International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(23s) 2026 pp. 686-701
https://doi.org/10.70917/ijcisim-2026-5632
Article

Landslide Segmentation from Multispectral Remote Sensing Data Using U-Net with Multichannel Topographic Information

Neeta M. Bajpai[1], Dr. Rahul M. Pethe[2]

[1,2]Department of Electronics and Telecommunication Engineering, S B Jain Institute of Technology Management and Research, Nagpur (M.S.)

Abstract: Landslides lie between high frequency normal hazards that cause moderate to high damage to infrastructure, ecosystems and human settlements, thus requiring precise and automated landslide detection and delineation. Conventional landslide mapping approaches are often dependent on manual interpretation, field surveys, and limited geospatial information, which can be time-consuming and problematic to scale over huge areas. This research presents a deep learning-based semantic segmentation framework using the U-Net architecture for pixel-level landslide delineation from multispectral and topographical remote sensing data. The projected method employs the Landslide4Sense dataset with 4,844 image patches of spatial size 128 × 128 pixels and 14 input channels including 12 Sentinel-2 multispectral bands, slope and Digital Elevation Model (DEM) information. The U-Net model is an encoder-decoder based feature extractor with skip connections to preserve spatial information and generate pixel level landslide masks. The model is trained and tested by the performance indicators of the Intersection over Union (IoU), F1-score, precision, recall and accuracy. The experimental results demonstrate that the proposed U-Net model provides 0.4300 IoU, 0.6014 F1-score, 0.5051 precision, 0.7431 recall and 0.9814 accuracy. The results show the efficiency of the U-Net in pixel detection for landslides, especially in terms of recall, while the lower precision indicates the existence of false-positive predictions. The developed framework provides a deep learning-based baseline for automated landslide segmentation using integrated multispectral and topographic information and establishes a foundation for further investigation of advanced Transformer- and attention-based architectures.
Keywords: — Landslide Segmentation, U-Net, Remote Sensing, Sentinel-2, Digital Elevation Model (DEM), Deep Learning

1. INTRODUCTION

One of the most important threats to hilly areas and environmentally sensitive areas is Landslides which can damage infrastructure, ecosystems, transportation systems, and human settlements. Accurate delineation and identification of landslide affected areas is therefore important for effective hazard assessment, hazard mitigation planning and disaster management. Traditional approaches for landslide mapping are based on field surveys, geological investigations and manual interpretation of remotely sensed images, which can be time consuming and labor-intensive, and are difficult to apply in a systematic way over large geographical areas. The development of satellite Earth Observation data sources has created a possibility to automatically map landslides, and to acquire spatially detailed information over large areas.

In recent years, with the rapid advancement of remote sensing technology, multispectral satellite images have been used in automated landslide detection and semantic segmentation with the aid of deep learning. Sentinel-2 images offer multispectral data which can be used to identify changes in land surface properties in landslide affected areas. However, spectral data may not be enough to characterize the terrain conditions that lead to landslides. Some additional information on the elevation and surface steepness can be derived from topographic data, e.g. Digital Elevation Model (DEM), terrain slope. This integration can then be used to create a more complete picture of the terrain for segmenting the landslides at the pixel level. The Landslide4Sense dataset is suitable for this purpose, as it is a combination of multispectral Sentinel-2 data and slope and DEM data in a multimodal input structure.

Traditional machine learning, Convolutional neural networks (CNNs), and more recent Transformer-based models are all methods used for detecting landslides. While geospatial conditioning factors are well-suited for inclusion in machine-learning models for susceptibility analysis, the models also need to learn spatial and contextual features from image data to delineate landslides at a pixel level. The U-Net architecture, which consists of an encoder-decoder module with omission connections, has shown great potential in semantic segmentation by performing hierarchical feature extraction and retaining spatial information during the decoding process. However, systematic evaluation of the performance of a traditional U-Net approach for a multimodal 14-channel landslide dataset composed of multispectral and topographic data is needed. In addition, accurate segmentation of landslides is difficult because of the irregularity of landslide boundaries, broken landslide regions, spectral similarity of the landslides to the surrounding terrain and variation in the terrain characteristics.

To tackle these issues, this paper presents a U-Net-based semantic segmentation framework for automatic landslide delineation using the Landslide4Sense dataset. The proposed framework takes as input a 14-channel patch consisting of twelve Sentinel-2 multispectral bands, terrain slope and Digital Elevation Model (DEM) information. The input patch has a spatial dimension of 128 × 128 pixels, while the corresponding binary ground-truth mask distinguishes landslide and non-landslide regions. The U-Net model performs hierarchical feature extraction by means of an encoder-decoder architecture and adopts skip connections to retain the spatial information required for accurate pixel-level segmentation. The model is trained and appraised using Intersection over Union (IoU), F1-score, Precision, Recall and Accuracy. The experimental results demonstrate an IoU of 0.4300, an F1-score of 0.6014, a Precision of 0.5051, a Recall of 0.7431, and an Accuracy of 0.9814.

The main contributions of this work are summarized as follows:

1. Expansion of a U-Net-based semantic segmentation outline for automated pixel-level landslide delineation using multispectral and topographic remote sensing information.
2. Integration of 14-channel multimodal input data including 12 Sentinel-2 multispectral bands, terrain slope and DEM information provides a complete representation of landslide related surface and terrain characteristics.
3. Implementation and evaluation of the U-Net architecture using the Landslide4Sense benchmark dataset containing 128 × 128 pixel image patches and corresponding binary landslide masks.
4. Quantitative evaluation of landslide segmentation performance using IoU, F1-score, Precision, Recall, and Accuracy, together with qualitative analysis of predicted segmentation masks.
5. Establishment of a U-Net performance baseline for subsequent investigation and comparative analysis of Transformer- and attention-based architectures for landslide segmentation.

The remainder of this paper is organized as follows. Section II presents a review of existing research on remote sensing-based landslide detection, machine learning and deep learning approaches, U-Net-based segmentation, and multispectral-topographic data integration. Section III discusses the Landslide4Sense dataset, multimodal input representation, preprocessing of data, dataset preparation, U-Net architecture, and evaluation methodology. The proposed work, experimental implementation, training process, quantitative results, qualitative segmentation results and observations on performance are described in Section IV. Section V concludes the paper with a summary of the key findings and potential directions for future research. Finally, the references used in this study are presented.

2. Literature Review

A report introduced the Landslide4Sense benchmark dataset and evaluated multiple deep learning models for landslide detection using remote sensing data [1]. The dataset combines multi-spectral Sentinel-2 data with topographic parameters such as slope and digital elevation model (DEM) and is accompanied by per-pixel landslide annotations. Their work offers a standardized benchmark for the evaluation of deep learning-based landslide segmentation and demonstrates the advantage of combining spectral and topographic information. This work is particularly relevant to the present study because the same Landslide4Sense dataset and multimodal input structure are used for evaluating the proposed U-Net framework.

An improved U-Net architecture proposed for landslide area segmentation by adding improved feature extraction and attention-based mechanisms into the conventional encoder-decoder structure [2]. Their approach was designed to advance the illustration of landslide features and the transfer of spatial information thru the network. The results showed improved segmentation performance compared to the original U-Net architecture. However, the addition of further architectural modules increases the model complexity whereas the present study aims to evaluate the conventional U-Net architecture as a baseline for landslide segmentation.

Similarly, U-Net-pyramid based semantic segmentation framework presented for landslides detection from remote sensing data. [3] The method consists of pyramid-based feature extraction and object-based image analysis for better representation of the landslide regions at different scales. The study suggested that the U-Net with multi-scale contextual information can improve the performance of landslide segmentation. Nevertheless, the proposed framework is an extension of the conventional U-Net architecture with additional feature processing components, while this study aims to explore the performance of a standard U-Net model with multimodal multispectral and topographic data.

An attention Swin Transformer U-Net for landslide segmentation reported for remote sensing images [4]. They used a U-Net segmentation framework with a Swin Transformer based feature extraction and attention mechanism to improve the model's ability to capture spatial details and larger contextual relationships. Their work shows the potential of Transformer-based and attention-based architectures to model complex landslide patterns and improve boundary representation. But the increased architectural complexity requires setting a conventional U-Net baseline first under the same dataset and evaluation conditions.

Another work reported the D2FLS-Net, a dual-stage DEM-guided fusion Transformer for landslide segmentation [5]. In this model, the DEM information was directly incorporated into the feature extraction and fusion process and Transformer-based modules were used to extract the contextual information. This work shows the potential of incrementally integrating topographical information into deep learning architectures for better delineation of landslides. However, the proposed framework is considerably more complex than a conventional encoder-decoder network, providing motivation for evaluating how effectively a standard U-Net can utilize the same type of multimodal information.

Other model, MCFI-Net, a multi-scale cross-layer feature interaction network for landslide segmentation proposed for remote sensing imagery [6]. The model was designed to improve the interaction between features extracted at different network levels and to enhance the representation of landslides with varying spatial scales. Their results show that multi-scale feature interaction can improve the segmentation of complex landslide regions. However, the added feature-interaction mechanisms increase the architectural complexity over the conventional U-Net, and a baseline comparison is needed before assessing the advantages of advanced architectures.

Similarly, other report presented the discovery of landslides by fusing spectral, textural and topographical information with deep learning [7]. Their study showed that the fusion of complementary sources of information can help to better represent the characteristics of landslides that may not be sufficiently distinguishable using the spectral information alone. The results encourage the use of multimodal remote sensing information for landslide analysis and are particularly relevant for the fusion of Sentinel-2 data with terrain-related features. The study reveals with the fusion of multisource features for landslide detection, while the present work assesses a standardized 14-channel input of Sentinel-2 bands, slope and DEM, within the framework of a U-Net segmentation.

A level-set-loss-guided semantic segmentation framework employed for landslide abstraction, named LSL-SS-Net [8]. The method solved the problem of difficulty in precisely representing the boundaries of landslides and objects with different spatial scales by introducing a special level-set-based loss function into semantic segmentation networks. The study emphasizes the importance of boundary preservation and spatial consistency in landslide segmentation. However, the proposed approach introduces a specialized loss strategy, while the present study first establishes the performance of a conventional U-Net architecture and provides a baseline for subsequent comparison with more advanced segmentation models.

Table 1: Literature Review Discoveries

S. No.	Author 's Name and Year	Main Perception	Discoveries	Limitation
1	Ghorbanzadeh et al. (2022) [1]	Benchmark Dataset and Deep Learning Models for Landslide4Sense	Proposed a multimodal benchmark with Sentinel 2, DEM and slope data and tested 11 deep learning segmentation models for landslide detection.	It provides a benchmark and several baseline models, but it does not primarily focus on enhancing the typical U-Net design.
5	Li et al. (2025) [2]	Enhanced U-Net for Landslide segmentation	It uses dilated convolution, EMA attention, and PAG-based skip connections, which improves the baseline U-Net in terms of mIoU, Precision, Recall and F1-score.	The extra modules make the architecture more complex than the traditional U-Net architecture.
2	Kaushal et al. (2024) [3]	UNet-Pyramid Landslide Segmentation	Showed improved semantic segmentation using remote sensing data.	Additional pyramid and object-based components increase architectural complexity compared with conventional U-Net.
3	Liu et al. (2024) [4]	Swin Transformer U-Net Attention	Integrated Swin Transformer with attention processes to learn better contextual features and retain spatial details for landslide segmentation.	Transformer and attention parts increase complexity and processing costs to the model.
7	Zhao et al. (2025) [5]	DEM-Guided Fusion Transformer (D2FLS-Net)	Proposed a dual-stage Transformer architecture that incorporates DEM information to improve terrain-aware landslide segmentation and border representation.	A transformer-based architecture is way more complicated than a standard U-Net.
8	Liao and Ye (2025) [6]	Multi-Scale Cross-Layer Feature Interaction Network	Multi-Scale Feature Extraction and Cross-Layer Feature Interaction for Landslide Segmentation in Remote Sensing Images.	The model complexity is increased due to additional modules for feature interactions compared with standard encoder-decoder networks.

4	Ren et al. (2024) [7]	Multisource Spectral, Textural and Topographical Data Fusion	The potential of merging spectral, texture and topography features for deep learning-based landslide detection was proved by the integration of multisource spectral, texture and topographic data.	Instead of the traditional baseline evaluation on Landslide4Sense, it focuses on multi-source feature fusion and a specific research area.
6	Hu et al. (2025) [9]	Hierarchical Cross-Attention Landslide Segmentation	Hierarchical Cross-Attention Landslide Segmentation Used cross-attention and hierarchical feature fusion for pixel-level landslide segmentation, with superior performance compared to baseline models in IoU and F1-score.	Designed for submeter optical imagery, based on a more complex attention-based architecture

3. Projected Method

3.1 Overall Research Workflow

The proposed methodology consists of sequential stages including problem identification, dataset acquisition, preprocessing, dataset preparation, multimodal feature integration, U-Net model development, training, and landslide segmentation. The Landslide4Sense dataset is used as the primary source of multispectral and topographic information. The input data are preprocessed and represented in a 14 channel format, including Sentinel-2 spectral bands, slope and DEM data. Finally, the prepared data is given to the U-Net architecture for pixel-wise segmentation of a landslide. [1], [7].

3.2 Dataset and Multimodel Input

The Landslide4Sense dataset is used in this study for pixel-level landslide segmentation. It contains 4,844 image patches of size 128 × 128 pixels with a spatial resolution of approximately 10 m/pixel. The dataset is available in HDF5 (.h5) format and contains both multispectral and topographic information [1].

The input sample contains 14 channels, including 12 Sentinel-2 spectral bands (B1–B12), slope and DEM information. The ground-truth data are binary segmentation masks that distinguish landslide pixels from non-landslide pixels. The combination of spectral and topographic information provides complementary features for landslide region identification [1], [7].

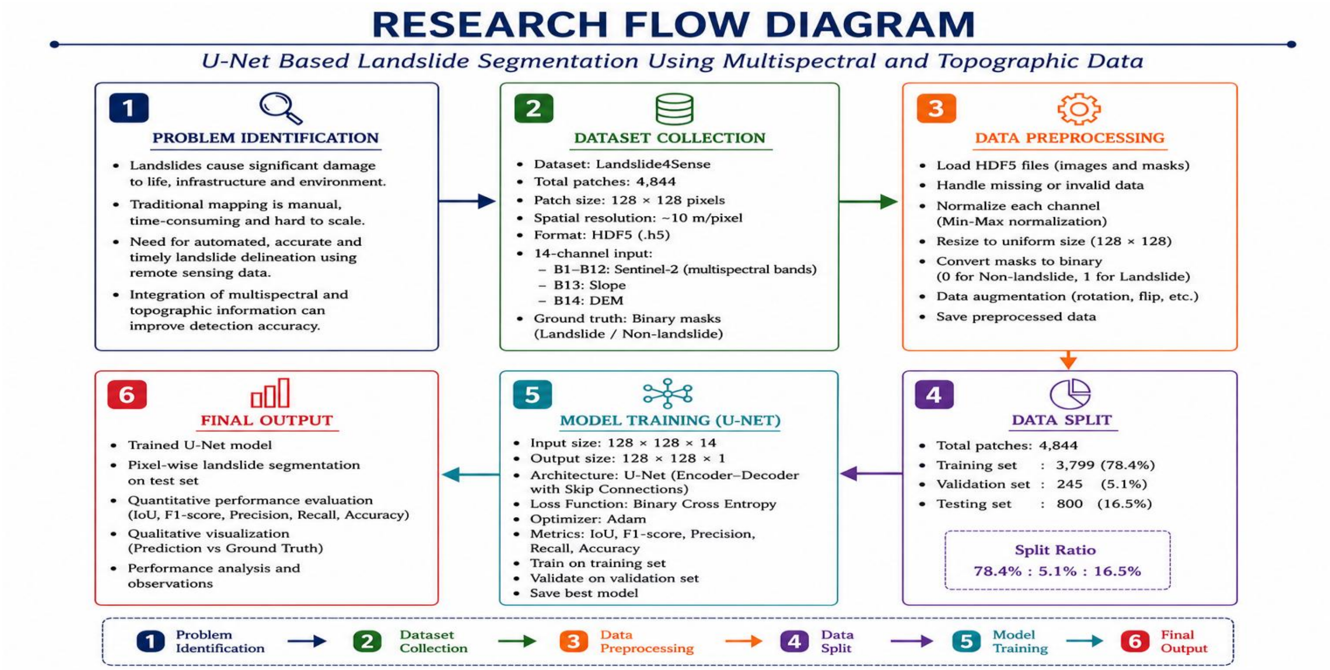

Figure 1: Research Flow Diagram

Table 2: Dataset and Multimodal Input Configuration

Parameter	Description
Dataset	Landslide4Sense
Number of image patches	4,844
Patch dimensions	128 × 128 pixels
Spatial resolution	~10 m/pixel
Data format	HDF5 (.h5)
Total input channels	14
Spectral information	Sentinel-2 B1–B12
Topographic information	Slope and DEM
Input tensor	128 × 128 × 14
Ground-truth	Binary landslide/non-landslide mask
Output tensor	128 × 128 × 1

The 14-channel multimodal input representation and the overall U-Net configuration are illustrated in Fig. 2. The input tensor is the concatenation of the Sentinel-2 spectral bands, slope and DEM, and is fed to the encoder-decoder architecture. The encoder excerpts spatial features in a hierarchical manner, while the decoder reconstructs the 3-D resolution by means of the up-convolution and skip connections. The final layer produces a pixel-level binary segmentation mask.[1] [2]

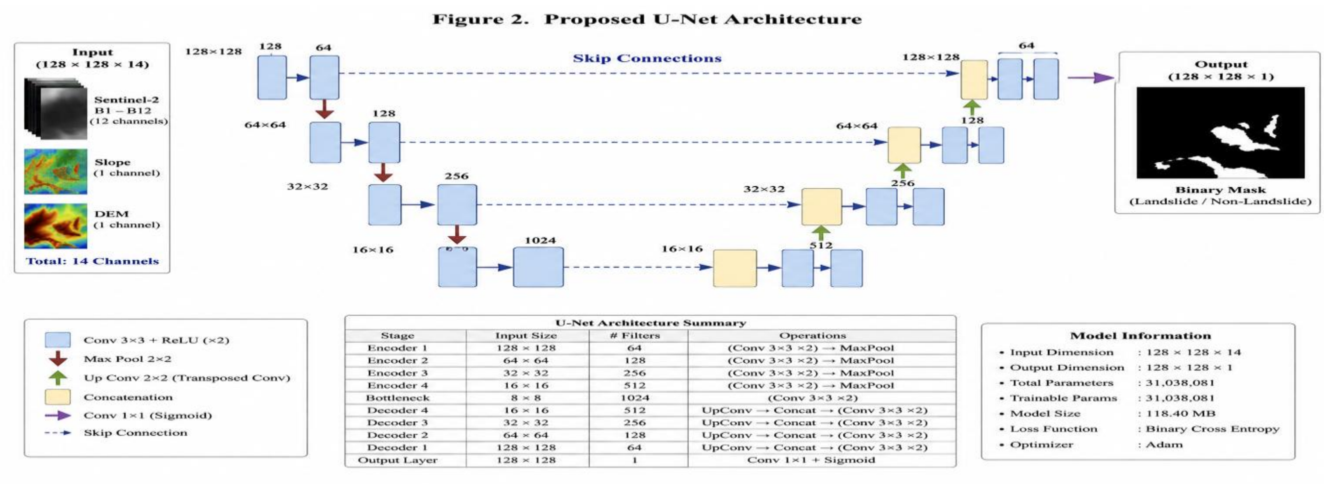

Figure 2: Proposed U-Net architecture with the 14-channel Landslide4Sense input and pixel-level landslide segmentation output.

The implemented U-Net receives an input of 128 × 128 × 14 and produces an output of 128 × 128 × 1. The network contains 31,038,081 trainable parameters with an approximate model size of 118.40 MB. The detailed stage-wise configuration of the encoder, bottleneck, decoder and output layer is depicted in Table 2. [9]

Table 3: U-Net Architecture Configuration

Stage	Input Size	Number of Filters	Operations
Encoder 1	128 × 128	64	Conv 3×3 × 2 → MaxPool
Encoder 2	64 × 64	128	Conv 3×3 × 2 → MaxPool
Encoder 3	32 × 32	256	Conv 3×3 × 2 → MaxPool
Encoder 4	16 × 16	512	Conv 3×3 × 2 → MaxPool
Bottleneck	8 × 8	1024	Conv 3×3 × 2
Decoder 4	16 × 16	512	UpConv → Concat → Conv 3×3 × 2
Decoder 3	32 × 32	256	UpConv → Concat → Conv 3×3 × 2
Decoder 2	64 × 64	128	UpConv → Concat → Conv 3×3 × 2
Decoder 1	128 × 128	64	UpConv → Concat → Conv 3×3 × 2
Output Layer	128 × 128	1	Conv 1×1 → Sigmoid

3.3 U-Net Based Segmentation Framework

The projected method employs a U-Net-based encoder-decoder framework for pixel-level landslide dissection. The encoder is a series of convolutional operations, which obtains hierarchical representations from the input features. The spatial down sampling enables the network to obtain higher-level contextual information [2,3].

The decoder reconstructs the spatial representation thru progressive up sampling. Skip connections between corresponding stages of the encoder and decoder are used to integrate the spatial details that are lost during the down sampling process. This is especially important for landslide delineation, in which the accurate delineation of irregular boundaries and small affected regions relies on both the contextual and the spatial information [2,4].

The last step is the reconstruction of the feature representation into a one channel probability map through a 1 × 1 convolution with sigmoid activation. The probability value of each pixel is thus assigned to the class of landslide, and a binary landslide mask can be generated.

The elaborate architectural configuration such as feature dimensions, number of filters and operations per stage is described in Table 3 and the complete network structure is shown in Fig. 2, [8].

4. Proposed Work

4.1 Experimental Dataset Preparation

The Landslide4Sense dataset was prepared for the experimental evaluation by loading the available HDF5 files and organizing the required input channels and corresponding ground-truth masks. The multispectral and topographic information was arranged into the required multimodal input representation, while the associated binary masks were prepared as the target outputs for pixel-level landslide segmentation. The prepared samples were subsequently organized into training, validation, and testing subsets for model development and independent evaluation [12,19].

Table 4: Dataset Distribution

Dataset	Number of Patches	Percentage
Training	3,799	78.4%
Validation	245	5.1%
Testing	800	16.5%
Total	4,844	100%

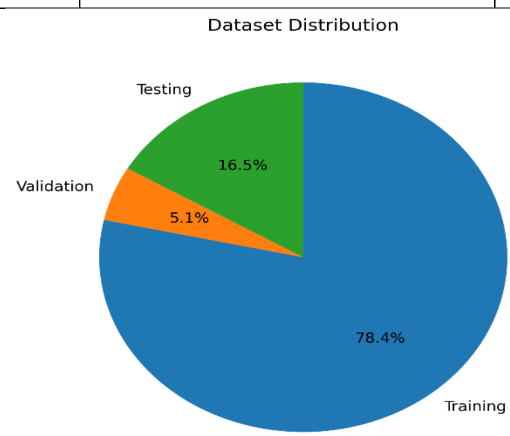

Figure 3: Dataset Distribution Pie Chart

For experimental evaluation, the complete Landslide4Sense dataset was divided into training, validation, and testing subsets. A total of 3,799 patches were used for training, 245 patches for validation, and 800 patches were retained for final test result. The corresponding delivery is offered in Table 4 and Fig. 3, [21].

4.2 Dataset Exploration and Input Analysis

We analyzed some representative samples of the Landslide4Sense dataset to analyze the characteristics of the multimodal data used for landslide segmentation before training the model. Fig. 4 presents four representative samples containing Sentinel-2 pseudo-RGB imagery, slope information, DEM-derived elevation information, and the corresponding ground-truth landslide masks.

Pseudo-RGB images are a graphical representation of surface characteristics Slope maps are a representation of terrain steepness The DEM layers provide complementary elevation information that describes the topographic structure of each sample. The ground-truth masks indicate the spatial distribution and extent of the annotated landslide regions.

The samples demonstrate considerable spatial variation in the appearance and distribution of landslide regions. The ground-truth masks also show that the target regions can vary substantially in size, shape, and spatial location across different image patches. The simultaneous presence of spectral, slope, and elevation information provides complementary spatial characteristics that can be utilized by the segmentation model for pixel-level landslide delineation [20].

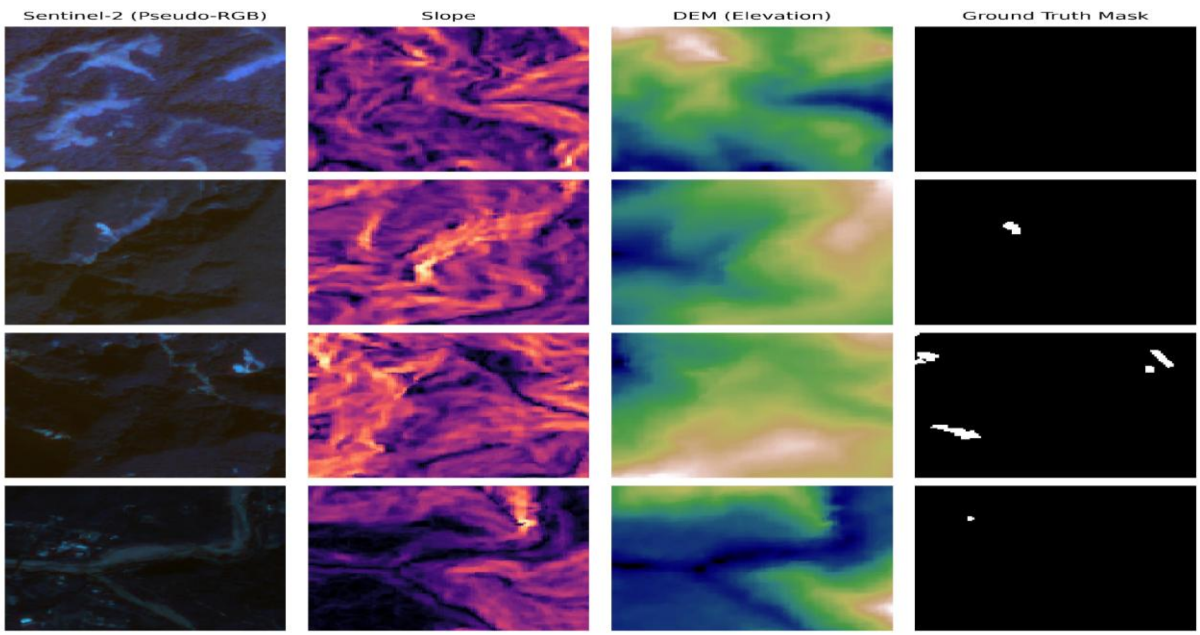

Figure 4: Representative Landslide4Sense samples showing Sentinel-2 pseudo-RGB imagery, slope, DEM-derived elevation, and corresponding ground-truth landslide masks.

4.3 U-Net Model Training and Experimental Configuration

Following dataset preparation, the U-Net model was trained using the training subset, while the authentication subset was used to display model performance during the training process. The testing subset was kept separate from the training procedure and was used for ultimate final performance evaluation. The experimental configuration was kept fixed to provide a reproducible training environment.

The model was implemented using Tensorflow/Keras (v2.10+) and trained using Adam optimizer with a learning rate of 0.0001. The batch size was 16 and the maximum number of training epochs was 50 . Binary cross-entropy was used as loss-function while a sigmoid activation was used at the output layer for binary pixel level segmentation. Early stopping with a patience of 10 epochs was added to avoid unnecessary training once the validation performance stopped improving. The complete experimental configuration is exposed in Table 5, [14].

Table 5: Experimental Training Configuration

Parameter	Configuration
Model	U-Net
Input size	128 × 128 × 14

Optimizer	Adam
Loss function	Binary Cross-Entropy
Batch size	16
Learning rate	0.0001
Maximum epochs	50
Early stopping	Patience = 10
Activation function	Sigmoid
Framework	TensorFlow / Keras (v2.10+)
Hardware	CPU (Intel/AMD x64 Native Windows)

Parameter Configuration Model U-Net Input size 128 × 128 × 14 Optimizer Adam Loss function Binary Cross-Entropy Batch size 16 Learning rate 0.0001 Maximum epochs 50 Early stopping Patience = 10 Activation function Sigmoid Framework TensorFlow / Keras (v2.10+) Hardware CPU (Intel/AMD x64 Native Windows) During the training the loss and IoU values for the training and validation subsets were logged at the end of each epoch. These learning curves were examined to evaluate the model's optimization behavior, convergence and possible overfitting.

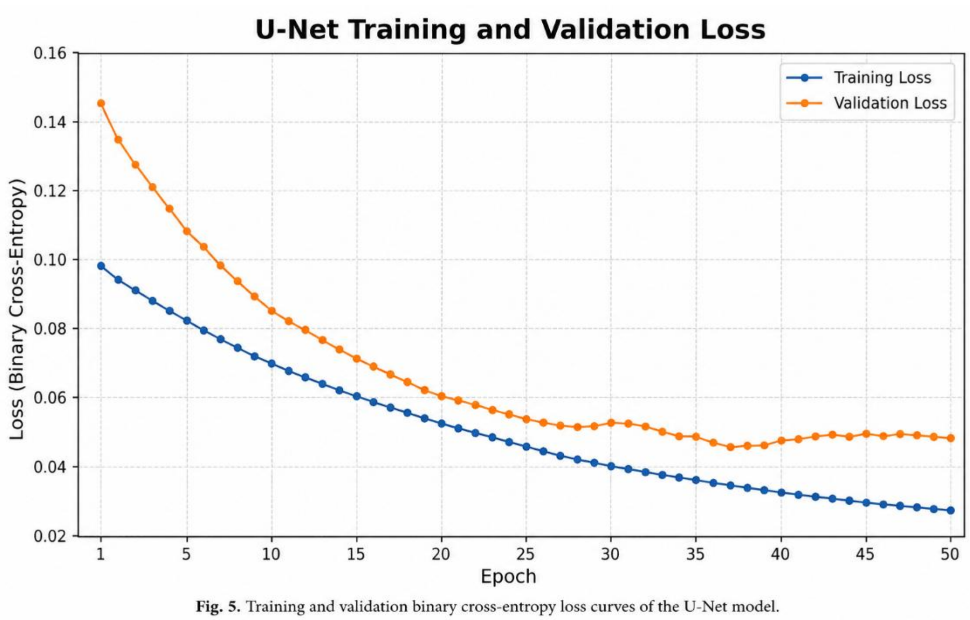

Fig. 5: Training and validation loss curves of the U-Net model.

The training loss decreases slowly during the training. The validation loss decreases significantly in the first few epochs and then stays relatively stable with small fluctuations. The separation between the training and validation curves toward the later epochs indicates a degree of generalization gap, although the validation loss does not show a sustained sharp increase.

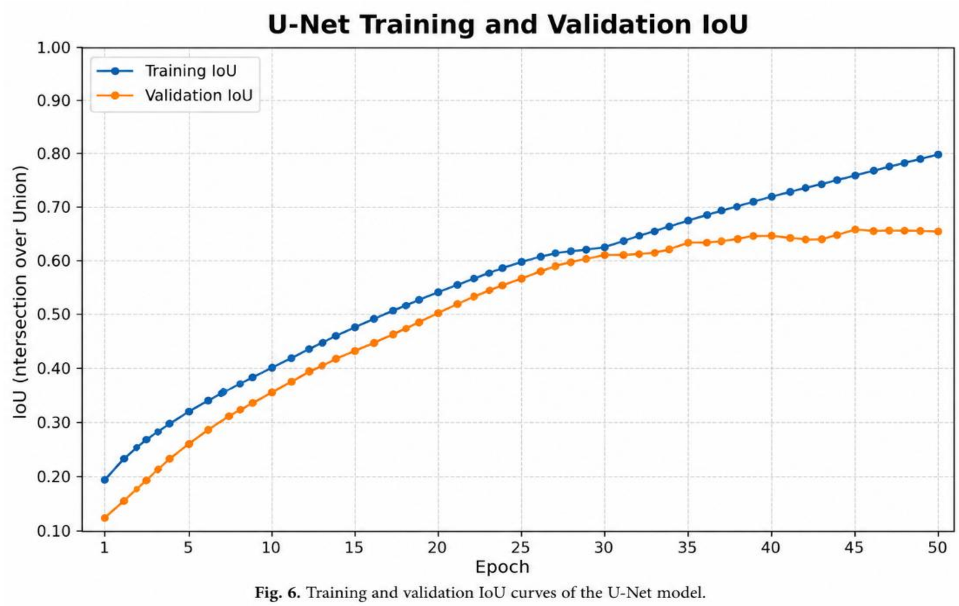

Fig. 6: Training and validation IoU curves of the U-Net model.

The IoU curves offer an alternative view of the model's learning behavior in segmentation by quantifying the overlap between predicted and reference regions. The rising tendency of the IoU indicates the incremental advancement of the spatial segmentation, and the flattening of the validation IoU indicates the model is close to convergence.[2] [3]

4.4 Quantitative Concert Assessment

The projected U-Net model performance was evaluated on the testing dataset using Intersection over Union (IoU), F1-score, precision, recall, and accuracy. These metrics provide complementary information regarding the overlap between the predicted and ground-truth landslide areas and the model's pixel-level sorting presentation. The obtained results are shown in Table 6, [2,3].

Table 6: U-Net Segmentation Performance

Metric	**U-Net**
IoU	0.4300
F1-score	0.6014
Precision	0.5051
Recall	0.7431
Accuracy	0.9814

U-Net model achieved an IoU value of 0.4300 and F1-score of 0.6014, which indicates moderate overlap between the predicted and reference landslide locations. The recall value of 0.7431 showed that a large proportion of the landslide pixels was properly detected. But the 0.5051 accuracy revealed that there were false positive predictions with certain non-landslide pixels misclassified as landslide regions.

The overall accuracy of 0.9814 demonstrates that most pixels were classified correctly. However, accuracy alone is not sufficient to characterize the segmentation performance because pixel-level landslide datasets may contain substantially more non-landslide pixels than landslide pixels. Therefore, IoU and F1-score are more informative metrics to evaluate the model's ability to delineate the real landslide regions [17].

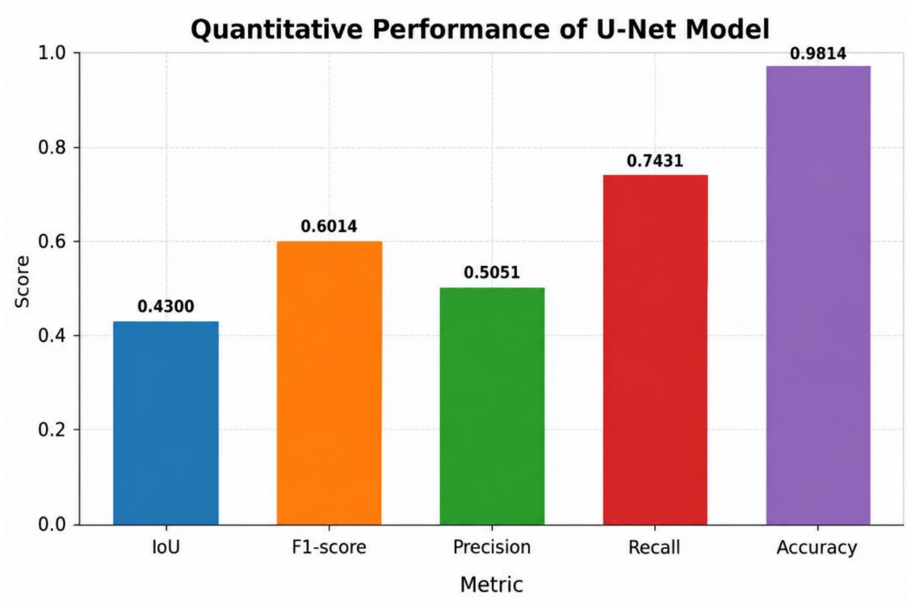

Fig. 7: Quantitative performance of U-Net model on IoU, F1-score, Precision, Recall and Accuracy.

4.5 Qualitative Results and Error Analysis

The predicted segmentation masks of the proposed U-Net model were compared with the corresponding ground-truth masks for representative samples of the testing dataset to further evaluate the qualitative presentation of the prototypical model. This analysis delivers a visual assessment of the model's ability to identify the spatial extent and structure of the landslide regions beyond the numerical metrics reported in Section 4.4. Fig. 8 presents four representative test samples using the Sentinel-2 pseudo-RGB image, ground-truth mask, and corresponding U-Net prediction [2, 4].

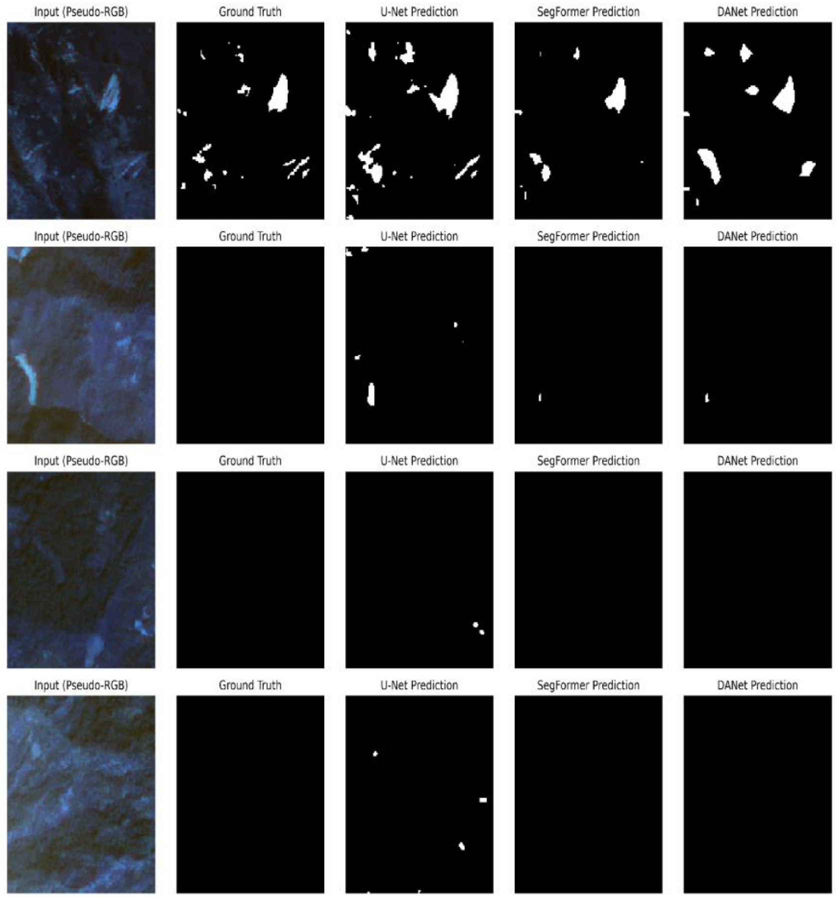

Fig. 8: Qualitative comparison of Sentinel-2 pseudo-RGB inputs, ground-truth landslide masks, and U-Net predictions on representative test samples.

The first sample includes several landslide areas of varying sizes and forms that are marked and annotated. When comparing with the ground truth mask, the U-Net prediction shows several of the major landslide regions, and reproduces the general spatial locations of these landslides, but also shows some additional landslides that might also be considered to be landslides. This means that the model can identify the key structures of a landslide, but also has the potential to label adjacent pixels as landslide.

In the second sample, there is no clearly marked landslide area in the ground truth mask, while the landslide area given by U-Net has only a small number of isolated regions. These predictions are false positive where the pixels which are not landslides are classified as landslides. The same behaviour can be seen for the third and fourth sample with no obvious landslide area in the ground-truth masks and small isolated predictions in the model.

The qualitative observations are consistent with the quantitative results as shown in Table 6. In particular, it has a higher recall of 0.7431 and a lower precision of 0.5051, meaning that there are false positive predictions. Qualitative examples demonstrate this behavior and demonstrate that the model's segmentation performance is related to the spatial properties of the regions in the input [22].

5. Error Analysis

To further analyse the observed segmentation behaviour, we conducted an error analysis by linking the predicted segmentation masks to ground-truth masks. The study focuses on the disparities in geographic extent between the anticipated and reference areas of landslides, including extra predicted regions, missing regions, conflicts in the boundaries and fragmented forecasts [2] [4].

A false positive region is created when the prototype model predicts landslide pixels that the ground-truth mask labels as non-landslide. As seen in Fig. 9, additional isolated predicted locations are detected in samples when there are no equivalent landslip regions in the ground-truth mask. These false positive predictions lead to a decrease in accuracy. The results show that certain non-landslide locations may be misclassified as landslide.

False negative areas are the landslip pixels in the ground-truth mask that are not identified by the U-Net model. Such variances may be detected when the anticipated mask covers just a portion of the reference landslip area. This is especially important for tiny, fragmented, or complicated landslip areas where the overlap between anticipated and reference masks may be reduced by the partial detection.

Boundary inconsistencies are also noticed if the anticipated landslip extent does not exactly mirror the limits of the ground-truth area. Additionally, the predicted results are also fragmented, which are shown as isolated/discontinuous patches relative to the associated reference mask.

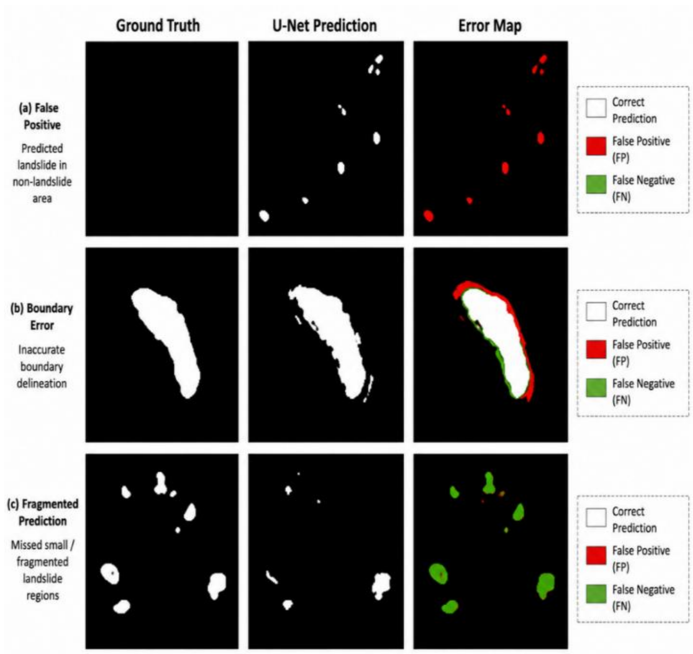

Fig. 9. Representative U-Net segmentation errors showing differences between ground-truth and predicted landslide regions.

Fig. 9. Representative U-Net segmentation errors showing differences between ground-truth and predicted landslide regions.

The error analysis reveals that the U-Net model can detect a number of key landslip structures but may provide some isolated predictions and spatial variations in difficult samples. The disparities found further highlight the difficulties in defining the borders of the landslip and recognising tiny or fragmented target areas at the pixel level.

To summarise, the qualitative error analysis complements the quantitative assessment by showing the geographical aspects of its segmentation mistakes as well as the numerical performance of the U-Net model. These findings provide useful paths for future enhancements of the segmentation framework, especially via better feature representations, better loss function design and more complicated segmentation structures.

6. Conclusion

We suggested a pixel-level landslide segmentation deep learning system based on U-Net utilising the multimodal Landslide4Sense dataset in this paper. The suggested technique combines the spectral information from Sentinel-2 with the topographic information from slope and DEM to generate a 14-channel input representation for the landslip delineation. The U-Net model was trained and assessed using IoU, F1-score, precision, recall and accuracy. The model attained the IoU of 0.4300, F1-score of 0.6014, precision of 0.5051, recall of 0.7431 and accuracy of 0.9814 . The findings reveal that the prototype model may be able to detect a large number of the landslip pixels, which is reflected in its recall. The lesser precision shows the prevalence of false-positive predictions. Qualitative and error analyses further indicate that the model can identify big landslip regions but may have difficulty detecting minor, fragmented, and border areas. To summarise, the proposed architecture exhibits the potential of U-Net for automatic segmentation of landslides using multimodal remote sensing data. These reported faults leave the door open for future improvements via better feature representation, loss functions and more advanced segmentation structures.

Future Work

Future study will concentrate on improving the proposed U-Net based landslip segmentation framework by exploring more sophisticated deep learning architectures and feature representation approaches for better pixel-level delineation. Further investigations may examine increased multimodal integration of Sentinel-2 spectral information with slope and DEM data for improved identification of tiny, fragmented and complicated landslip sites. Further research on optimising loss functions and training procedures may be carried out to decrease false positive predictions and improve IoU, F1-score and accuracy. Moreover, the comparative evaluation using SegFormer and DANet designs might be an indication of the bigger evaluation of the model performance in the same experimental settings. Future research might also explore the generalisation of the models on diverse geographical locations and remote sensing datasets to determine their robustness and practical application in automated landslip mapping.[15] [16].

REFERENCES

1. O. Ghorbanzadeh, Y. Xu, P. Ghamisi, M. Kopp, D. Kreil, Landslide4Sense: Reference Benchmark Data and Deep Learning Models for Landslide Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60, (2022) 1–17. https://doi.org/10.1109/TGRS.2022.3215209
2. G. Li, K. Li, G. Zhang, K. Pan, Y. Ding, Z. Wang, C. Fu, Z. Zhu, A landslide area segmentation method based on an improved UNet. *Scientific Reports*, 15(1), (2025) 11852. https://doi.org/10.1038/s41598-025-94039-5
3. A. Kaushal, A.K. Gupta, V.K. Sehgal, A semantic segmentation framework with UNet-pyramid for landslide prediction using remote sensing data. *Scientific Reports*, 14, (2024) 30071. https://doi.org/10.1038/s41598-024-79266-6
4. B. Liu, W. Wang, Y. Wu, X. Gao, Attention Swin Transformer UNet for Landslide Segmentation in Remotely Sensed Images. *Remote Sensing*, 16(23), (2024) 4464. https://doi.org/10.3390/rs16234464
5. C. Zhao, L. Li, Y. Wang, X. Li, C. Xu, Y. Song, D. Ren, C. Xiao, D2FLS-Net: Dual-Stage DEM-guided Fusion Transformer for landslide segmentation. *PLOS ONE*, 20(11), (2025) e0337412. https://doi.org/10.1371/journal.pone.0337412
6. J. Liao, L. Ye, MCFI-Net: Multi-Scale Cross-Layer Feature Interaction Network for Landslide Segmentation in Remote Sensing Imagery. *Electronics*, 14(21), (2025) 4190. https://doi.org/10.3390/electronics14214190
7. Y. Ren, J. Ma, J. Liu, X. Deng, G. Zhang, H. Guo, Enhancing deep learning-based landslide detection from open satellite imagery via multisource data fusion of spectral, textural, and topographical features: a case study of old landslide detection in the Three Gorges Reservoir Area (TGRA). *Geocarto International*, 39(1), (2024) 2421224. https://doi.org/10.1080/10106049.2024.2421224
8. Y. Yang, Z. Miao, X. Li, H. Zhang, S. Chen, LSL-SS-Net: level set loss-guided semantic segmentation networks for landslide extraction. *GIScience & Remote Sensing*, 61(1), (2024) 2416727. https://doi.org/10.1080/15481603.2024.2416727
9. W. Hu, G. Sun, X. Zeng, B. Tong, Z. Wang, X. Wu, P. Song, Hierarchical cross attention achieves pixel precise landslide segmentation in submeter optical imagery. *Scientific Reports*, 15(1), (2025) 21933. https://doi.org/10.1038/s41598-025-08695-8
10. S. Mihu, K.K.S. Tomar, A. Kumar, P.P. Choudhari, A. Raju, M. Gentilucci, M. Barbieri, Machine learning-based landslide susceptibility modeling in the Dibang Valley, NE India. *Earth Systems and Environment*, (2026). https://doi.org/10.1007/s41748-026-01036-3
11. M.A. Cifci, X. Hu, B. Öney, S. Misak, H. Moayedi, H.A. Dehrashid, Prediction of landslide susceptibility through ANN models optimized by evolutionary algorithms. *Scientific Reports*, 16, (2026) 9471. https://doi.org/10.1038/s41598-026-39458-8

12. C.F. Barella, J.L. Zêzere, N.F. Fernandes, Validation and prediction challenges in data-driven landslide susceptibility mapping: insights from random sampling of training and test data in machine learning models. *Bulletin of Engineering Geology and the Environment*, 85, (2026) 253. https://doi.org/10.1007/s10064-026-04911-5
13. W. Yang, Q. Ge, Z. Tao, D. Xu, Y. Wang, Z. Hao, Landslide susceptibility on the Qinghai-Tibet Plateau: Key driving factors identified through machine learning. *Journal of Geographical Sciences*, 36(1), (2026) 199–218. https://doi.org/10.1007/s11442-026-2444-6
14. Z. Liang, X. Zhang, Z. Liu, H. Huang, Y. Liu, Enhancing landslide susceptibility mapping through a hybrid model utilizing bivariate methods and convolutional neural networks. *Frontiers in Earth Science*, 14, (2026) 1792825. https://doi.org/10.3389/feart.2026.1792825
15. C. Liu, Q. Zhang, W. Yu, G. Li, Landslide Segmentation in High-Resolution Remote Sensing Images: The Van–UPerAttnSeg Framework with Multi-Scale Feature Enhancement. *Remote Sensing*, 17(7), (2025) 1265. https://doi.org/10.3390/rs17071265
16. N. Sharma, M. Saharia, G.V. Ramana, High resolution landslide susceptibility mapping using ensemble machine learning and geospatial big data. *CATENA*, 235, (2024) 107653. https://doi.org/10.1016/j.catena.2023.107653
17. S. Xu, Y. Song, X. Hao, A Comparative Study of Shallow Machine Learning Models and Deep Learning Models for Landslide Susceptibility Assessment Based on Imbalanced Data. *Forests*, 13(11), (2022) 1908. https://doi.org/10.3390/f13111908
18. S. Zhang, L. Bai, Y. Li, W. Li, M. Xie, Comparing Convolutional Neural Network and Machine Learning Models in Landslide Susceptibility Mapping: A Case Study in Wenchuan County. *Frontiers in Environmental Science*, 10, (2022) 886841. https://doi.org/10.3389/fenvs.2022.886841
19. G. Tang, Z. Fang, Y. Wang, Global landslide susceptibility prediction based on the automated machine learning (AutoML) framework. *Geocarto International*, 38(1), (2023) 2236576. https://doi.org/10.1080/10106049.2023.2236576
20. D. Sun, D. Chen, J. Zhang, C. Mi, Q. Gu, H. Wen, Landslide Susceptibility Mapping Based on Interpretable Machine Learning from the Perspective of Geomorphological Differentiation. *Land*, 12(5), (2023) 1018. https://doi.org/10.3390/land12051018
21. Z. Guo, F. Guo, Y. Zhang, J. He, G. Li, Y. Yang, X. Zhang, A python system for regional landslide susceptibility assessment by integrating machine learning models and its application. *Heliyon*, 9(11), (2023) e21542. https://doi.org/10.1016/j.heliyon.2023.e21542
22. H. Yu, W. Pei, J. Zhang, G. Chen, Landslide Susceptibility Mapping and Driving Mechanisms in a Vulnerable Region Based on Multiple Machine Learning Models. *Remote Sensing*, 15(7), (2023) 1886. https://doi.org/10.3390/rs15071886
23. C. Chen, L. Fan, CNN-LSTM-Attention Deep Learning Model for Mapping Landslide Susceptibility in Kerala, India. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-3/W1, (2022) 25–30. https://doi.org/10.5194/isprs-annals-X-3-W1-2022-25-2022
24. J.M. Habumugisha, N. Chen, M. Rahman, M.M. Islam, H. Ahmad, A. Elbeltagi, G. Sharma, S.N. Liza, A. Dewan, Landslide Susceptibility Mapping with Deep Learning Algorithms. *Sustainability*, 14(3), (2022) 1734. https://doi.org/10.3390/su14031734
25. C. Huang, F. Li, L. Wei, X. Hu, Y. Yang, Landslide Susceptibility Modeling Using a Deep Random Neural Network. *Applied Sciences*, 12(24), (2022) 12887. https://doi.org/10.3390/app122412887