Rain-triggered and co-seismic landslides cause substantial damage to settlements in mountainous regions (Huang and Fan 2013), destroying houses, transportation networks, and farmlands, and generating long-term socioeconomic impacts. In addition, landslides frequently threaten lifeline infrastructure such as power transmission corridors (e.g. transmission towers and access roads), leading to cascading disruptions to emergency response and regional economies (Jin et al. 2025). Furthermore, when landslides block river channels, the subsequent breach can trigger debris flows and floods, resulting in cascading hazards especially on the Qinghai-Tibet Plateau (Fan et al. 2019; Jiang et al. 2021, 2026). Recent studies further highlight that upslope urban expansion is driving population growth on hazardous slopes, thereby increasing exposure to risk and underscoring the need for accurate landslide inventory mapping (Ferrer et al. 2024).

Traditional landslide mapping has relied on field surveys to assess slope stability and terrain morphology (Brardinoni et al. 2003; Coe et al. 2003; Zhong et al. 2020). Although effective, these methods are labour-intensive and time-consuming, and often cannot provide timely and consistent representations of rapidly evolving slope conditions; these limitations became particularly evident during the 2008 Wenchuan earthquake (Gorum et al. 2011). The rapid development of remote sensing technology considerably advanced landslide mapping by employing diverse sensor modalities, including synthetic aperture radar (SAR) (Nava et al. 2022), multispectral (Udin et al. 2019), and hyperspectral sensors (Ye et al. 2019). However, large-area landslide mapping remains constrained by operator-dependent interpretation, which is strongly influenced by the mapper's experience and the level of detail imposed by the mapping scale, as well as by the absence of standardised mapping protocols (Guzzetti et al. 2012).

With advancements in artificial intelligence and current-generation computer systems, more effective solutions for landslide mapping from remote sensing imagery are becoming increasingly accessible and reliable (Xu et al. 2023). In contrast to many unsupervised approaches (Wang et al. 2022; Zhang et al. 2023) which are often less robust for very-high-resolution (VHR) imagery and complex backgrounds, supervised methods have shown strong capability in extracting landslide features and geomorphic traces from high-resolution images. In particular, Transformer-based architectures have achieved promising results in VHR landslide mapping by exploiting global receptive fields and dynamic attention to capture long-range spatial dependencies (Tang et al. 2022; Lv et al. 2023; Huang et al. 2024a). However, further improvements are still needed for rapid-response applications, especially when large-area inference must be conducted under constrained GPU memory and latency budgets, because the computational cost and memory demand of attention-based models increase rapidly with input size (Zhang et al. 2022; Zhu et al. 2024). Recent studies have explored several complementary directions, including weakly supervised learning to reduce annotation costs, domain adaptation and transfer learning to improve cross-region generalisation, and emerging foundation or promptable segmentation models to enhance adaptability across heterogeneous image distributions (Zhou et al. 2022; Zhang et al. 2023; Kirillov et al. 2023). In

parallel, block-based, modular, and hybrid architectures have also been investigated in landslide-related modelling, including hybrid block-modular neural networks for landslide susceptibility mapping (Abbaszadeh Shahri and Maghsoudi Moud 2021) and hybrid convolutional neural network (CNN)-Transformer frameworks for image-based landslide mapping (Wu et al. 2024). Such designs are particularly useful when a task requires explicit coordination between local detail and global context or benefits from multi-branch feature fusion, although their suitability depends on the dominant practical constraint, such as representational flexibility, cross-domain transfer, or deployment scalability.

To address this scalability challenge, Mamba, a selective state-space model (SSM) architecture, was adopted because it introduces input-dependent parameters that enable linear-time sequence modelling while preserving strong representation capacity (Dao and Gu 2024; Gu and Dao 2024). Its extension to two-dimensional imagery, Vision Mamba, has shown a favourable balance between performance and computational efficiency across a range of vision tasks (Liu et al. 2024; Zhang et al. 2024; Zhu et al. 2024), such as target recognition (Wang et al. 2026; Xing et al. 2026), change detection (Cheng et al. 2024), and land cover classification (Ma et al. 2024). Given the increasing use of satellite and unmanned aerial vehicle (UAV) imagery in landslide mapping, the development of a rapid and precise delineation strategy remains essential.

In this study, Landslide Mapping Network (LDM-Net), an efficient state space model for landslide mapping, was developed as an early application of the Mamba architecture in this field. LDM-Net employs a Vision State Space (VSS) module to extract spatial landslide features from remote sensing imagery and integrates a novel Global Feed-Forward Neural Network (G-FFN) module to fuse multi-scale features. This design improves landslide boundary delineation while maintaining computational efficiency, even when training samples are limited. Such efficiency is particularly important for post-event response, where rapid mapping must be performed over large VHR mosaics under constrained GPU memory and computational budgets. The proposed approach was further evaluated using a newly compiled 0.2 m UAV-based landslide dataset together with several public datasets. Comparative analyses against representative semantic segmentation baselines showed that LDM-Net achieves a favourable accuracy-efficiency trade-off, thereby supporting rapid landslide mapping. Finally, the utility of inventories generated by LDM-Net during inference was demonstrated in two independent regions, and their implications for landslide susceptibility modelling were discussed.

Methodology

Architecture overview

Landslide mapping aims to rapidly delineate the spatial distribution of slope failures from remote-sensing imagery. Compared with generic semantic segmentation tasks, landslide scars often exhibit spectral and textural ambiguity with bare soil and riverbeds, and their irregular runout geometry leads to fragmented or uncertain boundaries. To address these challenges under rapid-response constraints, LDM-Net is proposed as an efficient encoder-decoder framework based on Mamba-style visual state-space modelling (Dao and Gu 2024; Zhu et al. 2024). Unlike generic hybrid or multi-branch combinations, the modular design of LDM-Net is introduced for a task-specific purpose: to preserve long-range contextual aggregation and multi-scale delineation quality while maintaining feasible scalability for large-tile VHR post-event inference. In this design, the modular structure is not used merely for architectural combination, but to assign complementary roles to global-context modelling and local feature refinement, thereby improving the delineation of spatially continuous yet geometrically irregular landslide bodies during large-tile inference.

As illustrated in Fig. 1, LDM-Net comprises a hierarchical encoder and a lightweight decoder. The encoder consists of four stages of G_VSS blocks (Global Vision State Space blocks). Unlike Transformer self-attention, each G_VSS stage performs global-context aggregation through a visual state-space mechanism: two-dimensional feature maps are serialised through cross-directional scanning and processed by selective state-space operators, enabling near-linear scaling of computation and GPU memory with increasing input size (see Supplementary Note S1 and Supplementary Figs. S1–S4 for verified block diagrams and operator details).

Given an input RGB tile of size $H \times W \times 3$, the encoder first applies patch embedding (patch size = 4) and then progressively downsamples the spatial resolution by a factor of two at each stage, producing multi-scale feature tensors of $H/4 \times W/4 \times C$, $H/8 \times W/8 \times 2C$, $H/16 \times W/16 \times 4C$, and $H/32 \times W/32 \times 8C$, respectively. In the present implementation, the stage depths are [2, 2, 15, 2], with an initial channel width of $C = 96$. This hierarchical representation matches the typical scale spectrum of landslides, from small scar edges to large runout bodies, whereby shallow stages preserve fine boundary cues and deeper stages encode broader geomorphic context.

The decoder aggregates the four multi-level encoder features using a lightweight SegHead (Xie et al. 2021) and progressively upsamples them to generate the final landslide mask. Compared with symmetric U-Net decoders, this head is parameter-efficient (~ 0.5 M parameters) while retaining sharp boundary delineation. Overall, LDM-Net combines multi-scale localisation with efficient global-context modelling, making it suitable for large-tile inference and rapid post-event inventory mapping.

G_VSS block

The proposed G_VSS block consists of two components: a Visual State Space (VSS) module for global-context aggregation and a Global Feed-Forward Network (G_FFN) for feature calibration (Supplementary Fig. S1). To avoid ambiguity, the main text describes the functional workflow of these components, while the Supplementary Materials provide verified block-level diagrams and implementation details (Supplementary Note S1; Supplementary Figs. S1–S4), which are explicitly referenced below to facilitate replication.

Remote-sensing images are inherently two-dimensional, and directly applying 1D sequence modelling requires an appropriate serialisation strategy. Following recent visual state-space designs (Liu et al. 2024), VSS first applies a Cross-Scan Module (CSM) to convert a 2D feature map $\mathbb{R}^{H \times W \times C}$ into multiple 1D sequences by scanning along rows/columns in four directions (Supplementary Fig. S2a). These sequences are then processed by a selective state-space operator (S6/SS2D) to propagate long-range information

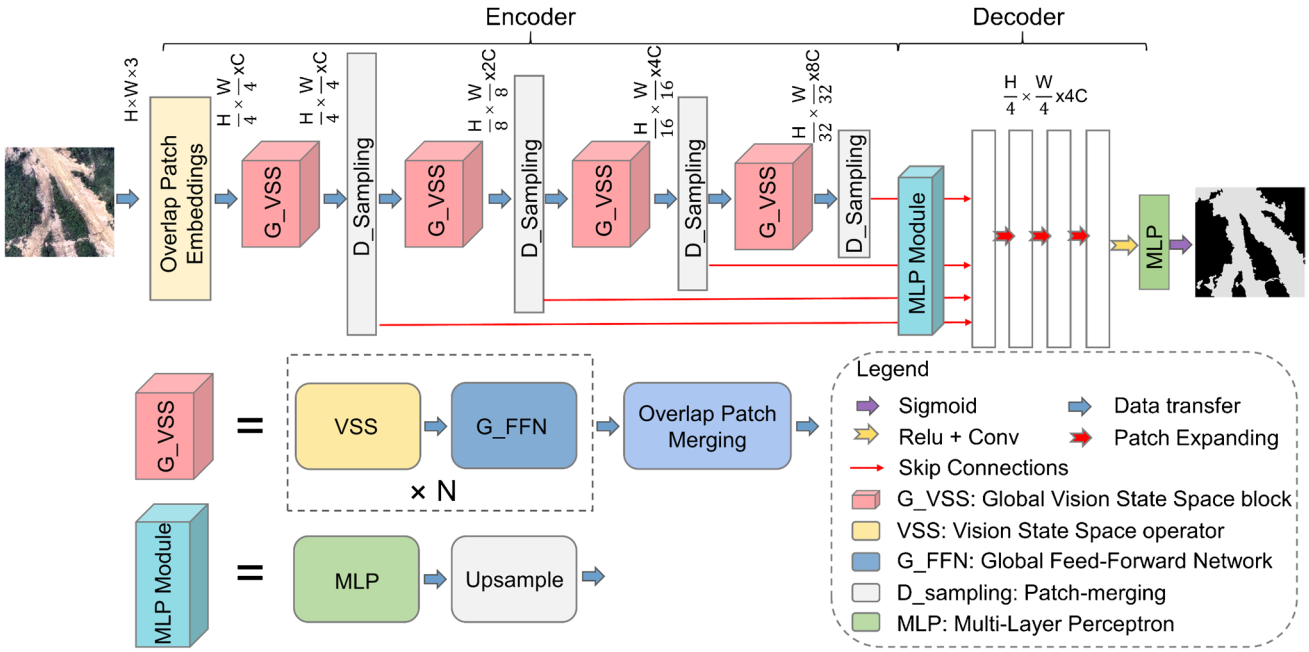

Fig. 1 Overview of the LDM-Net architecture. The encoder consists of four hierarchical stages of G_VSS blocks, with overlap patch embedding and patch merging for feature extraction and downsampling. The decoder employs an MLP module. G_VSS, Global Vision State Space block; VSS, Vision State Space operator; G_FFN, Global Feed-Forward Network; MLP, Multi-Layer Perceptron. Colours indicate different module types only, and arrows denote data transfer and feature-scale transitions

efficiently, and the outputs are merged back to a 2D representation (Supplementary Fig. S3). This design enables global information propagation over the feature map without the quadratic compute/memory growth typical of self-attention (Supplementary Fig. S2b). In practice, VSS strengthens the separability between landslides and spectrally similar backgrounds by integrating broader contextual cues while maintaining scalability for large input tiles.

After VSS, LDM-Net uses G_FFN to enhance feature discriminability by incorporating global non-linearity and stabilising high-dimensional representations. G_FFN consists of a normalisation layer, a dropout layer, and a residual calibration branch implemented by a global calibration submodule $G(\cdot)$ (Supplementary Fig. S4). Specifically, given the VSS output feature map $x \in \mathbb{R}^{H\times W\times C}$, $G(\cdot)$ first aggregates channel-wise global descriptors via the L2-norm pooling operators $g(x)$ (Eq. (1)) and then normalises each channel descriptor using $N(\|x_i\|)$ (Eq. (2)) (Woo et al. 2023). Based on the normalised descriptor, a lightweight gating function is computed using two successive 1×1 convolution layers, denoted by $C_1(\cdot) and C_2(\cdot)$ and a ReLU (Eq. (3)). Finally, the resulting calibration weights modulate the input features through a residual formulation with two learnable scalars γ and β, producing the calibrated output x_{out} (Eq. (4)). This process is formalised as follows:

$$g(x): = x \in \mathbb{R}^{H\times W\times C} \rightarrow gx = \{\|x_1\|, \|x_2\|, \cdots, \|x_c\|\} \in \mathbb{R}^C \quad (1)$$

$$N(\|x_i\|): = \|x_i\| \in R \rightarrow \frac{\|x_i\|}{\sum_{j=1,\cdots,c}\|x_j\|} \in R \quad (2)$$

$$F(x_i): = C_1(Max(0, C_2(x_i \times N(\|x_i\|)))) \quad (3)$$

$$x_{out} = \gamma \times (x_i \times F(x_i)) + \beta + x_i \quad (4)$$

where x_i is the feature map of the i-th channel and $\|\cdot\|$ is the L2-norm. γ and β are the two additional learnable parameters, which initialise to zero. $C_1(\cdot) and C_2(\cdot)$ denote two learnable 1×1 convolution layers.

Loss function and accuracy evaluation

Landslide mapping is formulated as a binary semantic segmentation task, where each pixel is classified as landslide (foreground, 1) or non-landslide (background, 0). Let $y_i = \{0, 1\}$ denote the ground-truth label of pixel i and $\hat{y}_i = [0, 1]$ the predicted landslide probability. The model is optimised using the binary cross-entropy (BCE) loss (De Boer et al. 2005):

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}[y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i)] \quad (5)$$

where N is the number of samples, y_i is the ground truth label (0 or 1), and $\hat{y}_i$ is the predicted probability of the i-th sample.

During inference, the network outputs a probability map which is binarised using a fixed threshold t (default t = 0.5) to obtain the final landslide mask. Model performance is evaluated on the held-out test set using confusion-matrix-based metrics computed at the pixel level. Importantly, only foreground (landslide) class metrics are reported; results are not averaged over landslide and background classes. This reduces the dominance of background true negatives in highly imbalanced segmentation settings, although it does not remove all interpretation issues. Specifically, Precision, Recall, F1-score (Chicco and Jurman 2020), and Intersection over

Union (IoU) (Rezatofighi et al. 2019) are reported for the landslide class only, rather than as averages over landslide and background classes:

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

$$F1 = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{8}$$

$$IoU = \frac{TP}{TP+FP+FN} \tag{9}$$

where TP, FP, TN, and FN represent the counts of true positives, false positives, true negatives, and false negatives, respectively.

For dataset-level reporting, each metric is first computed for each test image (tile), and an unweighted mean is then taken across all test images (image-wise macro average). Concretely, for a metric ∈ { Precision, Recall, F1-score, IoU}, the following is reported:

$$\overline{m} = \frac{1}{K}\sum_{k=1}^{K} m_k \tag{10}$$

where k is the number of test images and m_k is computed from pixel-wise TP, FP, FN and TN within the k-th image (If any test image contains no landslide pixels in the ground truth, foreground-only metrics may be undefined; such cases, if present, are excluded from the macro averaging.)

To complement fixed-threshold metrics, we report PR-AUC, a threshold-independent, score-based metric computed from predicted landslide probabilities. It is defined as the area under the precision-recall curve:

$$PR-AUC = \int_0^1 P(R)dR \tag{11}$$

where P and R are precision and recall, respectively. PR-AUC is computed for the foreground class and macro-averaged across test tiles.

Experimental setup and comparison baselines

LDM-Net was benchmarked against representative semantic-segmentation baselines spanning CNN, Transformer, and Mamba-style architectures (Supplementary Note S2). All models were trained and evaluated using identical data splits for each dataset and the same evaluation metrics (Section 2.3). All models were re-trained from scratch on the same splits; implementation details, hyperparameters, and code references are provided in Supplementary Note S2, whereas the experimental details of LDM-Net are provided in Supplementary Note S3. Here, "automation" is used in the narrow operational sense of deployed post-event inference from optical imagery, rather than in the broader sense of AutoML-style model search, uncertainty-aware predictive modelling, or self-updating automated workflows. Specifically, the workflow is considered automated only after the model architecture, training data split, and training procedure have been fixed and the trained model has been deployed. Under this definition, the deployed model can take post-event optical imagery as input and generate pixel-wise landslide masks independently, without manual digitisation or event-specific rule tuning. This usage does not refer to automated model selection, hyperparameter optimisation, uncertainty quantification, online self-updating, database expansion, or adaptive retraining during deployment. Accordingly, the automated component considered here is the deployed inference and mapping stage, rather than model design or model adaptation, and its practical purpose is to accelerate spatially consistent post-event landslide delineation when manual mapping is too slow for emergency-response needs.

Dataset

To rigorously evaluate the proposed method, this section introduces three publicly available landslide datasets spanning distinct spatial extents—small, medium, and large—which serve as benchmark testbeds for performance assessment. Specifically, HR-GLDD uses PlanetScope imagery at 3 m resolution with 128 × 128 tiles (approximately 0.15 km² per tile), with a landslide pixel ratio of 10.28%. GVLM provides 0.59 m bi-temporal VHR imagery cropped into 512 × 512 tiles (approximately 0.09 km² per tile), with a landslide pixel ratio of 6.72%. ULMD is a 0.2 m UAV dataset with raw 512 × 512 tiles (approximately 0.01 km² per tile), with a landslide pixel ratio of 7.16%. Notably, the Luding UAV survey covers approximately 493 km², whereas the Jiuzhaigou UAV survey covers approximately 110 km².

HR-GLDD

HR-GLDD is a global landslide dataset composed of PlanetScope imagery with a spatial resolution of 3 m, covering landslides triggered by six seismic events and six rainfall events, together with carefully curated manual annotations (Meena et al., 2023). Each sample is a 128 × 128 image tile spanning diverse geographic settings, ranging from subtropical mountainous regions and tropical hilly valleys to boreal areas, and includes abundant negative samples (e.g. bare land, vegetation, and impervious surfaces). The annotations are released as co-registered binary-mask GeoTIFFs (landslide vs. background) aligned with the imagery. Within each tile, the masks aim to capture all visually identifiable landslide pixels; however, very small or ambiguous scars close to the 3 m pixel size may be under-represented. In this study, the original training, validation, and test splits provided by the source dataset were used.

GVLM

The GVLM is a comprehensive resource spanning six continents and 17 landslide sites, encompassing a wide range of geological and climatic conditions (Zhang et al., 2023). It provides 17 pairs of bi-temporal VHR images at 0.59 m spatial resolution, together with corresponding co-registered binary-mask GeoTIFF labels for each site (landslide vs. background). Following the dataset protocol, each image was cropped into 512 × 512 tiles using a sliding window with 10% overlap, thereby capturing diverse background types such as vegetation, impervious surfaces, water/wetlands, and snow/ice, and providing abundant negative samples. For testing, 474 tiles from three events (the 2016 Kaikōura earthquake, the 2017 Kodagu rainfall event, and the 2010 Taitung typhoon) were collected, while the remaining 13 events yielded 1968 tiles, which were split 4:1 for training and validation. As with most large-area inventories, label

completeness is constrained by the mapped extent and may vary with scene conditions (e.g. shadow, snow/ice, and wetlands).

UAV landslide mapping dataset (ULMD)

ULMD is a new very-high-resolution UAV dataset comprising two comprehensive landslide inventories, which document co-seismic landslides triggered by the 2022 Luding earthquake and the 2017 Jiuzhaigou earthquake, respectively. For both inventories, the labels are released as co-registered binary-mask GeoTIFFs derived from detailed manual interpretation, thereby directly supporting pixel-wise segmentation.

On 5 September 2022, an Mw 6.6 earthquake struck Luding County, Sichuan Province, China, resulting in more than 10,000 co-seismic landslides. As part of a detailed investigation, Dai et al. (2023) conducted an extensive UAV survey covering an area of approximately 493 km² within the affected region. To create a high-quality landslide inventory, UAV imagery was integrated with pre-earthquake Google Earth imagery, enabling pixel-level landslide interpretation (Fig. S5). Negative samples, including buildings, bare land, dry riverbeds, and roads, were deliberately included to improve model robustness. The final inventory consists of 258 non-overlapping UAV tiles at a spatial resolution of 6000 × 6000 pixels. For model training, the tiles were randomly split 4:1 for training and validation and then cropped into smaller patches of 512 × 512 pixels, yielding a total of 50,810 training tiles.

On 8 August 2017, an Mw 6.4 earthquake struck Jiuzhaigou County, Sichuan Province, China, triggering more than 2500 co-seismic landslides. A targeted UAV survey by Fan et al. (2018) provided high-resolution imagery, which was subsequently refined using pre-earthquake Google Earth imagery for detailed pixel-level interpretation (Fig. S6). The Jiuzhaigou landslide inventory, segmented into 512 × 512 pixel tiles (totaling 5321 tiles), serves as an independent test set for evaluating the generalisation ability of models trained on the Luding dataset. Figure S7 illustrates representative remote-sensing tiles and corresponding landslide labels from ULMD. In these label images, black regions represent non-landslide background, whereas white regions indicate landslide features. To ensure robust model performance, the dataset includes diverse negative samples such as clouds, bare land, rivers, roads, and buildings, together with various landslide types, including avalanches (Fig. S7a), mass flows (Figs. S7b, S7c, and S7d), and slides (Fig. S7g). The inclusion of these diverse landslide and non-landslide samples improves model accuracy and generalisation.

Experiments and results

Landslide detection based on the datasets

LDM-Net was evaluated against representative baselines on HR-GLDD following the protocol described in Section 2.4, and the pixel-level results on the test set are summarised in Table 1, Supplementary Table 1, and Supplementary Table 2. LDM-Net achieved the highest F1 score (70.9%) and IoU (54.9%), indicating a favourable accuracy-efficiency trade-off for landslide pixel segmentation. Although its Recall and Precision were not individually the highest among all models, the balance between these metrics indicates robust performance on medium- to high-resolution imagery. Notably, traditional CNN architectures outperformed Transformer-based methods such as MaskFormer, SwinTrans, and SegFormer on this dataset, as the latter tended to achieve high Recall at the expense of Precision, thereby increasing false positives in spectrally ambiguous regions. By incorporating the G_FFN module, LDM-Net effectively balances this trade-off. Visual inspection (Fig. 2) further shows that LDM-Net delineates landslides accurately across diverse scenarios.

On the higher-resolution GVLM benchmark, the same evaluation protocol was applied (), and the pixel-level results on the test set are reported in Table 1, Supplementary Table 1 and Supplementary Table 2. LDM-Net attained the highest Recall (77.2%) and F1 score (81.6%), together with an IoU of 67.2%. Although its Precision (86.6%) was slightly lower than that of some traditional CNN-based methods, the overall balance of metrics indicates robust performance on complex high-resolution imagery. The limited receptive fields of CNN-based models constrained their ability to capture holistic landslide features, whereas Transformer- and Mamba-based models, benefiting from global context, delivered improved results (Fig. 3). These findings indicate that LDM-Net is particularly effective in delineating landslide features in high-resolution settings.

On ULMD, derived from UAV imagery, LDM-Net also achieved strong performance under the same evaluation protocol, as summarised in Table 1, Supplementary Table 1 and Supplementary Table 2. It achieved the highest Recall (77.2%) and a Precision of 86.3%, resulting in an F1 score of 81.5% and an IoU of 68.1%. Comparisons with other methods show that, although VMamba and SegFormer performed competitively, LDM-Net provided the most balanced overall performance across Precision, Recall, F1, and IoU. The progression from 3 m to 0.59 m and finally to 0.2 m imagery further highlights the increasing difficulty of accurately capturing fine-grained spectral and spatial details. Traditional CNN architectures struggle under these high-resolution, large-area conditions, whereas the Mamba-based design of LDM-Net effectively balances Recall and Precision, thereby supporting more accurate landslide detection and segmentation (Fig. 4).

Ablation experiments

A novel G_VSS block was introduced within the Mamba-style encoder to enhance landslide detection through efficient global-context modelling. The G_VSS block builds on a visual state-space (VSS) design with cross-directional (four-way) scanning and further incorporates the proposed G_FFN module to strengthen the representation of large-scale landslide structures such as scarp and runout. Ablation experiments on ULMD using 512 × 512 inputs (Table 2), evaluated with the same pixel-level protocol as in Table 1, show that although the G_FFN module achieves only a slight (0.3%) increase in Precision compared to a standard FFN module, it improves Recall by 2.4%, resulting in gains of 1.5% in F1, a 0.8% rise in landslide IoU. Despite a 12% increase in model parameters, G_FFN reduces encoder FLOPs by 54% and GPU memory usage by 10% for a single forward pass under a 512 × 512 input setting (Table 2; "; Supplementary Note S3), making it especially suitable for VHR landslide detection where efficiency is important.

Feature heatmaps (Fig. S8) clearly demonstrate that the G_FFN better concentrates salient landslide features while reducing information redundancy, thereby mitigating issues such as feature

Table 1 Accuracy of landslide detection by LDM-Net and comparative methods on HR-GLDD, GVLM, and ULMD. Metrics are computed for the landslide (foreground) class only at the pixel level. The best scores are highlighted in bold

Dataset	Methods	Recall (%)	Precision (%)	F1 (%)	IoU (%)	PR-AUC	Reference
HR-GLDD	FCN	72.0	66.1	68.9	52.5	0.704	Long et al. 2015
	U-Net	72.6	67.4	69.9	53.8	0.723	Ronneberger et al. 2015
	ResU-Net	72.9	67.0	69.8	53.7	0.721	Diakogiannis et al. 2020
	PSP-Net	73.2	65.6	69.2	52.9	0.706	Zhao et al. 2017
	Deeplabv3+	74.7	61.9	68.4	51.9	0.700	Chen et al. 2018
	LinkNet	74.0	65.0	69.2	52.9	0.706	Chaurasia and Culurciello 2017
	UperNet	74.6	62.6	68.1	51.6	0.696	Xiao et al. 2018
	HRNet	70.1	**70.2**	70.1	53.9	0.724	Wang et al. 2021
	MaskFormer	**74.6**	63.5	68.6	52.2	0.703	Cheng et al. 2021
	SwinTrans	74.1	64.4	68.9	52.1	0.704	Liu et al. 2021
	SegFormer	73.8	65.6	69.5	53.3	0.715	Xie et al. 2021
	VMamba	73.7	67.8	70.6	54.6	0.729	Liu et al. 2024
	LDM-Net	74.5	67.6	**70.9**	**54.9**	**0.733**	**This study**
GVLM	FCN	53.2	89.2	66.6	50.1	0.637	Long et al. 2015
	U-Net	54.0	89.0	67.2	50.6	0.646	Ronneberger et al. 2015
	ResU-Net	54.4	88.8	67.5	51.0	0.648	Diakogiannis et al. 2020
	PSP-Net	57.6	89.8	70.2	53.9	0.691	Zhao et al. 2017
	Deeplabv3+	63.0	**91.8**	74.6	59.9	0.728	Chen et al. 2018
	LinkNet	62.2	91.5	74.0	58.9	0.724	Chaurasia and Culurciello 2017
	UperNet	65.6	87.0	74.8	60.1	0.731	Xiao et al. 2018
	HRNet	66.1	86.9	75.1	60.3	0.734	Wang et al. 2021
	MaskFormer	70.1	86.0	77.2	62.9	0.750	Cheng et al. 2021
	SwinTrans	71.9	86.1	78.4	64.2	0.762	Liu et al. 2021
	SegFormer	76.4	83.7	79.9	66.3	0.776	Xie et al. 2021
	VMamba	74.4	87.0	80.2	66.5	0.781	Liu et al. 2024
	LDM-Net	**77.2**	86.6	**81.6**	**67.2**	**0.795**	**This study**

Table 1 (continued)

Dataset	Methods	Recall (%)	Precision (%)	F1 (%)	IoU (%)	PR-AUC	Reference
ULMD	FCN	51.0	81.3	62.6	45.5	0.733	Long et al. 2015
	U-Net	50.8	81.8	62.7	45.5	0.733	Ronneberger et al. 2015
	ResU-Net	51.3	81.9	63.0	46.0	0.735	Diakogiannis et al. 2020
	PSP-Net	64.8	72.8	68.6	52.1	0.761	Zhao et al. 2017
	Deeplabv3+	71.9	82.4	76.8	62.3	0.797	Chen et al. 2018
	LinkNet	73.2	72.9	73.1	57.5	0.780	Chaurasia and Culurciello 2017
	UperNet	71.2	79.2	74.9	60.2	0.788	Xiao et al. 2018
	HRNet	71.7	81.7	76.3	61.6	0.795	Wang et al. 2021
	MaskFormer	76.1	77.7	76.9	62.4	0.798	Cheng et al. 2021
	SwinTrans	76.6	78.2	77.4	65.5	0.801	Liu et al. 2021
	SegFormer	77.0	81.5	79.2	66.7	0.809	Xie et al. 2021
	VMamba	74.8	86.0	80.0	67.3	0.813	Liu et al. 2024
	LDM-Net	**77.2**	**86.3**	**81.5**	**68.1**	**0.820**	**This study**

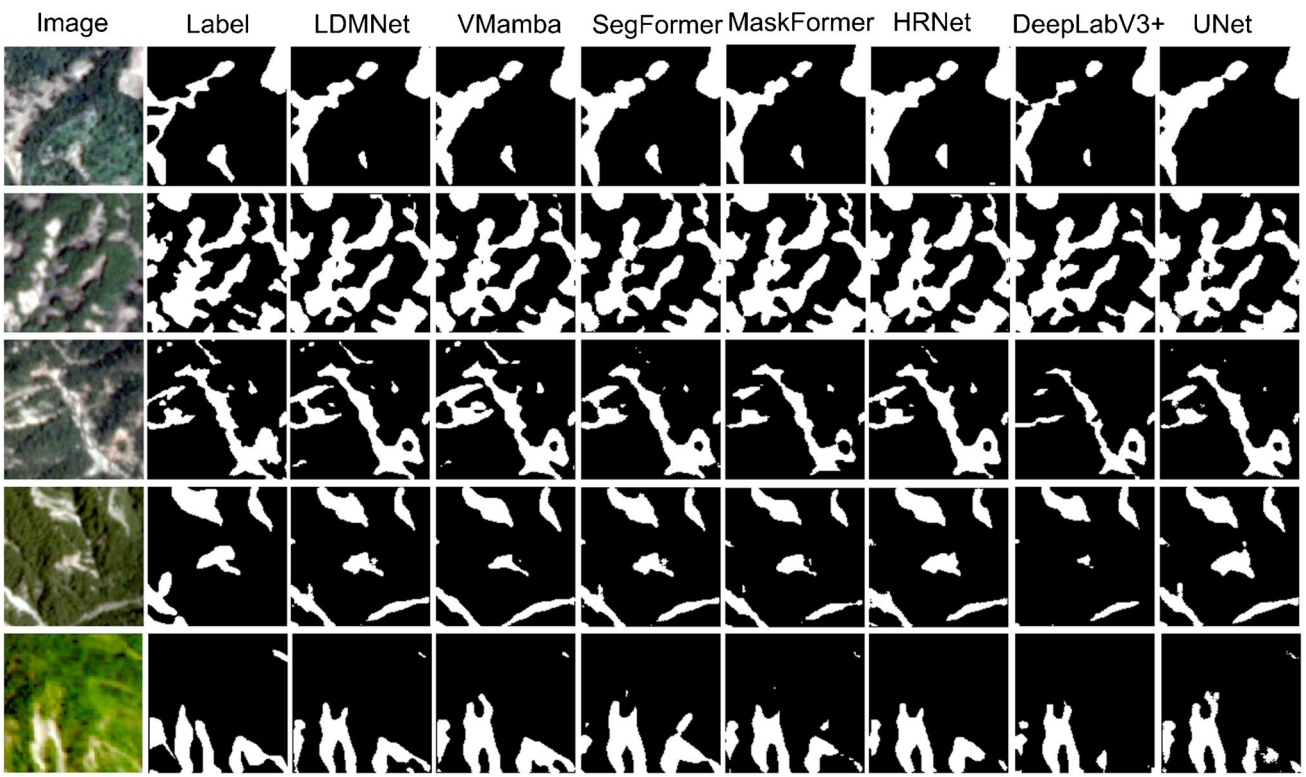

Fig. 2 Detection results of five landslide samples in the HR-GLDD test set using LDM-Net and comparison algorithms

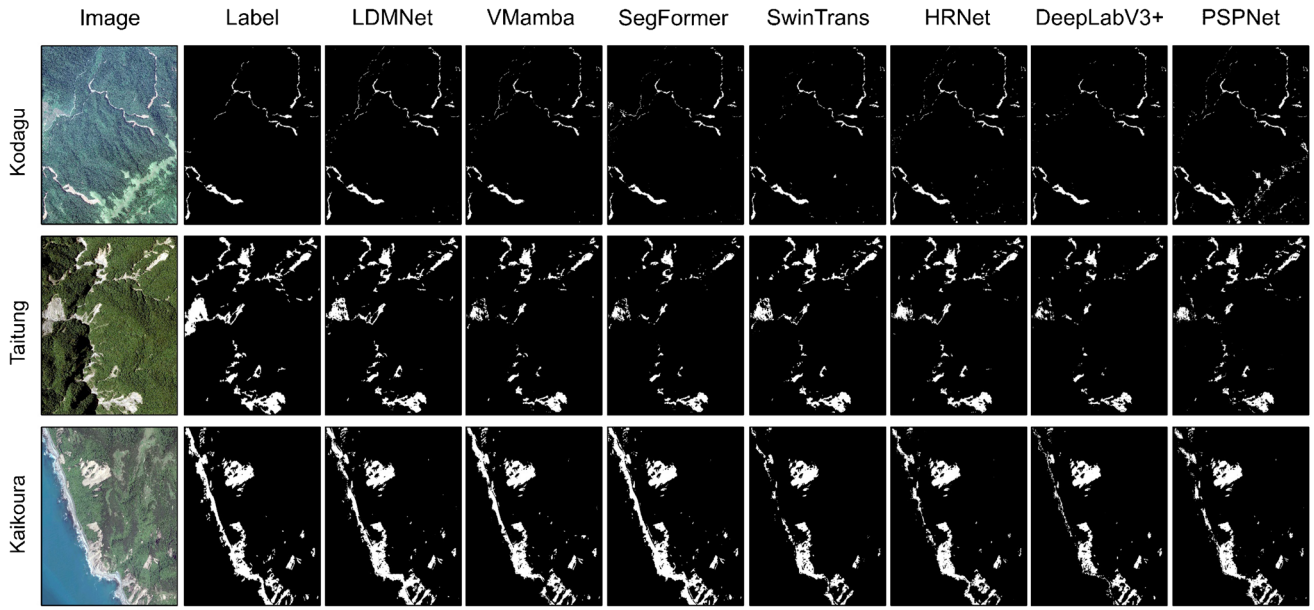

Fig. 3 Detection results of three landslide samples in the GVLM test set using LDM-Net and comparison algorithms

collapse and misclassification. Compared with the standard FFN and the baseline VMamba block, the proposed G_VSS with G_FFN improves feature representation and segmentation accuracy, particularly for partially overlapping or small landslides.

Verification for independent regions

To assess the effectiveness of the proposed LDM-Net for disaster management, independent landslide events were selected for evaluation. This experiment was designed to test cross-event generalisation under fixed post-event inference conditions using temporally and geographically independent cases (Supplementary Note S3). All cross-event evaluations used the same fixed RGB normalisation. Temporal radiometric variability was not adaptively corrected and is therefore treated here as a limitation rather than as a resolved component of the framework. The evaluation cases included earthquake-triggered landslides associated with the Mw 6.0 Mesetas earthquake in the Eastern Cordillera of Colombia on 24 December 2019 (Poveda et al. 2022), earthquake-triggered landslides associated with the Mw 7.8 and 7.5 Pazarcık earthquakes in south-eastern Türkiye on 6 February 2023 (Görüm et al., 2023), and rainfall-triggered landslides in Meizhou, Guangdong, China, on 14 May 2024 (Yang et al. 2025). For all events, PlanetScope imagery was manually interpreted to generate landslide labels. To mitigate the limitations associated with single-source training data (Xu et al., 2024; Fang et al. 2024), the HR-GLDD, GVLM, and ULMD datasets were combined into a multi-source landslide dataset and standardised to a common RGB input format and a unified tile size of 512 × 512 pixels (Supplementary Note S3). This standardisation ensured only a consistent deployment input representation and should not be interpreted as temporal normalisation, dynamic sensor harmonisation, or sequential database updating.

As shown in Table 3 and Supplementary Table 3, models trained on the multi-source landslide dataset outperformed those trained solely on HR-GLDD. In the Guangdong rainfall event, the model trained on the multi-source dataset improved the F1 score and IoU by 5.4% and 6.9%, reaching 76.5% and 61.9%, respectively. In the Mesetas earthquake event, gains of 3.3% and 4.0% were observed in F1 score and IoU, respectively. In the two Pazarcık earthquake cases, F1 score and IoU also increased by 6.8% and 6.1%, respectively. Visual results (Figs. 5, 6, and 7) further show that LDM-Net effectively identifies landslides while excluding non-landslide features such as buildings, roads, and impervious surfaces. These results indicate that, when trained on a diverse dataset, LDM-Net shows strong generalisation capability for rapid and accurate landslide detection across different disaster contexts.

Discussion

Computational efficiency advantage of LDM-Net

LDM-Net provides a favourable accuracy-efficiency trade-off for very-high-resolution (VHR) post-event mapping, delivering competitive segmentation performance while requiring substantially less GPU memory and fewer floating-point operations than attention-heavy Transformer-based architectures (Table 4; Fig. 8). Although the parameter count of LDM-Net is comparable to that of several commonly used backbones (e.g. ResNet101-, Swin-S-, and MiT-B4-based variants), its peak GPU memory demand is typically much lower as input size increases, and its FLOPs grow more gradually than those of Transformer-based models (Fig. 8). Compared with lightweight CNN baselines such as LinkNet, LDM-Net may exhibit slightly higher peak memory usage, but it remains markedly more compute-efficient (e.g. ~ 194.3 G FLOPs at 1600 × 1600 pixels, approximately one quarter of that of LinkNet). Therefore, the distinction of LDM-Net from existing approaches lies not only in its competitive segmentation accuracy, but also in its more favourable deployment-stage trade-off among accuracy, efficiency, and scalability for large-field-of-view VHR post-event mapping.

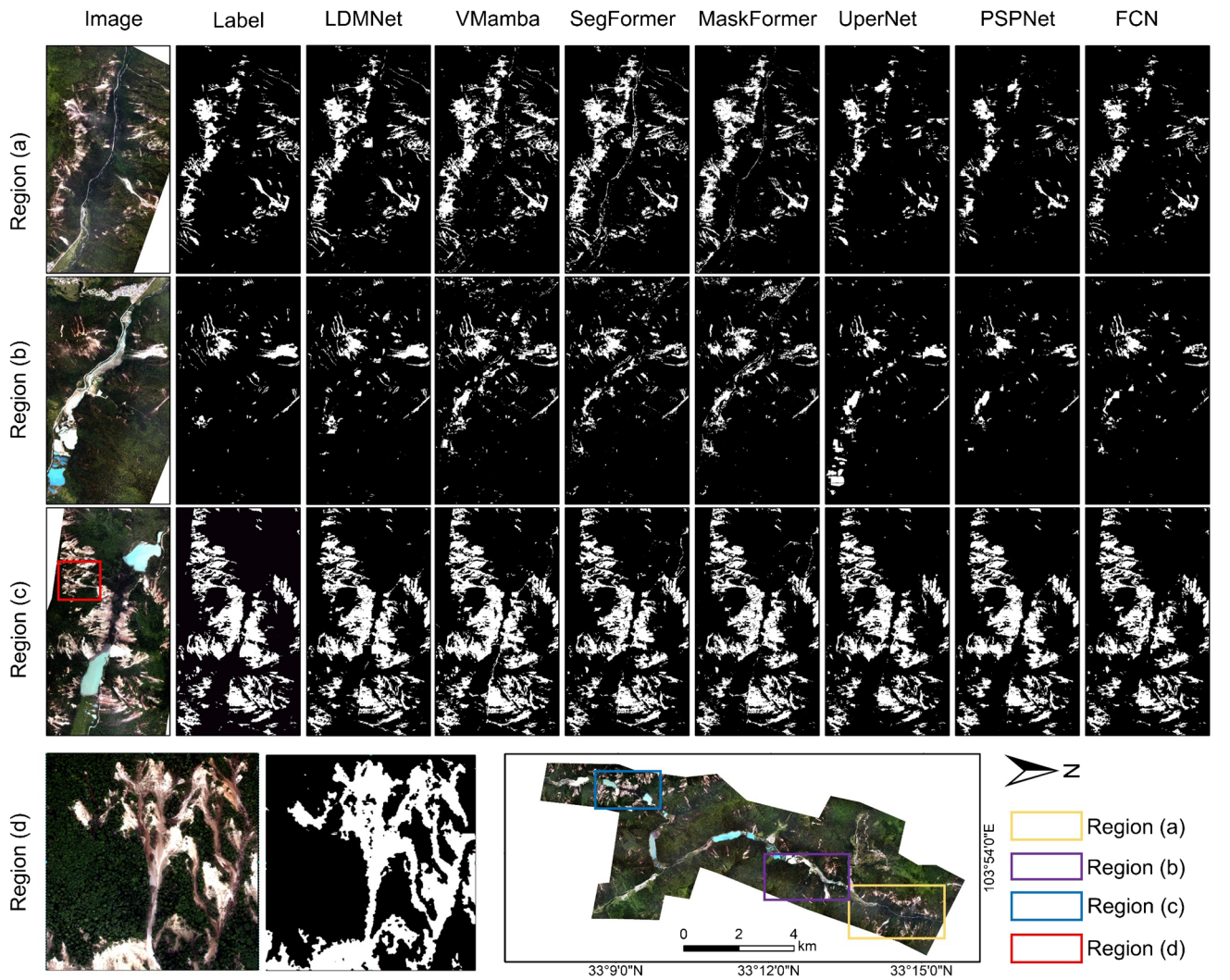

Fig. 4 Detection results of the 2017 Jiuzhaigou event in the ULMD test set using LDM-Net and comparison algorithms

An important operational implication is that LDM-Net can support substantially larger input tiles, i.e. a larger field of view under a fixed native image resolution, under limited GPU resources, which is often critical for rapid mapping over extensive VHR mosaics. Larger tiles provide broader contextual cues, including surrounding morphology, runout context, and background heterogeneity, and may reduce patch-wise fragmentation. Both effects are beneficial when delineating landslides as coherent objects rather than as isolated pixels. To examine this operational field-of-view effect in a controlled setting, a tile-size experiment was conducted on ULMD by re-tiling the same imagery and binary masks into patches of 256 × 256, 512 × 512, 1024 × 1024, and 1600 × 1600 pixels, while preserving the native image resolution and grid spacing. No resampling or resizing was applied. For each patch size, LDM-Net was trained from scratch and evaluated independently using the same training configuration.

In the present profiling setting, Transformer-based architectures showed steeper growth in memory use and computation as input tile size increased, which can limit their deployment for large-area, high-resolution mapping (Fig. 8). By contrast, the Mamba-style backbone in LDM-Net exhibited substantially more scalable behaviour with increasing input size in the present implementation, making large-tile training and inference more feasible. Consistent with the expected benefit of a broader field of view, increasing the ULMD tile size yielded modest but consistent improvements in pixel-level accuracy (Table 5; Fig. 8; Supplementary Table S4), with IoU increasing from 66.7% (256 × 256) to 68.1% (512 × 512), 69.3% (1024 × 1024), and 70.3% (1600 × 1600). The observed improvement was mainly reflected in Recall, whereas Precision remained broadly stable. This pattern is consistent with better recovery of landslide pixels that might otherwise be missed, rather than with a substantial increase in false positives. Qualitative examples (Supplementary Fig. S9) further suggest improved landslide integrity and reduced fragmentation or truncation in complex backgrounds when larger tiles are used.

At the same time, these pixel-level metrics are inherently sensitive to boundary placement and object geometry. Even without resampling, the proportion of boundary pixels varies

Table 2 Ablation comparison of FFN (VMamba) and the proposed G_FFN (LDM-Net) with 512×512 input on ULMD. Encoder parameters/FLOPs/GPU memory are measured for a single forward pass under the same input setting, metrics are computed for the landslide (foreground) class only at the pixel level, the best scores are highlighted in bold

Metrics	FFN (VMamba)	G_FFN (LDM-Net)
Recall (%)	74.8	**77.2**
Precision (%)	86.0	**86.3**
F1 (%)	80.0	**81.5**
IoU (%)	67.3	**68.1**
PR-AUC	0.813	**0.820**
Encoder parameters (M)	49.38	**55.7**
Encoder Flops (G)	41.3	**19.0**
Encoder GPU memory footprint (M)	2101	**1916**

with landslide size and shape (e.g. small, elongated, or highly irregular landslides), and re-tiling can change how frequently landslide bodies are truncated at patch borders. As a result, some variation in Recall, Precision, F1-score, IoU, and PR-AUC across input sizes may reflect boundary quantisation and tile-border effects, in addition to genuine differences in the contextual information available to the model. Therefore, the values reported in Table 5 should be interpreted as an empirical ULMD-specific trend rather than as a universal effect size. Their main practical significance lies in showing that larger tiles, when made computationally feasible by LDM-Net, can improve landslide-body completeness and contextual consistency in rapid VHR post-event mapping, even if the absolute metric gains remain modest.

In addition to the tile-size effect discussed above, object geometry relative to patch extent is another important source of uncertainty when interpreting pixel-level segmentation performance, as further discussed in Section 5.3.

Application in recent landslide events

Identifying landslide-prone slopes is important because it helps civil protection agencies allocate resources when new events occur. In this context, landslide susceptibility assessment helps identify slopes with a high propensity to fail on the basis of past or recent landslide occurrences. However, such analysis is only possible when reliable information on recent failures is available. Systematic landslide mapping therefore has value beyond supporting rapid emergency response, as it also provides the basis for anticipating which hillslopes are most likely to fail in future events.

To evaluate this facet, the utility of LDM-Net-derived inventories for landslide susceptibility modelling was examined, as this is an important component of both landslide hazard mitigation and risk management. For susceptibility modelling, a Random Forest classifier was implemented in scikit-learn (n_estimators = 50, max_depth = 50, with all remaining parameters left at their default values). To ensure a fair comparison between the manual and LDM-Net inventories, identical conditioning factors and sampling strategies were adopted. The conditioning factors and sampling strategy for earthquake-induced landslide modelling followed Dai et al. (2023), whereas those for rainfall-induced landslide modelling followed Huang et al. (2024a, b). For earthquake-triggered landslides, susceptibility results based on the LDM-Net inventory reproduced the main patterns of unstable slopes in complex mountainous terrain, although their discriminative performance was lower than that of the manually interpreted benchmark. The main discrepancies occurred in non-mountainous areas such as riverbeds and bare soil, as observed in the south-eastern part of Fig. 9b, where spectrally similar deposits were occasionally misidentified as landslides. These commission errors introduced additional positive samples in low-relief or depositional settings and thereby weakened the geomorphic contrast between landslide and non-landslide conditions, resulting in a 13% reduction in AUC relative to the manually curated inventory.

For rainfall-induced landslides, the susceptibility surface derived from the LDM-Net inventory reproduced almost all major spatial patterns obtained from the manual inventory when the same conditioning factors were applied (Fig. 10). The remaining discrepancies were mainly confined to narrow river corridors, where sandy or unvegetated channel deposits may be confused

Table 3 Performance of LDM-Net in independent rainfall- and earthquake-triggered landslide events. Metrics are computed for the landslide foreground class only at the pixel level

Events	Dataset	Recall (%)	Precision (%)	F1 (%)	IoU (%)	PR-AUC
Guangdong Rainfall	Multi-dataset	74.7	78.4	76.5	61.9	0.759
	HR-GLDD	72.1	70.1	71.1	55.0	0.701
Mesetas Earthquake	Multi-dataset	74.1	72.3	73.2	57.7	0.726
	HR-GLDD	70.1	69.7	69.9	53.7	0.683
Turkey Earthquake	Multi-dataset	79.3	48.1	59.9	42.5	0.621
	HR-GLDD	66.5	44.1	53.1	36.4	0.574

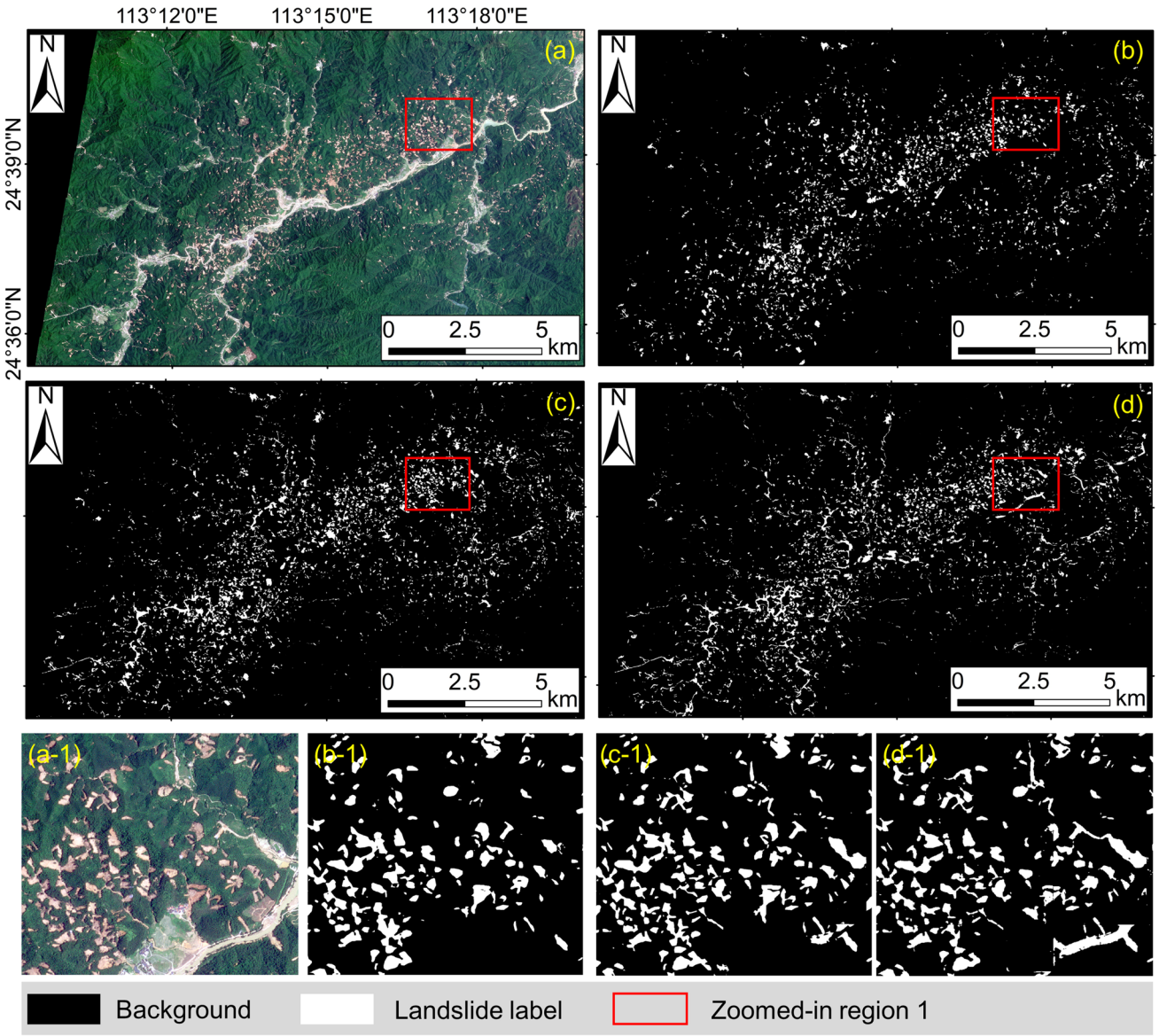

Fig. 5 Prediction results of landslides induced by rainfall events in Guangdong in 2024. **a** PlanetScope image, **b** ground truth, **c** prediction results based on multi-source datasets with LDM-Net, and (**d**) prediction results based on HR-GLDD with LDM-Net. (a-1), (b-1), (c-1), and (d-1) represent the image, label, prediction results based on multi-source datasets with LDM-Net in region 1, and prediction results based on HR-GLDD with LDM-Net in region 1, respectively

with landslide material, leading to a smaller AUC decrease of 7%. Overall, although susceptibility modelling based on the LDM-Net inventory was somewhat less accurate than that based on manual interpretation, it preserved the first-order spatial pattern of landslide-prone areas and was sufficiently reliable for rapid regional screening and emergency-oriented hazard assessment when timely manual mapping was not feasible (Table 6).

Future directions

The present workflow yields binary semantic inventories that often amalgamate adjacent landslide bodies. Such amalgamations, resulting from spatial proximity, overlapping failure surfaces, or progressive reactivation, often lead to overestimation of landslide extent and misclassification of individual landslide kinematics (Hungr et al. 2014). Under the present binary semantic-segmentation labels, rigorous size-stratified, object-level evaluation remains difficult because individual landslide instances are not consistently preserved. As shown in, adjacent landslides may be merged in inventories generated by the deployed LDM-Net during inference, which can limit downstream analyses. Future work should therefore move beyond purely semantic segmentation towards instance-aware delineation, for example by integrating instance-segmentation frameworks (e.g. Mask R-CNN-type models) and/or boundary-aware learning to separate touching or overlapping landslide bodies. In addition, lightweight post-processing strategies,

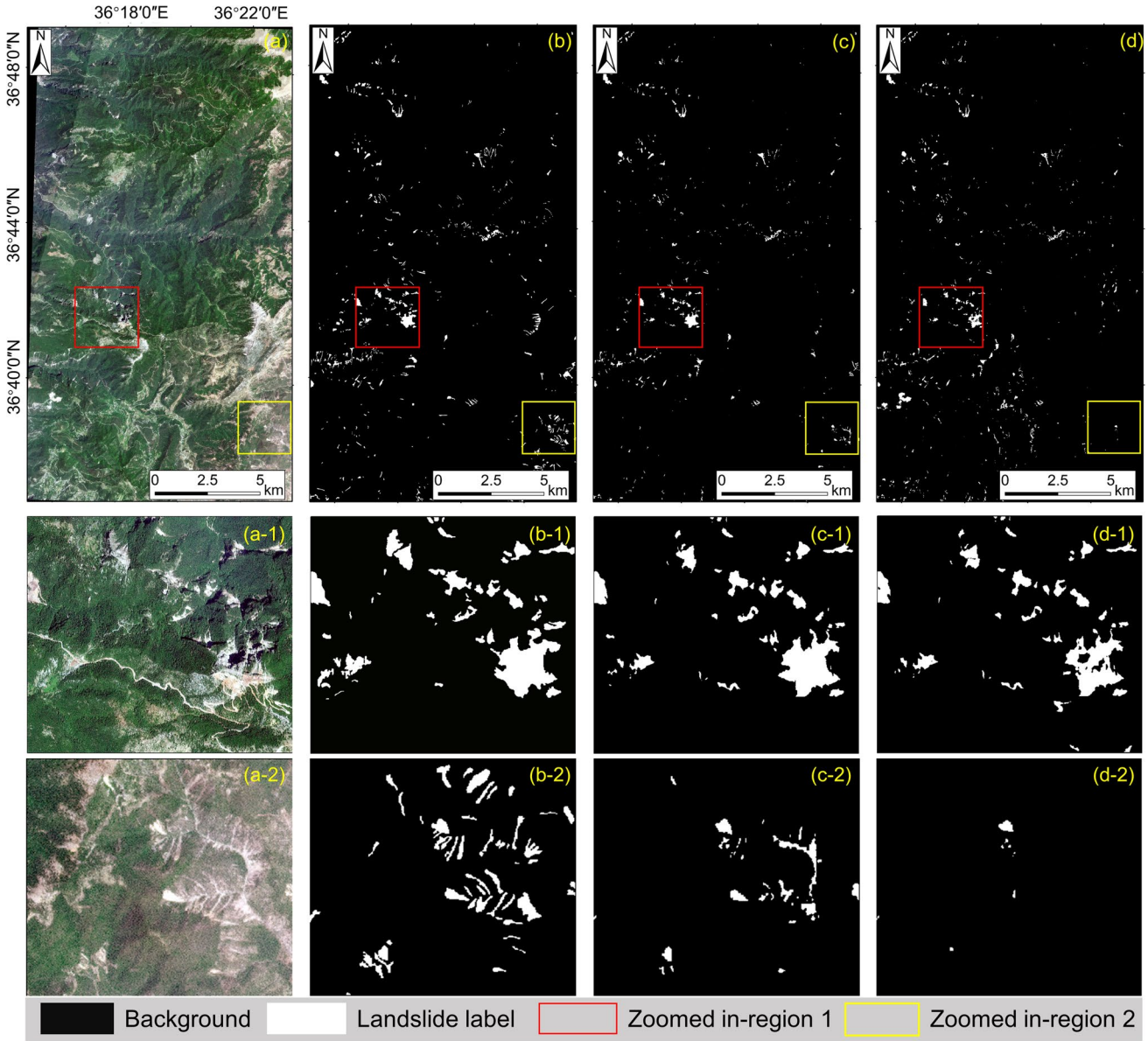

Fig. 6 Prediction results of landslides induced by earthquake events in Türkiye in 2023, **a** PlanetScope image, **b** ground truth, **c** prediction results based on multi-source datasets and LDM-Net, and (**d**) prediction results based on HR-GLDD and LDM-Net, (a-1), (b-1), (c-1), and (d-1) represent the image, label, prediction results based on multi-source datasets with LDM-Net, and prediction results based on HR-GLDD with LDM-Net in region 1 respectively, (a-2), (b-2), (c-2), and (d-2) represent the image, label, prediction results based on multi-source datasets with LDM-Net, and prediction results based on HR-GLDD with LDM-Net in region 2 respectively

such as distance-transform-guided marker-controlled watershed splitting, topology-constrained refinement, or object-based correction using DEM-derived morphometric cues, may further improve the disaggregation of merged units. Such instance-consistent inventories would directly support advanced applications including runout modelling, hazard assessment, and risk analysis (Liu et al. 2024).

A further limitation concerns the relationship between ground sampling distance and real-world landslide size. Although LDM-Net was evaluated on multiple benchmark datasets, these datasets have different ground sampling distances (e.g. HR-GLDD, GVLM, and ULMD), and the same landslide may therefore occupy substantially different numbers of pixels across datasets. This scale mismatch is particularly important for small landslides close to the minimum mappable size, for which one- or two-pixel boundary shifts, mixed-pixel effects, and annotation uncertainty may lead to disproportionately large changes in pixel-level metrics (Precision, Recall, F_1, and IoU). The effect may be more pronounced for elongated or highly irregular landslides because of their higher boundary-to-area ratios. Accordingly, the reported metrics should be interpreted as dataset-specific indicators of operational performance under each dataset's resolution and annotation conditions, rather than as strictly

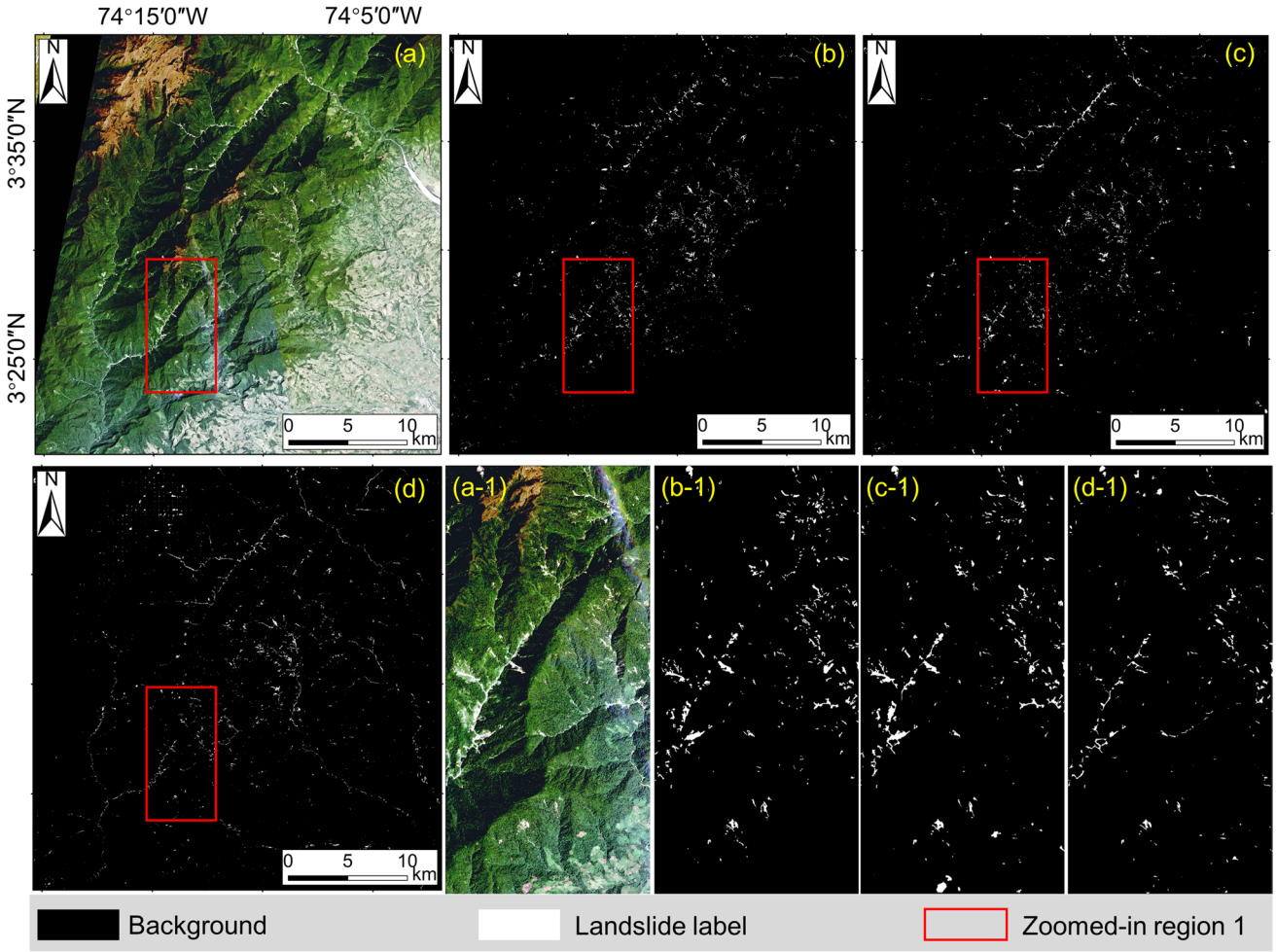

Fig. 7 Prediction results of landslides induced by earthquake events in Mesetas in 2019, **a** PlanetScope image, **b** ground truth, **c** prediction results based on multi-source datasets and LDM-Net, and (**d**) prediction results based on HR-GLDD and LDM-Net, (a-1), (b-1), (c-1), and (d-1) represent the image, label, prediction results based on multi-source datasets with LDM-Net, and prediction results based on HR-GLDD with LDM-Net in region 1 respectively

Table 4 Computational characteristics of representative segmentation models

Models	Encoder	Parameters (M)	Encoder parameters (M)	Reference
LinkNet	ResNet101	50.2	42.5	Chaurasia and Culurciello 2017
MaskFormer	VIT-B	118.7	86.5	Cheng et al. 2021
SwinTrans	Swin-S	81.7	48.8	Liu et al. 2021
SegFormer	MiT-B4	63.99	60.8	Xie et al. 2021
LDM-Net	G_VSS	58.88	55.7	**This study**

scale-invariant measures. Future work should include size-stratified object-level evaluation (e.g. small-, medium-, and large-landslide classes), boundary-aware metrics, and cross-resolution sensitivity analyses to quantify more explicitly the effect of pixel-to-reality scale on delineation accuracy. Further work should also examine robustness to interpolation artifacts, image-quality variation, and heterogeneous sensor characteristics more systematically, and may explore enhanced image-processing or image-fusion strategies to

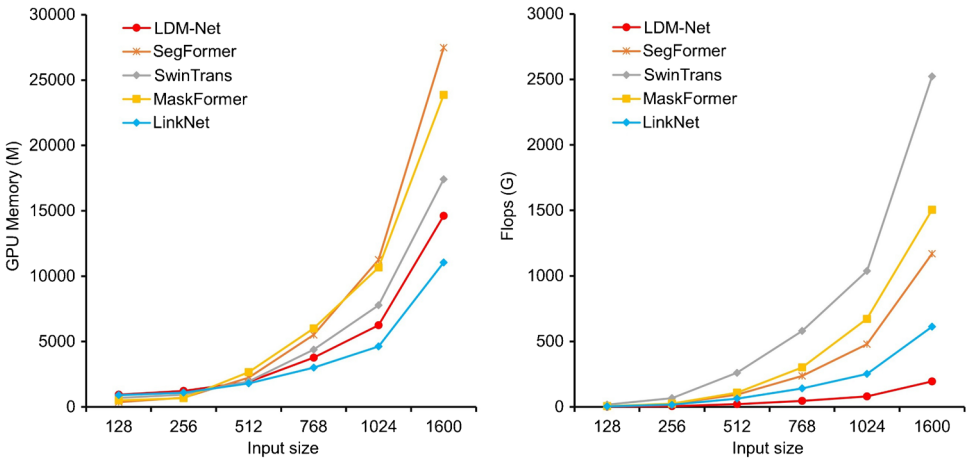

Fig. 8 GPU memory footprint and FLOPs of representative models with varying input sizes, including LinkNet, MaskFormer, Swin Transformer, SegFormer, and the proposed LDM-Net (Chaurasia and Culurciello 2017; Cheng et al. 2021; Liu et al. 2021; Xie et al. 2021; this study)

improve deployment stability across heterogeneous optical conditions and resolutions (e.g. Mishra et al. 2023).

The present framework should therefore be interpreted as an event-based post-disaster mapping model rather than as a dynamically updated monitoring system. It should likewise not be interpreted as an AutoML, uncertainty-quantification, or self-updating automated modelling workflow. This focus on single-time post-event imagery is deliberate for rapid response, because suitable cloud-free and geometrically matched pre-event imagery is not always available within the required emergency time window. However, this design also limits applicability to progressive or slow-moving landslides, for which multi-temporal monitoring, pre-post change detection, or time-series deformation analysis may be more appropriate (Hervás et al. 2003; Lu et al. 2019; Bekaert et al. 2020). As illustrated by the landslide mapping results for the Pazarcık earthquake (Fig. 6), model performance diverges markedly across contrasting surface conditions. On the vegetated western flank of the scene, the network delineates slide boundaries clearly (Fig. 6a-1), recovering both shape and extent. By contrast, the nearly bare eastern slopes reveal a clear limitation: only the larger failures that form conspicuous scars are captured, whereas shallow slides, which produce subtler spectral contrasts in the exposed regolith, may remain undetected (Fig. 6a-2). Future operational extensions should therefore consider preprocessing and normalisation schemes that better accommodate heterogeneous image quality, radiometric inconsistency, and temporal variability across deployment settings (Abbaszadeh Shahri et al. 2024). Explicit temporal normalisation and time-series modelling may further support dynamic landslide monitoring and time-dependent hazard assessment, particularly because sequential landslide prediction is sensitive to data preparation, sampling frequency, temporal splitting, and time-varying external forcings (Ebrahim et al. 2024; Huang et al. 2025).

These disparities further illustrate how variations in topography, lithology, climate, vegetation cover, triggering mechanism, and sensor-specific spectral response can each pose substantial challenges to universal landslide segmentation. Models trained on data from specific regions or sensor types often experience performance reductions when applied to new regions or instruments, largely because of differences in landslide statistical characteristics, spectral signatures, and texture features. Domain-adaptation strategies, particularly transfer learning, offer promising approaches for addressing these issues. By fine-tuning pre-trained models on smaller, region-specific datasets (Bhuyan et al. 2023), feature representations may be adapted to regional morphological variation, including differences in scarp formation, material composition, and sliding behaviour. Future work should explore these techniques further to improve model universality and operational robustness.

Table 5 Effect of input tile size (field of view) on ULMD under fixed native image resolution. Metrics are computed at the pixel level for the landslide (foreground) class only. The same imagery and binary masks were re-tiled into different patch sizes without resampling or resizing. The best scores are highlighted in bold

Input sizes	Recall (%)	Precision (%)	F1 (%)	IoU (%)	PR-AUC
256×256	75.1	84.5	79.5	66.7	0.810
512×512	77.2	86.3	81.5	68.1	0.820
1024×1024	82.5	82.7	82.6	69.3	0.826
1600×1600	83.6	82.7	83.2	70.3	0.833

Conclusion

In this study, Landslide Mapping Network (LDM-Net) was developed for robust and efficient post-event landslide delineation from remote-sensing imagery. The main conclusions are as follows.

1. Methodologically, LDM-Net provides an efficient state-space framework for landslide mapping by integrating a Mamba-based Global Vision State Space module (G_VSS) with a Global Feed-Forward Network (G_FFN). This design enables long-range spatial-context modelling, multi-scale feature fusion, and efficient feature extraction within a lightweight encoder-decoder architecture, thereby improving landslide delineation

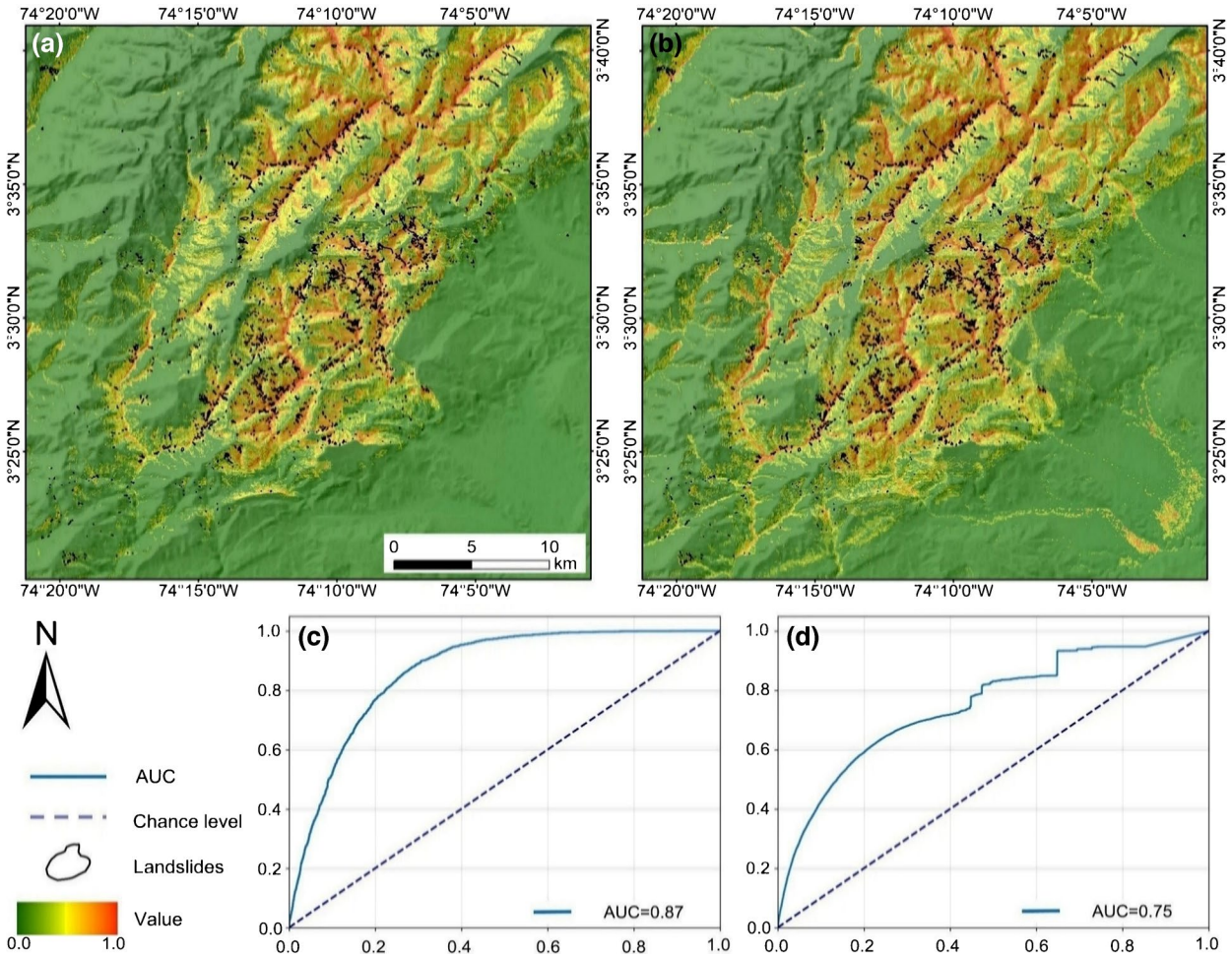

Fig. 9 Landslide susceptibility results for the 2019 Mesetas earthquake. **a** Susceptibility map derived from the manually interpreted inventory; **b** susceptibility map derived from the LDM-Net inventory; **c** receiver operating characteristic (ROC) curve corresponding to (a); and (**d**) ROC curve corresponding to (b). The area under the ROC curve (AUC) summarises the discrimination ability of each susceptibility model, with higher values indicating better performance

Table 6 Comparison of susceptibility models built from manual vs. LDM-Net inventories

Event	Inventory	Presence points	Total landslide area (km^2)	Susceptibility model	AUC
Meizhou	Manual	2845	8.34	RF	0.89
Meizhou	LDM-Net	3321	9.63	RF	0.82
Mesetas	Manual	2536	8.93	RF	0.87
Mesetas	LDM-Net	2945	10.21	RF	0.75

from very-high-resolution imagery under rapid-response conditions.

2. In terms of segmentation performance and computational efficiency, LDM-Net shows a favourable trade-off among accuracy, efficiency, and scalability relative to representative CNN- and Transformer-based baselines. Across three benchmark datasets, namely HR-GLDD, GVLM, and the newly compiled 0.2 m ULMD, LDM-Net achieved strong performance in terms of Precision, Recall, F1 score, and IoU. In addition, GPU memory use and computational cost increased more gradually with input size than in attention-heavy Transformer-based models, making large-field-of-view training and inference more feasible under practical disaster-response constraints. Independent evaluations in geographically and tectonically distinct regions further indicate good cross-event generalisation.

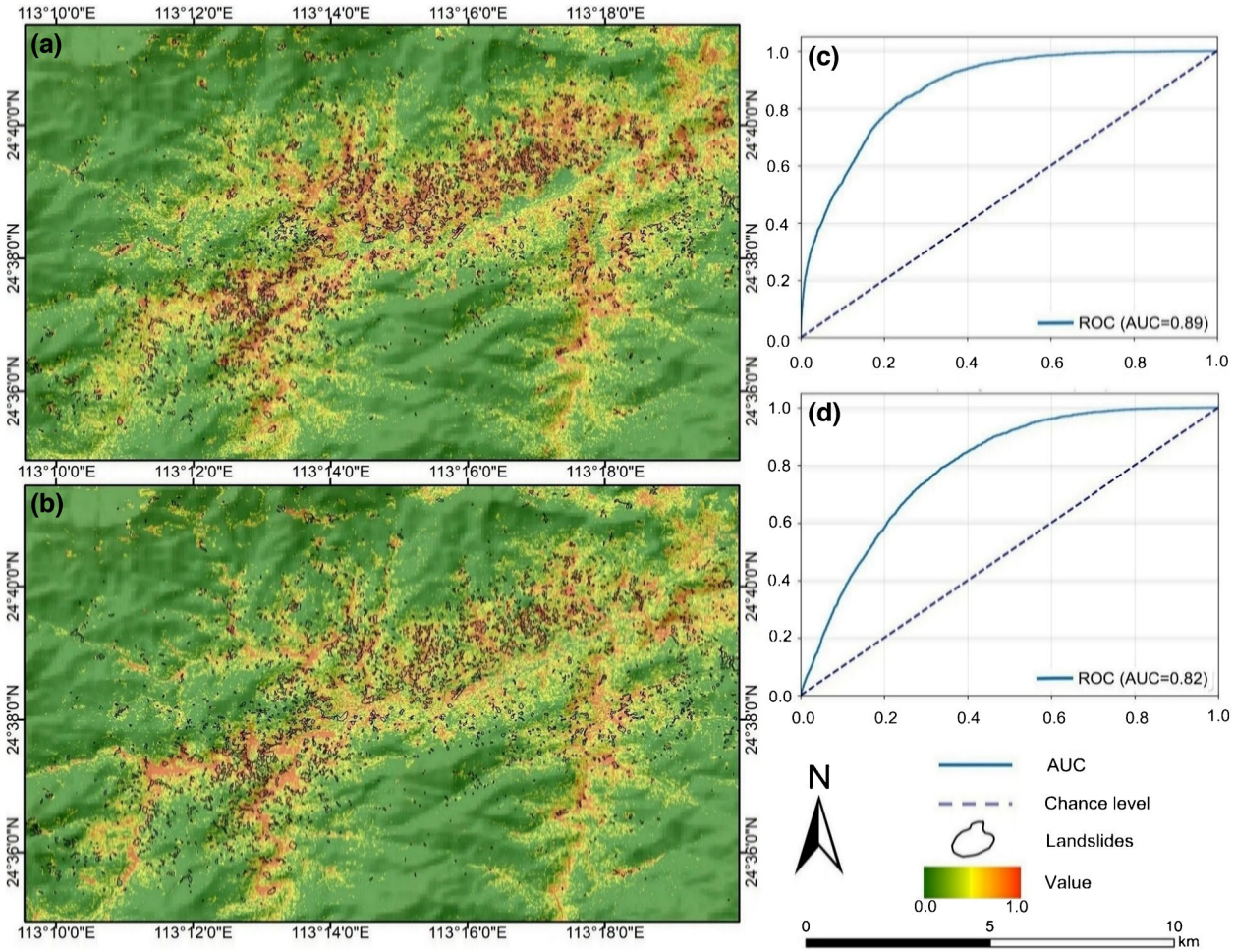

Fig. 10 Landslide susceptibility results for the 2024 Guangdong rainfall event. **a** Susceptibility map derived from the manually interpreted inventory; **b** susceptibility map derived from the LDM-Net inventory; **c** receiver operating characteristic (ROC) curve corresponding to (a); and (**d**) ROC curve corresponding to (b). The area under the ROC curve (AUC) summarises the discrimination ability of each susceptibility model, with higher values indicating better performance

3. In terms of practical application, inventories generated by the deployed LDM-Net support subsequent landslide susceptibility modelling for both earthquake-triggered and rainfall-induced cases. Although susceptibility results derived from LDM-Net inventories remain slightly less accurate than those based on manual interpretation, they reproduce the first-order spatial pattern of landslide-prone areas, particularly in complex mountainous terrain. This indicates that LDM-Net can help bridge the gap between research-oriented segmentation models and practical post-event inventory production by providing timely and spatially detailed landslide information for emergency response, susceptibility assessment, and hazard management.

Overall, LDM-Net should be understood as a deployed post-event segmentation framework under a fixed experimental protocol, rather than as an AutoML, self-updating, or uncertainty-aware automated modelling workflow. Within this scope, its combination of delineation accuracy, computational efficiency, and deployment-stage scalability makes it a promising tool for rapid emergency response, risk assessment, and preparedness for cascading geomorphic hazards.

Acknowledgements

We would like to thank the State Key Laboratory of Geohazard Prevention and Geoenvironment Protection for providing UAV data.

Funding

The research is supported by the National Science Fund for Distinguished Young Scholars of China (grant 42125702), the National Natural Science Foundation of China (grants 42307263 and W2412050), the New Cornerstone Science Foundation through the XPLORER PRIZE (grant XPLORER-2022-1012), the National Science and Technology Major Project of Sichuan Province (grant 2024ZDZX0020), the National Key R&D Program of China (grant 2022YFF0800604), the China Scholarship Council (grant 202409230002), and Chengdu University of Technology Postgraduate Innovative Cultivation Program (program 2024BJCX007).

Data and code availability

The data supporting this study are freely available from Zenodo at https://doi.org/10.5281/zenodo.12792725. The code are freely available at https://github.com/fshutong/LDM-Net.

Declarations

Competing interests The authors declare no competing interests.

References

Abbaszadeh Shahri A, Maghsoudi Moud F (2021) Landslide susceptibility mapping using hybridized block modular intelligence model. Bulletin of Engineering Geology and the Environment 80:267–284. https://doi.org/10.1007/s10064-020-01922-8

Abbaszadeh Shahri A, Shan C, Larsson S, Johansson F (2024) Normalizing large scale sensor-based MWD data: an automated method toward a unified database. Sensors 24:1209. https://doi.org/10.3390/s24041209

Bekaert DPS, Handwerger AL, Agram P, Kirschbaum DB (2020) InSAR-based detection method for mapping and monitoring slow-moving landslides in remote regions with steep and mountainous terrain: an application to Nepal. Remote Sens Environ 249:111983. https://doi.org/10.1016/j.rse.2020.111983

Bhuyan K, Tanyaş H, Nava L, Puliero S, Meena SR, Floris M, van Westen C, Catani F (2023) Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. Scientific Reports 13:162. https://doi.org/10.1038/s41598-022-27352-y

Brardinoni F, Slaymaker O, Hassan MA (2003) Landslide inventory in a rugged forested watershed: a comparison between air-photo and field survey data. Geomorphology 54:179–196. https://doi.org/10.1016/S0169-555X(02)00355-0

Chaurasia A, Culurciello E (2017) LinkNet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE Visual Communications and Image Processing (VCIP), pp 1–4. https://doi.org/10.1109/VCIP.2017.8305148

Chen L, Zhu Y, Papandreou G, Schroff F, Adam H (2018) Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp 801–818. https://doi.org/10.1007/978-3-030-01234-2_49

Cheng B, Schwing AG, Kirillov A (2021) Per-pixel classification is not all you need for semantic segmentation. Adv Neural Inf Process Syst 34:17864–17875. https://doi.org/10.5555/3540261.3541628

Cheng G, Huang Y, Li X, Lyu S, Xu Z, Zhao H, Zhao Q, Xiang S (2024) Change detection methods for remote sensing in the last decade: a comprehensive review. Remote Sensing 16:2355. https://doi.org/10.3390/rs16132355

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

Coe JA, Ellis WL, Godt JW, Savage WZ, Savage JE, Michael JA, Kibler JD, Powers PS, Lidke DJ, Debray S (2003) Seasonal movement of the Slumgullion landslide determined from Global Positioning System surveys and field instrumentation, July 1998-March 2002. Eng Geol 68:67–101. https://doi.org/10.1016/S0013-7952(02)00199-0

Dai L, Fan X, Wang X, Fang C, Zou C, Tang X, Wei Z, Xia M, Wang D, Xu Q (2023) Coseismic landslides triggered by the 2022 Luding Ms 6.8 earthquake, China. Landslides 20:1277–1292. https://doi.org/10.1007/s10346-023-02061-3

Dao T, Gu A (2024) Transformers are SSMs: generalized models and efficient algorithms through structured state space duality. In: Proceedings of the 41st International Conference on Machine Learning (PMLR), 235:10041–10071. https://proceedings.mlr.press/v235/dao24a.html

De Boer P-T, Kroese DP, Mannor S, Rubinstein RY (2005) A tutorial on the cross-entropy method. Ann Oper Res 134:19–67. https://doi.org/10.1007/s10479-005-5724-z

Diakogiannis FI, Waldner F, Caccetta P, Wu C (2020) ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. ISPRS J Photogramm Remote Sens 162:94–114. https://doi.org/10.1016/j.isprsjprs.2020.01.013

Ebrahim KMP, Fares A, Faris N, Zayed T (2024) Exploring time series models for landslide prediction: a literature review. Geoenvironmental Disasters 11:25. https://doi.org/10.1186/s40677-024-00288-3

Fan X, Scaringi G, Xu Q, Zhan W, Dai L, Li Y, Pei X, Yang Q, Huang R (2018) Coseismic landslides triggered by the 8th August 2017 M s 7.0 Jiuzhaigou earthquake (Sichuan, China): factors controlling their spatial distribution and implications for the seismogenic blind fault identification. Landslides 15:967–983. https://doi.org/10.1007/s10346-018-0960-x

Fan X, Scaringi G, Korup O, West AJ, van Westen CJ, Tanyas H, Hovius N, Hales TC, Jibson RW, Allstadt KE, Zhang L, Evans SG, Xu C, Li G, Pei X, Xu Q, Huang R (2019) Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. Rev Geophys 57:421–503. https://doi.org/10.1029/2018RG000626

Fang C, Fan X, Wang X, Nava L, Zhong H, Dong X, Qi J, Catani F (2024) A globally distributed dataset of coseismic landslide mapping via multi-source high-resolution remote sensing images. Earth Syst Sci Data 16:4817–4842. https://doi.org/10.5194/essd-16-4817-2024

Ferrer JV, Samprogna Mohor G, Dewitte O, Pánek T, Reyes-Carmona C, Handwerger AL, Hürlimann M, Köhler L, Teshebaeva K, Thieken AH, Tsou C-Y, Urgilez Vinueza A, Demurtas V, Zhang Y, Zhao C, Marwan N, Kurths J, Korup O (2024) Human settlement pressure drives slow-moving landslide exposure. Earths Future 12:e2024EF004830. https://doi.org/10.1029/2024EF004830

Gorum T, Fan X, van Westen CJ, Huang RQ, Xu Q, Tang C, Wang G (2011) Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake. Geomorphology 133:152–167. https://doi.org/10.1016/j.geomorph.2010.12.030

Görüm T, Tanyas H, Karabacak F, Yılmaz A, Girgin S, Allstadt KE, Süzen ML, Burgi P (2023) Preliminary documentation of coseismic ground failure triggered by the February 6, 2023 Türkiye earthquake sequence. Eng Geol 327:107315. https://doi.org/10.1016/j.enggeo.2023.107315

Gu A, Dao T (2024) Mamba: linear-time sequence modeling with selective state spaces. In: First Conference on Language Modeling (COLM 2024). OpenReview

Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang K-T (2012) Landslide inventory maps: new tools for an old problem. Earth Sci Rev 112:42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Hervás J, Barredo JI, Rosin PL, Pasuto A, Mantovani F, Silvano S (2003) Monitoring landslides from optical remotely sensed imagery: the case history of Tessina landslide, Italy. Geomorphology 54(1–2):63–75. https://doi.org/10.1016/S0169-555X(03)00056-4

Huang R, Fan X (2013) The landslide story. Nat Geosci 6:325–326. https://doi.org/10.1038/ngeo1806

Huang F, Xiong H, Jiang S-H, Yao C, Fan X, Catani F, Chang Z, Zhou X, Huang J, Liu K (2024a) Modelling landslide susceptibility prediction: a review and construction of semi-supervised imbalanced theory. Earth-Sci Rev 250:104700. https://doi.org/10.1016/j.earscirev.2024.104700

Huang Y, Zhang J, He H, Jia Y, Chen R, Ge Y, Ming Z, Zhang L, Li H (2024b) MAST: an earthquake-triggered landslides extraction method combining morphological analysis edge recognition with Swin-Transformer deep learning model. IEEE J Sel Top Appl Earth Obs Remote Sens 17:2586–2595. https://doi.org/10.1109/JSTARS.2023.3342989

Huang M, Nishimura S-I, Shibata T, Wang ZZ (2025) End-to-end time-dependent probabilistic assessment of landslide hazards using hybrid deep learning simulator. Computers and Geotechnics 178:106920. https://doi.org/10.1016/j.compgeo.2024.106920

Hungr O, Leroueil S, Picarelli L (2014) The Varnes classification of landslide types, an update. Landslides 11:167–194. https://doi.org/10.1007/s10346-013-0436-y

Jiang R, Zhang L, Peng D, He X, He J (2021) The landslide hazard chain in the Tapovan of the Himalayas on 7 February 2021. Geophys Res Lett 48(17):e2021GL093723. https://doi.org/10.1029/2021GL093723

Jiang R, Zhang L, Lu W, Peng D, He X, Xiao S, Wei M (2026) A numerical model for cascading glacier mass flow analysis (GMFA): erosion-deposition dynamics, phase changes, and multi-hazard chain transformations. J Rock Mech Geotech Eng 18(1):299–316. https://doi.org/10.1016/j.jrmge.2025.02.010

Jin B, Zeng T, Liu S, Li Y, Gui L, Zhao B, Yin K, Catani F, Peduto D (2025) Quantitative risk assessment for rainfall-induced landslides of transmission line towers: the case of Chongqing national transmission protection regions. Int J Disaster Risk Reduct 128:105715. https://doi.org/10.1016/j.ijdrr.2025.105715

Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo W-Y, Dollár P, Girshick R (2023) Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp 4015–4026. https://doi.org/10.1109/ICCV51070.2023.00371

Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B (2021) Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp 10012–10022. https://doi.org/10.1109/ICCV48922.2021.00986

Liu Y, Tian Y, Zhao Y, Yu H, Xie L, Wang Y, Ye Q, Jiao J, Liu Y (2024) VMamba: visual state space model. Adv Neural Inf Process Syst 37:103031–103063. https://doi.org/10.52202/079017-3273

Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965

Lu P, Qin Y, Li Z, Mondini AC, Casagli N (2019) Landslide mapping from multi-sensor data through improved change detection-based Markov random field. Remote Sens Environ 231:111235. https://doi.org/10.1016/j.rse.2019.111235

Lv P, Ma L, Li Q, Du F (2023) ShapeFormer: a shape-enhanced vision transformer model for optical remote sensing image landslide detection. IEEE J Sel Top Appl Earth Obs Remote Sens 16:2681–2689. https://doi.org/10.1109/JSTARS.2023.3253769

Ma X, Zhang X, Pun M-O (2024) RS3Mamba: visual state space model for remote sensing images semantic segmentation. IEEE Geosci Remote Sens Lett 21:1–5. https://doi.org/10.1109/LGRS.2024.3414293

Meena SR, Nava L, Bhuyan K, Puliero S, Soares LP, Dias HC, Floris M, Catani F (2023) HR-GLDD: a globally distributed dataset using generalized DL for rapid landslide mapping on HR satellite imagery. Earth Syst Sci Data 15:3283–3298. https://doi.org/10.5194/essd-15-3283-2023

Mishra VK, Nareti U, Kumar R, Pant T, Aleem A, Singh A, Biable SE (2023) GDF: a novel image fusion approach for compelling depiction of earthly features. J Sensors 2023:9429505. https://doi.org/10.1155/2023/9429505

Nava L, Monserrat O, Catani F (2022) Improving landslide detection on SAR data through deep learning. IEEE Geosci Remote Sens Lett 19:1–5. https://doi.org/10.1109/LGRS.2021.3127073

Poveda E, Pedraza P, Velandia F, Mayorga E, Plicka V, Gallovič F, Zahradník J (2022) 2019 Mw 6.0 Mesetas (Colombia) earthquake sequence: insights from integrating seismic and morphostructural observations. Earth Space Sci 9(12):e2022EA002465. https://doi.org/10.1029/2022EA002465

Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S (2019) Generalized intersection over union: a metric and a loss for bounding box regression. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp 658–666. https://doi.org/10.1109/CVPR.2019.00075

Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Tang X, Tu Z, Wang Y, Liu M, Li D, Fan X (2022) Automatic detection of coseismic landslides using a new transformer method. Remote Sens 14:2884. https://doi.org/10.3390/rs14122884

Udin W, Norazami N, Sulaiman N, Zaudin NC, Ma'ail S, Nor AM (2019) UAV based multi-spectral imaging system for mapping landslide risk area along Jeli-Gerik highway, Jeli, Kelantan. In: 2019 IEEE 15th International Colloquium on Signal Processing & Its Applications (CSPA), pp 162–167. https://doi.org/10.1109/CSPA.2019.8695995

Wang J, Sun K, Cheng T, Jiang B, Deng C, Zhao Y, Liu D, Mu Y, Tan M, Wang X, Liu W, Xiao B (2021) Deep high-resolution representation learning for visual recognition. IEEE Trans Pattern Anal Mach Intell 43(10):3349–3364. https://doi.org/10.1109/TPAMI.2020.2983686

Wang J, Lei D, Zhang Y, Yuan J, Liu C, Luo B, Liu Q, Wang G (2026) Med-Mamba: multi-scale deformable attention via state space models for robust medical image segmentation. Biomed Signal Process Control 112:108363. https://doi.org/10.1016/j.bspc.2025.108363

Wang X, Fan X, Xu Q, Du P (2022) Change detection-based co-seismic landslide mapping through extended morphological profiles and ensemble strategy. ISPRS Journal of Photogrammetry and Remote Sensing 187:225–239. https://doi.org/10.1016/j.isprsjprs.2022.03.011

Woo S, Debnath S, Hu R, Chen X, Liu Z, Kweon IS, Xie S (2023) ConvNeXt V2: co-designing and scaling ConvNets with masked autoencoders. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp 16133–16142. https://doi.org/10.1109/CVPR52729.2023.01548

Wu L, Liu R, Ju N, Zhang A, Gou J, He G, Lei Y (2024) Landslide mapping based on a hybrid CNN-transformer network and deep transfer learning using remote sensing images with topographic and spectral features. Int J Appl Earth Obs Geoinf 126:103612. https://doi.org/10.1016/j.jag.2023.103612

Xiao T, Liu Y, Zhou B, Jiang Y, Sun J (2018) Unified perceptual parsing for scene understanding. In: Proceedings of the European Conference on Computer Vision (ECCV), pp 418–434. https://doi.org/10.1007/978-3-030-01228-1_26

Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P (2021) SegFormer: simple and efficient design for semantic segmentation with transformers. Adv Neural Inf Process Syst 34:12077–12090. https://doi.org/10.5555/3540261.3541185

Xing Z, Ye T, Yang Y, Cai D, Gai B, Wu X-J, Gao F, Zhu L (2026) Seg-Mamba-V2: long-range sequential modeling mamba for general 3-D medical image segmentation. IEEE Trans Med Imaging 45(1):4–15. https://doi.org/10.1109/TMI.2025.3589797

Xu Q, Zhao B, Dai K, Dong X, Li W, Zhu X, Yang Y, Xiao X, Wang X, Huang J, Lu H, Deng B, Ge D (2023) Remote sensing for landslide investigations: a progress report from China. Eng Geol 321:107156. https://doi.org/10.1016/j.enggeo.2023.107156

Xu Y, Ouyang C, Xu Q, Wang D, Zhao B, Luo Y (2024) CAS landslide dataset: a large-scale and multisensor dataset for deep learning-based landslide detection. Sci Data 11:12. https://doi.org/10.1038/s41597-023-02847-z

Yang G, Zhao L, Qin Y, Yang T, Chen S (2025) Clustered landslides induced by rainfall in Jiangwan Town, Shaoguan City, Guangdong Province, China. Landslides 22:1325–1338. https://doi.org/10.1007/s10346-025-02463-5

Ye C, Li Y, Cui P, Liang L, Pirasteh S, Marcato J, Goncalves WN, Li J (2019) Landslide detection of hyperspectral remote sensing data based on deep learning with constrains. IEEE J Sel Top Appl Earth Obs Remote Sens 12:5047–5060. https://doi.org/10.1109/JSTARS.2019.2951725

Zhang C, Jiang W, Zhang Y, Wang W, Zhao Q, Wang C (2022) Transformer and CNN hybrid deep neural network for semantic segmentation of very-high-resolution remote sensing imagery. IEEE Trans Geosci Remote Sens 60:4408820. https://doi.org/10.1109/TGRS.2022.3144894

Zhang X, Yu W, Pun M-O, Shi W (2023) Cross-domain landslide mapping from large-scale remote sensing images using prototype-guided domain-aware progressive representation learning. ISPRS J Photogramm Remote Sens 197:1–17. https://doi.org/10.1016/j.isprsjprs.2023.01.018

Zhang H, Zhu Y, Wang D, Zhang L, Chen T, Wang Z, Ye Z (2024) A survey on visual Mamba. Appl Sci Basel 14(13):5683. https://doi.org/10.3390/app14135683

Zhao H, Shi J, Qi X, Wang X, Jia J (2017) Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 6230–6239. https://doi.org/10.1109/CVPR.2017.660

Zhong C, Liu Y, Gao P, Chen W, Li H, Hou Y, Nuremanguli T, Ma H (2020) Landslide mapping with remote sensing: challenges and opportunities. Int J Remote Sens 41:1555–1581. https://doi.org/10.1080/01431161.2019.1672904

Zhou Y, Wang H, Yang R, Yao G, Xu Q, Zhang X (2022) A novel weakly supervised remote sensing landslide semantic segmentation method: combining CAM and cycleGAN algorithms. Remote Sens 14(15):3650. https://doi.org/10.3390/rs14153650
Zhu Q, Cai Y, Fang Y, Yang Y, Chen C, Fan L, Nguyen A (2024) Samba: semantic segmentation of remotely sensed images with state space model. Heliyon 10(19):e38495. https://doi.org/10.1016/j.heliyon.2024.e38495

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10346-026-02789-8.

Chengyong Fang · Xuanmei Fan (✉) · Xin Wang (✉) · Kushanav Bhuyan · Xiangyang Dou · Mingyao Xia
State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China
Email: fxm_cdut@qq.com

Xin Wang
Email: wangxin@cdut.edu.cn

Chengyong Fang
Email: 2475372620@qq.com

Kushanav Bhuyan
Email: kushanavb@gmail.com

Xiangyang Dou
Email: douxyz1994@gmail.com

Kushanav Bhuyan · Filippo Catani
Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, Padua 35129, Italy
Email: filippo.catani@unipd.it

Chengyong Fang · Xiangyang Dou
Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede 7522 NH, Netherlands

Hao Zhong
College of Information Science and Technology, Chengdu University of Technology, Chengdu 610059, China
Email: zhonghao@cdut.edu.cn