https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

Enhanced Landslide Detection from Remote Sensing Imagery Using an Attention-Optimized UNet-CBAM

Haiyang Li[1], Jing Wang[2], Shuguang Wu[3], Yawei Wang[4,5], Yehemin Gao[1], Guigen Nie[6]

[1]Jiangsu Hydraulic Research Institute, Nanjing, China.
[2]Guangzhou Institute of Technology, Xidian university, Guangzhou, China
[3]School of Electrical Engineering, Naval University of Engineering, Wuhan, China;
[4]The First Surveying and Mapping Institute of Hunan Province, Changsha 410114, China;
[5]Hunan Engineering Research Center of 3D Real Scene Construction and Application Technology, Changsha 410114, China;
[6]GNSS Research Center, Wuhan University, Wuhan, China;

Correspondence to: Jing Wang (wangjing01@xidian.edu.cn)

Abstract. Landslide deformation monitoring is crucial for disaster prevention and protecting infrastructure, ecosystems, and lives in vulnerable regions. Traditional methods, though useful, often lack the precision required for complex terrains, limiting their effectiveness in landslide-prone areas. This study presents the UNet-Convolutional Block Attention Module (CBAM) framework, which combines the UNet architecture with CBAM to enhance landslide detection and segmentation in remote sensing imagery. The integration of CBAM improves the model's ability to focus on spatially significant features, leading to more accurate and efficient extraction of landslide-related information. Experimental results demonstrate that the UNet-CBAM outperforms the baseline UNet by 10% in performance over the UNet, with a notable improvement in the Area Under Curve (AUC) metric. The proposed model shows robustness in diverse and challenging landscapes, proving its effectiveness for landslide monitoring. This enhancement offers significant potential for improving early-warning systems, disaster preparedness, and risk management strategies in landslide-prone areas.

1 Introduction

Landslides pose significant threats to human life, infrastructure, and environmental stability, especially in mountainous and geologically unstable regions (Dai et al., 2002; Tehrani et al.,2022; Fang et al., 2023). Monitoring landslide deformations accurately and in real time is crucial for disaster prevention and mitigation, as it allows for timely warnings and protective measures. Given the increased incidence of landslides due to climate change and human activities, reliable monitoring has become an essential aspect of geohazard management. By capturing subtle shifts and precursors to major landslide events, effective monitoring systems provide critical information for safeguarding lives and property.

Various methods are currently employed for landslide deformation monitoring, including traditional displacement sensors, Global Navigation Satellite Systems (GNSS) (Huang et al.,2023; Wang et al., 2021), and Light Detection and Ranging (LiDAR) (Zhang et al., 2024; Fang et al., 2022). These techniques are widely used due to their accuracy and direct measurement

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

capabilities; however, each has its limitations. GNSS provides precise positional data but is restricted to specific measurement points and lacks comprehensive spatial coverage. LiDAR offers high-resolution surface data but can be costly and sensitive to environmental factors, such as vegetation cover and weather conditions, which can obstruct data accuracy.

Image recognition technologies, particularly those based on remote sensing and satellite imagery, have gained increasing attention in landslide monitoring research (Shi et al., 2020). These techniques enable large-scale, continuous observation of landslide-prone areas without direct field measurements, thus complementing traditional methods. Recent developments in machine learning and deep learning have further enhanced the accuracy of image-based landslide monitoring, with several studies demonstrating effective landslide detection and classification in high-resolution satellite and aerial images (Li et al., 2022; Yu et al., 2017; Ghorbanzadeh et al., 2022; Ofli et al., 2023). However, challenges remain, such as variations in lighting, terrain, and seasonal conditions that can impact image quality and recognition performance.

UNet, originally developed for biomedical image segmentation, has proven highly effective for complex image recognition tasks due to its encoder-decoder structure and skip connections, which help preserve spatial information (Devara et al., 2024; Dang et al., 2024; Fu et al., 2023). In the field of remote sensing, UNet has shown potential in segmenting and classifying regions with irregular boundaries, making it a promising tool for landslide deformation monitoring. Zhang et al. proposed a landslide extraction model using multi-scale feature fusion and attention mechanism, which proves that the new model can extract landslides accurately and effectively (Dong et al., 2022). Niu et al. proposed a lightweight landslide detection network based on single-temporal images, which reduced model dependence on original data (Niu et al., 2022). However, applying UNet to landslide imagery still faces limitations in distinguishing fine details under diverse environmental conditions, indicating room for further enhancements to improve recognition accuracy and robustness.

This study proposes an enhanced landslide monitoring method by integrating the CBAM with the UNet architecture, forming the UNet-CBAM. The CBAM module adds both channel and spatial attention mechanisms to UNet, allowing the network to better focus on critical areas within each image. This modification is expected to improve the model's capability to distinguish landslide-prone regions from surrounding terrain, especially under challenging visual conditions. Combining UNet's structure with CBAM's attention mechanisms, the proposed method improves segmentation accuracy for more reliable landslide deformation monitoring.

The remainder of this study is organized as follows: Section 2 presents the methodology, detailing the design and implementation of the UNet-CBAM. Section 3 covers the experiments and results, including evaluations of the proposed model and comparisons with other methods. Section 4 provides a discussion of the experimental findings, emphasizing the advantages

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

and potential limitations of the proposed model. Finally, Section 5 concludes the study, summarizing the key findings and suggesting directions for future research in landslide deformation monitoring.

2 Methodology

This section introduces a UNet-CBAM for landslide deformation monitoring, designed to improve segmentation accuracy in remote sensing images. As shown in Fig. 1, the workflow includes data preprocessing, attention-enhanced segmentation, and post-processing to refine results and reduce noise. This structured pipeline ensures accurate and robust landslide identification across complex terrain.

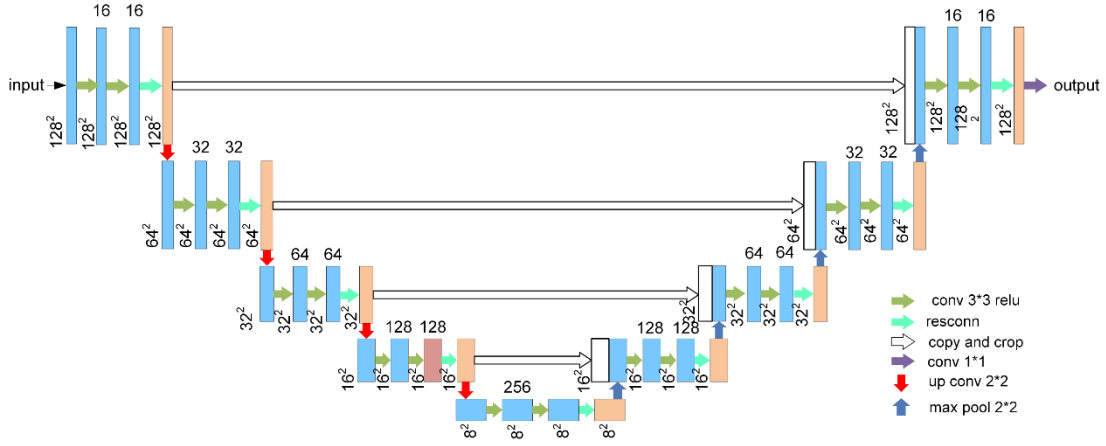


Fig. 1. UNet-CBAM Workflow for Landslide Monitoring

2.1 UNet Architecture

The UNet architecture, originally developed for biomedical image segmentation, has gained significant attention for its ability to segment complex, irregular shapes within images, making it well-suited for landslide deformation monitoring in remote sensing imagery (Meena et al., 2022; Ghorbanzadeh et al., 2021). The UNet employs an encoder-decoder structure with skip connections, allowing it to capture both local and global context, thus preserving critical spatial information while enhancing feature extraction.

The encoder block is responsible for capturing hierarchical feature representations from input images through a series of convolutional and down-sampling layers. In each layer, the encoder progressively captures spatial and contextual information,

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

allowing the model to learn meaningful features at different levels of abstraction (Lu et al., 2023). The encoding process is defined by convolution operations:

$$x' = f(W * x + b) \tag{1}$$

where x represents the input feature map, W is the weight matrix for the convolutional filter, b is the bias term, and f denotes a non-linear activation function such as $ReLU$. This convolution operation allows the network to learn spatial patterns within the image, capturing features relevant to landslide identification, while down-sampling increases receptive field size and reduces spatial dimensions.

The bottleneck layer sits at the core of the UNet, serving as a bridge between the encoder and decoder by distilling critical information into a compressed form. This layer captures the most salient features through a series of convolutional operations, focusing on the essential patterns that define landslide-prone areas (Wei et al., 2023). The bottleneck operation is also governed by convolutional transformations, similar to the encoder:

$$z = f(W_b * x' + b_b) \tag{2}$$

where z is the encoded feature map, W_b and b_b are the weights and bias specific to the bottleneck layer, and f denotes the activation function. By concentrating critical features at the bottleneck, the model captures and compresses essential patterns, enhancing the robustness of landslide segmentation in complex backgrounds.

The decoder block reconstructs the feature representations from the bottleneck layer, progressively restoring the spatial resolution of the image. By up-sampling and concatenating features from the encoder, the decoder refines predictions, enabling precise localization of landslide boundaries (Vega et al., 2024). The decoding process includes up-sampling and convolutional operations:

$$y = f(W_d * z' + b_d) \tag{3}$$

where y is the reconstructed feature map, W_d and b_d are the decoder's weights and bias terms, and z' represents the up-sampled bottleneck feature. The decoder's up-sampling process, combined with feature concatenation from skip connections, ensures that the model accurately restores spatial details necessary for distinguishing landslide regions from surrounding terrain. Skip connections link corresponding layers of the encoder and decoder, allowing low-level features to be preserved throughout the model. They facilitate the flow of spatial information directly from encoder to decoder, helping maintain fine-grained details that are essential for precise landslide boundary segmentation (Chen et al., 2023). Mathematically, skip connections are represented as:

$$y_{skip} = y + x \tag{4}$$

where y denotes the decoder feature map and x is the corresponding encoder feature map. By bypassing certain layers, skip connections address information loss during down-sampling, enhancing feature resolution and ensuring the output closely aligns with the input image's structure.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

The UNet, with its encoder-decoder structure, bottleneck compression, and skip connections, effectively segments landslide regions by capturing both low- and high-level features. Its flexibility also makes it ideal for integrating attention mechanisms to enhance segmentation accuracy and robustness in complex remote sensing scenarios.

2.2 CBAM Integration for Enhanced Feature Attention

The CBAM enhances feature selection in Convolutional Neural Network (CNN) by sequentially applying channel and spatial attention mechanisms. Channel attention highlights important feature channels, while spatial attention refines focus on key image regions. Integrated into the UNet, CBAM enables more effective localization of landslide-prone areas, improving segmentation performance in complex terrain.

2.2.1 Channel Attention Mechanism

The Channel Attention Mechanism (CAM) in CBAM enhances the model's ability to focus on the most informative channels in the feature map, selectively emphasizing important features while suppressing irrelevant ones. This attention mechanism dynamically assigns weights to each channel based on its relevance, allowing the model to prioritize essential features for landslide detection in remote sensing imagery (Huang et al., 2020). By focusing on significant channels, CAM improves segmentation accuracy and helps the model concentrate on landslide-related patterns.

CAM begins by using global average pooling and max pooling to summarize each channel's spatial information across the feature map. Global average pooling captures the mean activation for each channel, while max pooling identifies the highest feature values within each channel (Zhou et al., 2021). Together, these operations provide a comprehensive view of each channel's importance. The pooling operations are represented as:

$$F_{avg} = AvgPool(F)$$
$$F_{max} = MaxPool(F) \quad (5)$$

where F is the input feature map, F_{avg} is the result of global average pooling, and F_{max} represents the output from max pooling. These pooled outputs serve as a condensed representation of each channel's contribution, highlighting both the average and peak responses, which are critical for identifying important channels in landslide imagery.

The pooled outputs are then processed through shared fully connected (FC) layers to compute a refined channel importance score. By mapping the pooled values through FC layers with shared weights, CAM generates a unified attention score for each channel, capturing its relative importance across spatial features (Chen et al., 2023). The channel attention map M_c is calculated as follows:

$$M_c = \sigma(W_1 ReLU(W_0 F_{avg}) + W_0 ReLU(W_1 F_{max})) \quad (6)$$

where M_c is the channel attention map, W_0 and W_1 are weight matrices of the FC layers, and σ is the $Sigmoid$ activation function. This calculation produces a channel attention map that assigns a unique importance weight to each channel, allowing the model to enhance focus on channels most relevant to landslide features.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

The computed channel attention map is then applied to the original feature map by element-wise multiplication, amplifying essential features. By multiplying each channel by its corresponding attention weight, the model selectively enhances or suppresses channel features, based on their contribution to landslide identification (Tolooshams et al., 2020). The formula is expressed as follows:

$$F_{out} = M_c \otimes F \tag{7}$$

where F_{out} represents the output feature map with channel attention applied, and $\otimes$ denotes element-wise multiplication. This step refines the feature representation, enabling the model to focus on the most relevant channels, which is critical for accurately distinguishing landslide areas from surrounding terrain.

CAM dynamically prioritizes relevant channels, enhancing critical features for landslide detection while suppressing less important ones. This targeted focus improves segmentation accuracy and robustness, especially in complex remote sensing image.

2.2.2 Spatial Attention Mechanism

The Spatial Attention Mechanism (SAM) in CBAM refines the network's focus by emphasizing critical spatial regions within each feature map. Unlike channel attention, which enhances specific channels, spatial attention operates across spatial dimensions, helping the model to prioritize areas within each image that are essential for landslide detection (Aghdam et al., 2023; Zhang et al., 2022). By selectively weighting spatial locations, SAM enables more accurate segmentation of landslide-prone regions by directing attention to spatial features of interest.

SAM begins by aggregating channel information to form a spatial feature map that combines key channel insights across spatial dimensions. Aggregation is achieved through max pooling and average pooling operations applied along the channel axis, producing two spatial maps that capture overall spatial activation across channels (Cao et al., 2022). The formula is expressed as follows:

$$\begin{aligned} F_{avg}^{spatial} &= AvgPool(F, axis = channels) \\ F_{max}^{spatial} &= MaxPool(F, axis = channels) \end{aligned} \tag{8}$$

where F is the input feature map, $F_{avg}^{spatial}$ and $F_{max}^{spatial}$ represent the spatial feature maps derived from average and max pooling across channels, respectively. These spatial maps serve as condensed representations of spatial relevance by pooling information across all channels, which helps highlight important spatial areas within the feature map.

The pooled spatial maps are then concatenated and passed through a convolutional layer to generate a unified spatial attention map (Zhao et al., 2021). A convolutional layer is applied to capture local dependencies and refine the spatial focus within the feature map. The formula is expressed as follows:

$$M_s = \sigma(Conv_{7\times7}([F_{avg}^{spatial}, F_{max}^{spatial}])) \tag{9}$$

where M_s is the spatial attention map, $Conv_{7\times7}$ denotes the convolutional layer with a 7×7 kernel size, and σ is the sigmoid activation function. The concatenation operation is represented by [ ]. The spatial attention map assigns weights to each spatial

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

location, determining areas of the feature map that are most significant for landslide detection. This step ensures that the model's attention is focused on regions of high spatial relevance.

The generated spatial attention map is applied to the original feature map through element-wise multiplication, focusing the model's attention on specific spatial regions. By multiplying each spatial location in the feature map with its corresponding attention weight, the model enhances or suppresses spatial regions based on their relevance to landslide features (Chen et al., 2023). The formula is expressed as follows:

$$F_{out} = M_s \otimes F \tag{10}$$

where F_{out} is the output feature map with spatial attention applied, and $\otimes$ denotes element-wise multiplication. Applying the spatial attention map refines the spatial feature representation, making it easier for the model to identify regions critical for landslide detection.

By enabling location-specific focus, SAM enhances the model's ability to distinguish landslide-prone areas from the background, improving segmentation precision and robustness. Its spatial refinement complements CAM's channel-level attention, enabling more effective analysis of complex remote sensing data.

2.3 Integrating CBAM with UNet

The integration of CBAM with the UNet significantly enhances the network's capacity to capture and emphasize critical spatial and channel-wise information in remote sensing images for landslide detection. By merging CBAM's attention mechanisms with the Encoder-Decoder structure, the model effectively concentrates on key features across multiple levels, leading to improved segmentation accuracy and robustness. This approach enables UNet to extract intricate spatial details while prioritizing features most indicative of landslide-prone areas, resulting in a more precise and targeted analysis.

CBAM is embedded within the Encoder and Decoder paths of UNet to dynamically refine feature maps at various stages of the network. By integrating CBAM modules into UNet, the model applies attention to both channels and spatial dimensions, thus enhancing feature representation in the encoding and decoding processes, where essential landslide characteristics can be accentuated. At each CBAM layer, the feature map F undergoes channel attention M_c and spatial attention M_s, resulting in an updated feature map F_{out} as:

$$F_{out} = M_s \otimes (M_c \otimes F) \tag{11}$$

where F_{out} is the final output with spatial attention. This attention-enhanced feature map F_{out} at each Encoder and Decoder level ensures that UNet learns the most relevant features for landslide detection, focusing the model's efforts on critical areas in the image.

CBAM is applied to the skip connections in UNet, which bridge the encoder and decoder layers to retain detailed spatial information. By applying CBAM within skip connections, the model can highlight informative spatial and channel-wise details in the features that flow between Encoder and Decoder, maintaining crucial information about landslide-prone regions across UNet's levels.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

$$F_{skip-out} = M_c \otimes M_s \otimes F_{skip} \quad (12)$$

where F_{skip} is the skip connection feature map, and $F_{skip-out}$ is the output of the skip connection after applying both channel and spatial attention. Enhanced skip connections allow UNet to preserve detailed information across the network's levels, improving feature consistency and enabling the model to retain essential landslide features even in deep layers (Wang et al., 2024).

The combined UNet and CBAM architecture provides a multi-level attention mechanism that amplifies relevant features for landslide detection throughout the network. By applying CBAM at multiple levels of UNet, the model can distinguish critical landslide features from noise, optimizing feature extraction and segmentation across spatial and channel dimensions.

$$F_{final} = U_{CBAM}(F) \quad (13)$$

where F_{final} represents the output of the U-Net with integrated CBAM, and U_{CBAM} denotes the UNet augmented with CBAM layers. This enhanced output ensures that the model has focused on the most critical features for landslide segmentation, resulting in high accuracy and improved robustness against variations in remote sensing data.

The integration of CBAM into UNet significantly improves the model's capability to capture both spatially and channel-wise relevant features. This dual-attention approach allows the UNet to focus selectively on the most informative aspects of the image, thereby enhancing landslide segmentation accuracy. By refining UNet's ability to identify landslide patterns, CBAM integration provides a more effective and robust solution for remote sensing-based landslide detection.

2.4Auxiliary Feature Calculation

This section details the calculation of auxiliary features designed to enhance the model's capability to identify landslide-prone areas by quantifying terrain and vegetation characteristics. Key indices include the Normalized Difference Vegetation Index (NDVI), Slope, Elevation, Normalized Difference Moisture Index (NDMI), Green Normalized Difference Vegetation Index (GNDVI), Brightness, and Bare Soil Index (BSI). Each index provides specific insights into environmental attributes essential for distinguishing landslide regions from stable terrain. By calculating and incorporating these indices, the model benefits from enriched, multi-dimensional data that significantly boost predictive accuracy. A detailed explanation of each feature, including its formula, definitions of terms, and its role in the landslide detection process, is provided below.

NDVI represents the density and health of vegetation, calculated using near-infrared and red bands (Matas et al., 2022). The formula is as follows:

$$NDVI = \frac{NIR - R}{NIR + R} \quad (14)$$

where NIR is the near-infrared reflectance, and R is the red reflectance. This index helps distinguish vegetated areas, which can reveal disrupted vegetation as a characteristic of landslide zones.

The slope is derived from digital elevation data, representing terrain inclination (Sujatha et al., 2023). The formula is expressed as follows:

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

$$Slope = \arctan\left(\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\right) \quad (15)$$

where $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ denote elevation changes along the x and y axes, respectively. Steep slopes are more susceptible to landslides, making this feature valuable for identifying high-risk regions.

Elevation data provides absolute height information on the terrain, essential for analyzing topographic variations (Liu et al., 2023). Elevation values are directly obtained from digital elevation models. Areas with significant elevation changes are often more prone to landslides, highlighting the role of elevation in risk assessment.

NDMI measures vegetation or surface moisture using near-infrared and short-wave infrared bands (Delgado et al., 2022). It is calculated as follows:

$$NDMI = \frac{NIR - SWIR}{NIR + SWIR}) \quad (16)$$

where $SWIR$ represents the short-wave infrared reflectance. Higher moisture levels can destabilize terrain, making $NDMI$ a critical index for identifying potential landslide sites.

GNDVI is a vegetation index that emphasizes greenness, offering greater sensitivity than NDVI in some cases (Vaca et al., 2023). It is defined by:

$$GNDVI = \frac{NIR - G}{NIR + G}) \quad (17)$$

where G represents the green band reflectance. This index aids in capturing vegetation characteristics, which contribute to assessing slope stability in vegetated regions.

Brightness is computed by combining multiple bands to reflect overall surface reflectivity (Pellarin et al., 2016). Brightness information aids in distinguishing surface types, which helps in separating landslide-affected regions from stable ground.

BSI quantifies the presence of bare soil (Alam et al., 2023), calculated as follows:

$$BSI = \frac{(SWIR + R) - (NIR + B)}{(SWIR + R) + (NIR + B)} \quad (18)$$

where B is the blue reflectance. BSI is particularly useful in identifying areas with exposed soil, a characteristic often seen in landslide-prone areas.

The calculated auxiliary features serve to capture multi-dimensional environmental information, integrating vegetation health, moisture, elevation, and surface characteristics into the model. This enriched feature set allows the model to effectively distinguish landslide regions from non-landslide regions, thereby improving detection accuracy and robustness across varied terrains and environmental conditions.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

2.5 Model Training and Evaluation Metrics

This section outlines the model training strategy and evaluation metrics employed to assess the performance of the proposed UNet-CBAM. Effective training and rigorous evaluation are essential to ensure that the model generalizes well and accurately detects landslide-prone areas in remote sensing imagery.

The training strategy focuses on reducing overfitting and enhancing the model's ability to accurately learn the distinguishing features of landslides across varied landscapes (Levi et al., 2024). The model training objective is to minimize the cross-entropy loss, given as:

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)) \tag{19}$$

where N represents the number of pixels, y_i is the ground truth label (1 for landslide, 0 for non-landslide), and $\hat{y}_i$ is the predicted probability. This loss function evaluates the difference between the predicted and true labels for each pixel, allowing the model to iteratively adjust weights and improve segmentation accuracy for training.

Accuracy and Intersection over Union (IoU) are key metrics used to measure the segmentation quality of landslide regions against ground truth data (Baheti et al., 2020). Accuracy provides an overall measure of the proportion of correctly classified pixels, while IoU evaluates the overlap between predicted and actual landslide regions, providing insight into the model's precision in capturing relevant areas.

$$\begin{aligned} Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\ IoU &= \frac{TP}{TP + FP + FN} \end{aligned} \tag{20}$$

where $TP, TN, FP,$ and FN represent true positives, true negatives, false positives, and false negatives, respectively. These metrics offer a comprehensive evaluation of the model's performance, with IoU specifically tailored to measure segmentation tasks by calculating the area of overlap between predicted landslides and the actual annotated regions.

Precision and recall are essential to understanding the model's ability to correctly identify landslide pixels, emphasizing its robustness in landslide detection without misidentifying non-landslide areas (Huang et al., 2024; Zhang et al., 2021). Precision reflects the accuracy of positive predictions, while recall evaluates the completeness of landslide detections.

$$\begin{aligned} Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \end{aligned} \tag{21}$$

where TP and FP denote true positives and false positives, respectively, and FN represents false negatives. High precision and recall ensure that the model not only identifies most landslide pixels but also minimizes false positives, leading to a balanced and effective landslide segmentation performance.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

The F1 score provides a harmonic mean of precision and recall, offering a balanced metric to assess the model's accuracy in identifying landslides (Liu et al., 2022). This metric is particularly useful in cases of imbalanced data, ensuring that both precision and recall are considered equally in evaluating model performance.

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{22}$$

The $F1$ score provides a single, comprehensive metric that assesses the model's detection performance, making it useful for evaluating landslide segmentation, where both positive detection accuracy and completeness are crucial.

The training strategy and evaluation metrics collectively enable a robust assessment of the proposed UNet-CBAM. Through this approach, we ensure that the model is optimized for landslide detection and validated rigorously across multiple dimensions, resulting in a reliable, high-performance segmentation tool capable of accurately identifying landslide-prone regions from remote sensing data.

3 Experiment and Results

This study utilizes the Landslide4Sense-2022 dataset, a high-resolution remote sensing dataset released by the International Advanced Research Institute (IARAI) for landslide detection and monitoring. It includes diverse, well-annotated optical and SAR images of landslide events across various geographic and environmental conditions. With 128×128 pixel resolution and multi-band information stored in .h5 format, the dataset supports supervised learning with binary mask annotations, providing a rich and diverse resource for developing and evaluating deep learning models in geological disaster prevention.

3.1 Dataset and feature visualization

For feature extraction, we selected the fourth image from the Landslide4Sense-2022 dataset, and its distinct characteristics are showcased through multiple indices, as shown in Fig. 2. The RGB image presents a true color representation, capturing visible details and land cover variations. The NDVI highlights vegetation density and health, revealing high vegetation in green-dense areas and low vegetation in barren zones. The Slope map represents terrain gradients, accentuating steep versus gentle slopes, which are critical in landslide-prone regions. Elevation data shows altitude variations, providing insight into the topographic relief that influences landslide dynamics. The NDMI illustrates moisture levels within vegetation, useful for assessing soil and vegetation wetness. GNDVI focuses on chlorophyll concentration in vegetation, refining vegetation health analysis. The Brightness Index indicates land surface reflectance, which differentiates light and dark surfaces effectively, while the BSI enhances the detection of bare soil and non-vegetated surfaces. Together, these features provide a comprehensive depiction of the region's terrain, vegetation, moisture, and surface characteristics, aiding in nuanced landslide detection and analysis.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

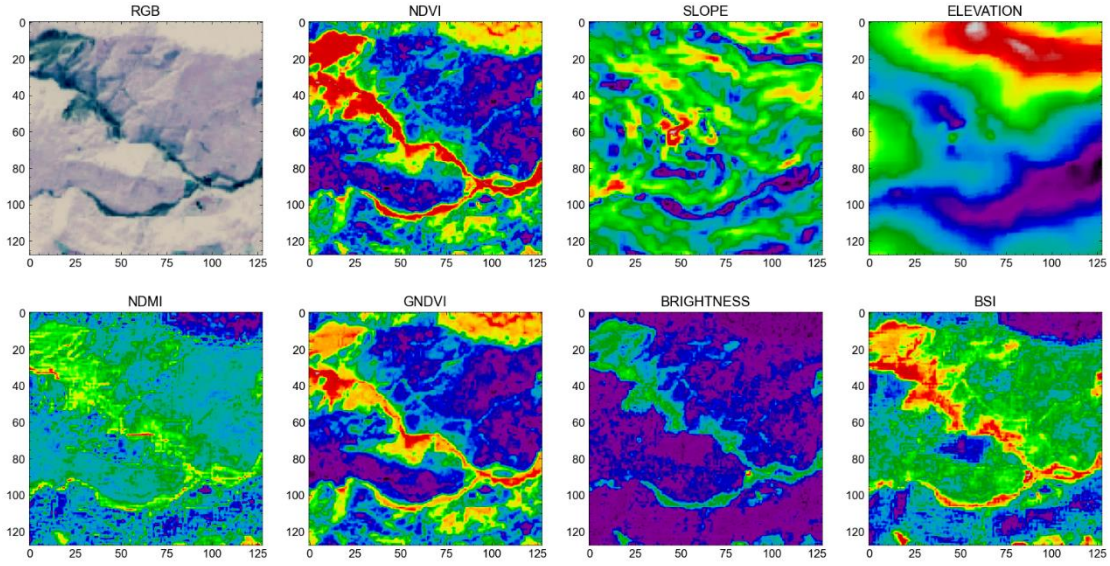


Fig. 2. Feature Extraction Maps for Landslide-Prone Region

3.2 Model Development and Stepwise Construction

To further elucidate the advantages of the proposed UNet-CBAM, this section presents a detailed analysis of its internal feature representations and a stepwise comparison with the baseline UNet through ablation experiments. First, the evolution of feature maps across different layers is visualized to highlight the enhanced spatial and semantic extraction capabilities introduced by the CBAM attention mechanism. Subsequently, a series of ablation studies are conducted to quantify the contribution of each architectural component, thereby validating the effectiveness of the CBAM integration in improving model performance and robustness.

As shown in Fig. 3, the feature map generated by the ‘conv2d_6’ layer comprises 32 sub-images, each representing distinct learned spatial patterns and feature activations from the input data. These sub-images capture a range of visual cues—such as edges, textures, and region-specific details—that reflect the model's capacity to extract meaningful representations of complex landscapes. Each feature map emphasizes different aspects of the scene, including object boundaries, surface textures, and high-contrast regions, all of which are critical for identifying landslide characteristics. Collectively, these 32 feature maps demonstrate the layer ability to encode multi-scale, context-rich information, thereby enhancing the model effectiveness in accurately distinguishing landslide areas from surrounding terrain.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

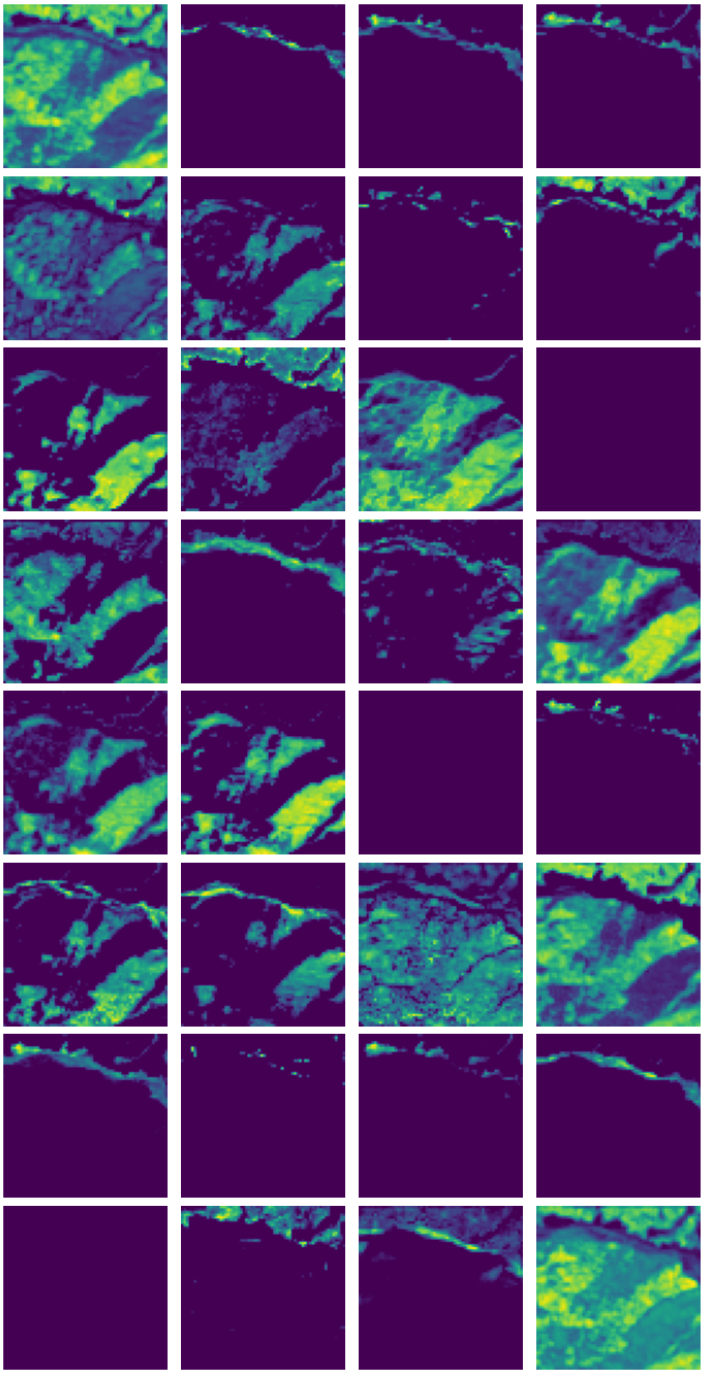
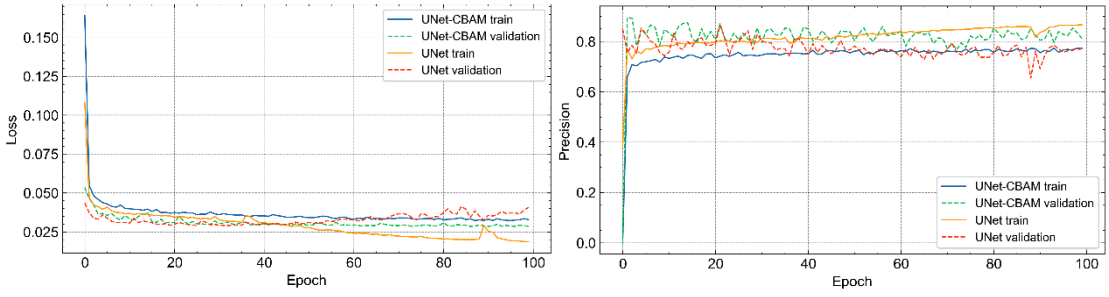


Fig. 3. Feature maps generated by the 'conv2d_6' layer

Fig. 4 illustrates the training and validation performance of UNet and UNet-CBAM using four evaluation subplots, focusing on the Loss Function and Precision metrics. In subplot (a), all four curves stabilize around the 20th epoch. UNet-CBAM

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

exhibits consistent training and validation loss, whereas UNet shows a slight decrease in training loss and a marginal increase in validation loss, suggesting potential overfitting. In subplot (b), the precision curves converge rapidly, stabilizing around the 4th epoch. UNet-CBAM maintains steady training precision with minor oscillations in validation precision, while UNet shows a gradual improvement in training precision accompanied by mild fluctuations in validation precision. Overall, UNet-CBAM demonstrates more stable performance in both loss and precision, with less fluctuation and a lower tendency for overfitting compared to UNet.



(a) (b)

Fig. 4. Comparative Analysis of Loss and Precision for UNet and UNet-CBAM

The ROC curve in Fig. 5 the strong predictive performance of both the UNet and UNet-CBAM. Beginning at the origin, the curve ascends steeply, reaching a true positive rate of 1.0 at a false positive rate of approximately 0.05. This indicates that the model achieves near-perfect sensitivity with minimal false positives at an early stage. Both models attain an AUC of 0.99, reflecting excellent overall classification performance. Despite their identical AUC values, the shapes of the ROC curves reveal important differences. The UNet-CBAM reaches optimal performance more quickly, as evidenced by its earlier and steeper rise compared to the baseline UNet. This suggests that UNet-CBAM is more efficient in distinguishing landslide areas from non-landslide regions at lower false positive rates, which is especially valuable in practical applications where early and accurate detection is critical. Overall, Fig. 5 highlights the outstanding classification performance of both models, while emphasizing UNet-CBAM's advantage in achieving faster convergence and improved early-stage discrimination, further validating its effectiveness in landslide detection tasks.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

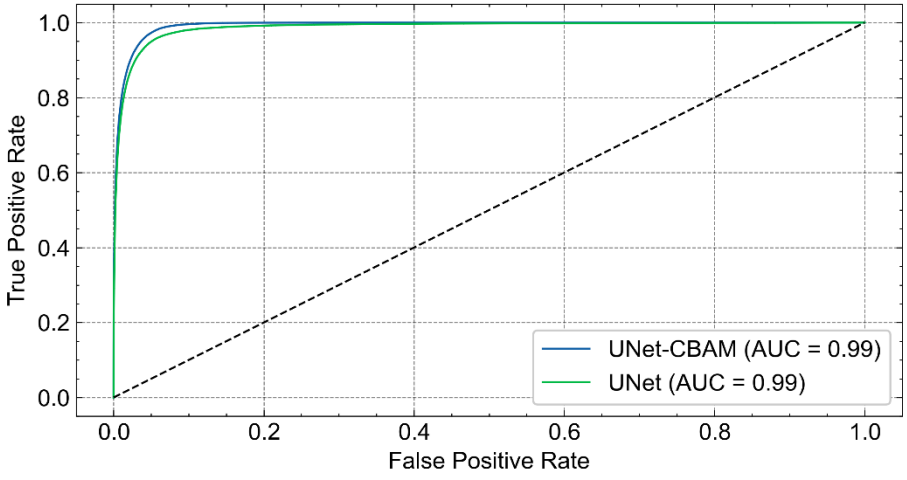
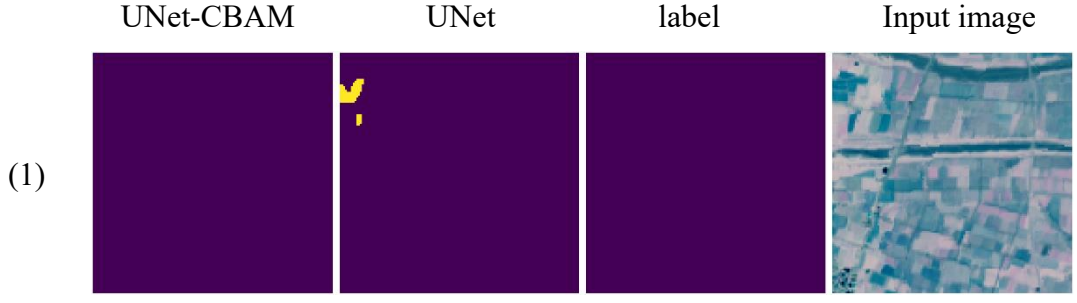


Fig. 5. ROC curve illustrating the predictive performance of the proposed model

Fig. 6 presents the ablation experiment results, showcasing a comparative analysis between UNet and UNet-CBAM through six sets of predicted landslide maps. In the first set, the model incorrectly identifies the upper-left corner as a landslide area, whereas the UNet-CBAM yields predictions that perfectly align with the ground truth. In the second set, UNet detects only a small portion in the lower-right region, while UNet-CBAM not only accurately captures this area but also successfully identifies the landslide region in the upper-left, with overall results more consistent with the label. The third set shows UNet performing relatively well in the central region; however, its detection in the lower-right is limited. In contrast, UNet-CBAM achieves better alignment with the ground truth in both the lower-right and lower-left regions. In the fourth set, UNet fails to detect any landslide area, whereas UNet-CBAM successfully identifies the landslide in the lower-left corner. The fifth set highlights significant discrepancies between UNet's prediction and the ground truth in both the upper-left and central areas. Conversely, UNet-CBAM demonstrates a markedly improved prediction that closely matches the label. In the final set, UNet shows poor detection in the lower and upper-right regions, while UNet-CBAM once again outperforms it with more accurate localization in both areas. In summary, Fig. 6 clearly demonstrates that the introduction of the CBAM module significantly enhances the model's ability to detect landslide areas, leading to more precise and reliable predictions, especially in regions where the baseline UNet struggles.



https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

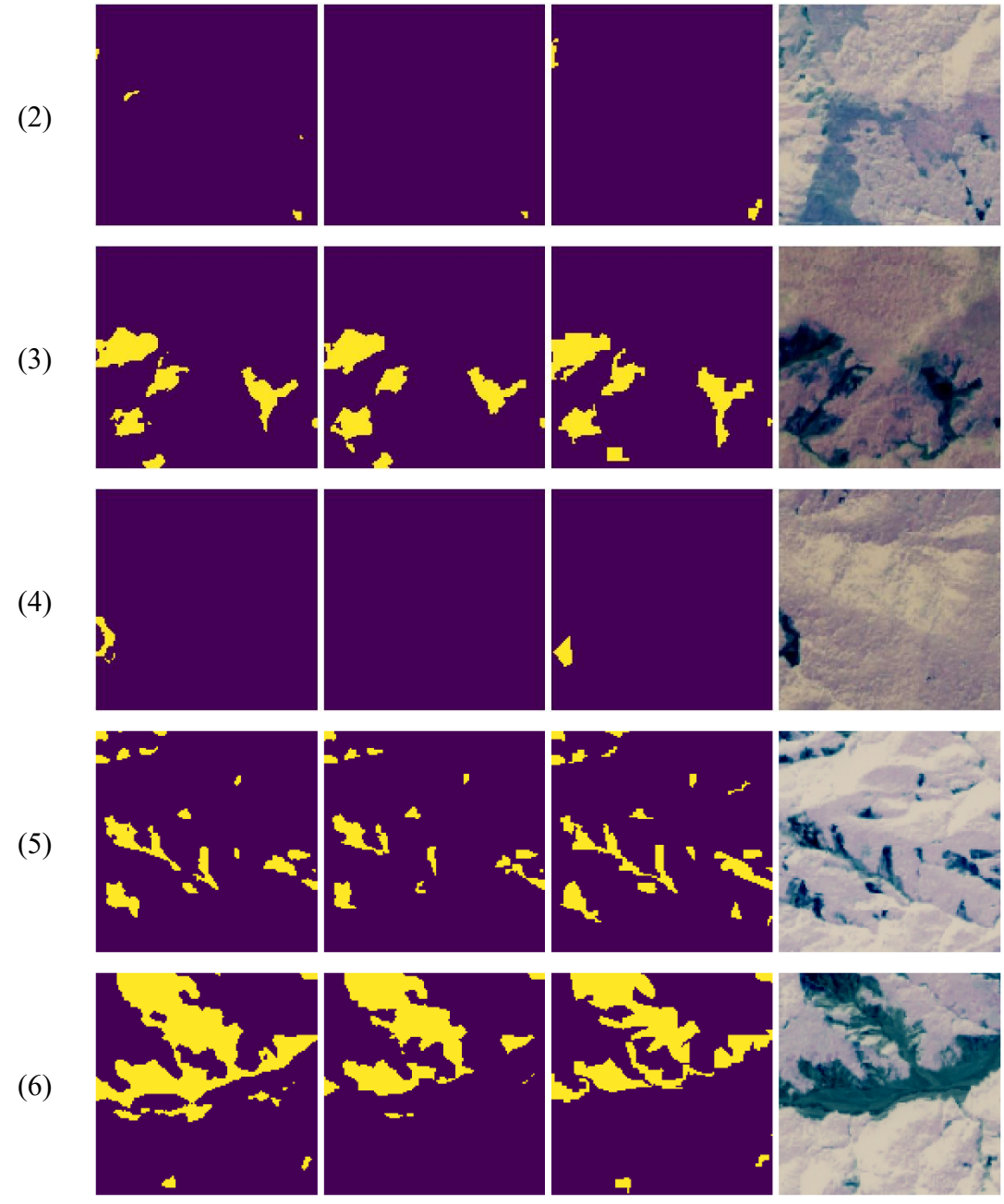


Fig. 6. Ablation results comparing UNet and UNet-CBAM for landslide detection.

3.3 Performance Comparison with Other Methods

To further validate the effectiveness and robustness of the proposed UNet-CBAM, this section conducts a comprehensive performance comparison with two representative semantic segmentation models: Deep Pyramid Fully Convolutional Network

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

(DP-FCN) (Lv et al., 2020) and SegNet (Antara et al., 2019). These methods serve as strong baselines, and the comparative analysis—based on quantitative metrics and visual results—demonstrates the superior accuracy and generalization capability of the proposed approach in landslide detection.

Fig. 7 illustrates the training and validation performance of DP-FCN, SegNet, and UNet-CBAM across four evaluation subplots, focusing on the loss function and precision. In subplot (a), all curves, except for the SegNet validation curve, reach stability around the 20th epoch. The SegNet validation loss initially increases rapidly, then decreases sharply before eventually stabilizing. In subplot (b), the UNet-CBAM test precision stabilizes around the 4th epoch, with minor fluctuations observed in the validation set. SegNet exhibits a similar trend but with slightly lower precision and more pronounced oscillations in validation. In contrast, DP-FCN shows minimal change during the first 38 epochs, followed by a gradual increase, reaching stability around the 50th epoch, but ultimately records the lowest precision among the three models. Overall, the UNet-CBAM demonstrates faster convergence, higher precision, and better generalization performance, while SegNet and DP-FCN exhibit slower or less stable training dynamics.

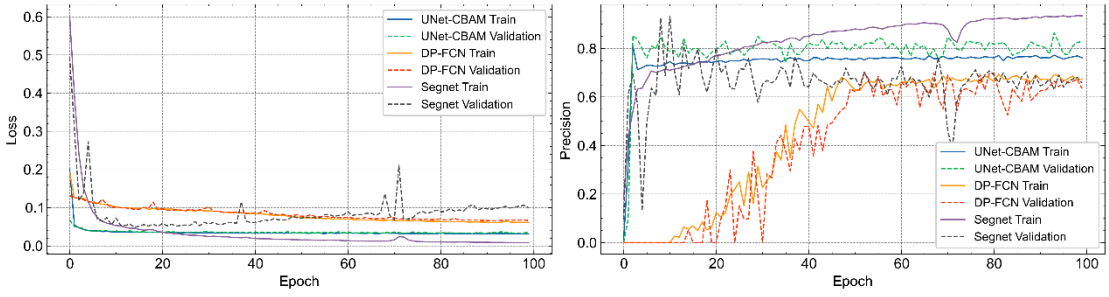
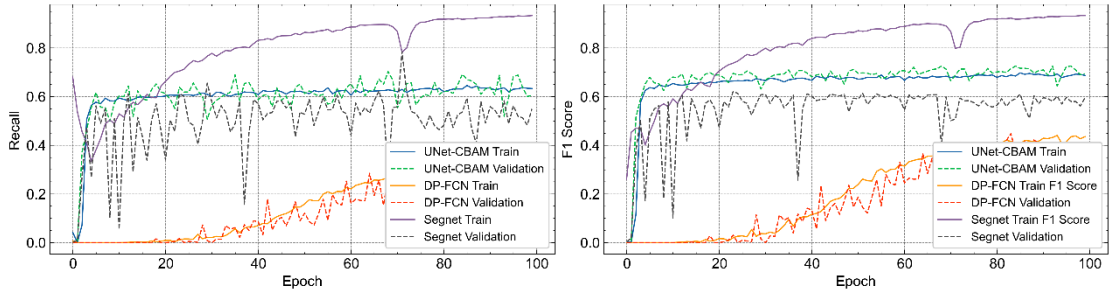


Fig. 7. Comparative Analysis of Loss and Precision for DP-FCN, SegNet, and UNet-CBAM

Fig. 8 presents the training and validation performance of DP-FCN, SegNet, and UNet-CBAM based on Recall and F1 Score. In subplot (a), UNet-CBAM's recall stabilizes around the 6th epoch with minor validation fluctuations. SegNet shows a similar trend but with slightly lower recall and more noticeable variability. DP-FCN shows little improvement until around the 38th epoch, gradually increasing and stabilizing near the 80th epoch, but with the lowest recall overall. In subplot (b), UNet-CBAM's F1 score also stabilizes by the 6th epoch, with consistent results and minimal variation. SegNet again follows a similar pattern with slightly lower scores and greater fluctuations. DP-FCN mirrors its recall behavior, improving only after the 38th epoch and stabilizing later with the lowest F1 score. Overall, UNet-CBAM consistently achieves faster convergence, higher recall and F1 scores, and better validation stability, confirming its superior performance in both classification accuracy and robustness compared to SegNet and DP-FCN.

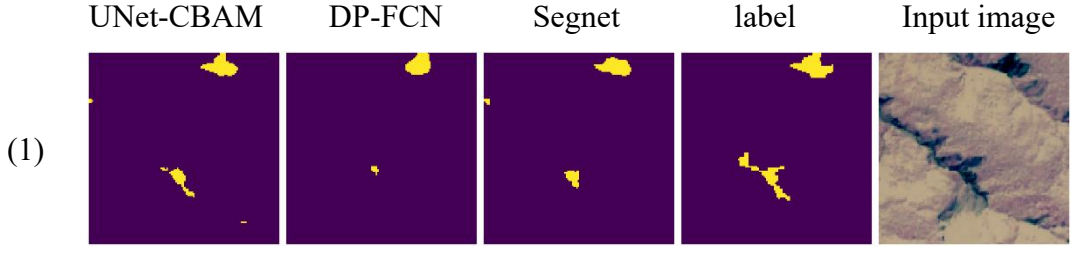
https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025



(a) (b)

Fig. 8. Comparative Analysis of recall and F1 Score for DP-FCN, SegNet, and UNet-CBAM

Fig. 9 presents a comparative analysis of the landslide prediction results from UNet-CBAM, DP-FCN, and SegNet across eight image samples. In the first image, SegNet incorrectly predicts the upper-left region as a landslide, while DP-FCN underperforms in the lower-left area. UNet-CBAM delivers superior results, accurately identifying landslide regions in both the upper-right and lower-left. In the second image, DP-FCN performs poorly, SegNet yields slightly better results, and UNet-CBAM's prediction aligns most closely with the ground truth. In the third image, DP-FCN fails to detect the landslide in the lower-right corner, and SegNet misclassifies the upper-right as a landslide area. UNet-CBAM again provides the most accurate prediction. In the fourth image, SegNet demonstrates poor performance by misclassifying the lower-left region, while DP-FCN detects the landslide but overestimates its extent. UNet-CBAM delivers a more precise and label-consistent prediction. In the fifth image, DP-FCN misses the upper-right landslide region entirely. SegNet slightly improves on this but falsely identifies the upper-left as a landslide. UNet-CBAM remains the most accurate. In the sixth image, DP-FCN fails to detect landslides in both the lower-left and upper-right areas, and SegNet misses the central landslide region. UNet-CBAM yields the best results. In the seventh image, DP-FCN again fails to detect landslides in the upper-right and central regions. SegNet and UNet-CBAM perform comparably overall, though SegNet mistakenly labels part of the central area as a landslide. UNet-CBAM remains closer to the ground truth. Finally, in the eighth image, DP-FCN misses the landslide in the upper part of the image. Both SegNet and UNet-CBAM produce similar predictions, but SegNet again misclassifies parts of the central region. UNet-CBAM shows the highest consistency with the label. Overall, Fig. 9 demonstrates that UNet-CBAM consistently outperforms DP-FCN and SegNet in landslide detection, providing more accurate and reliable predictions across various spatial patterns.



https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

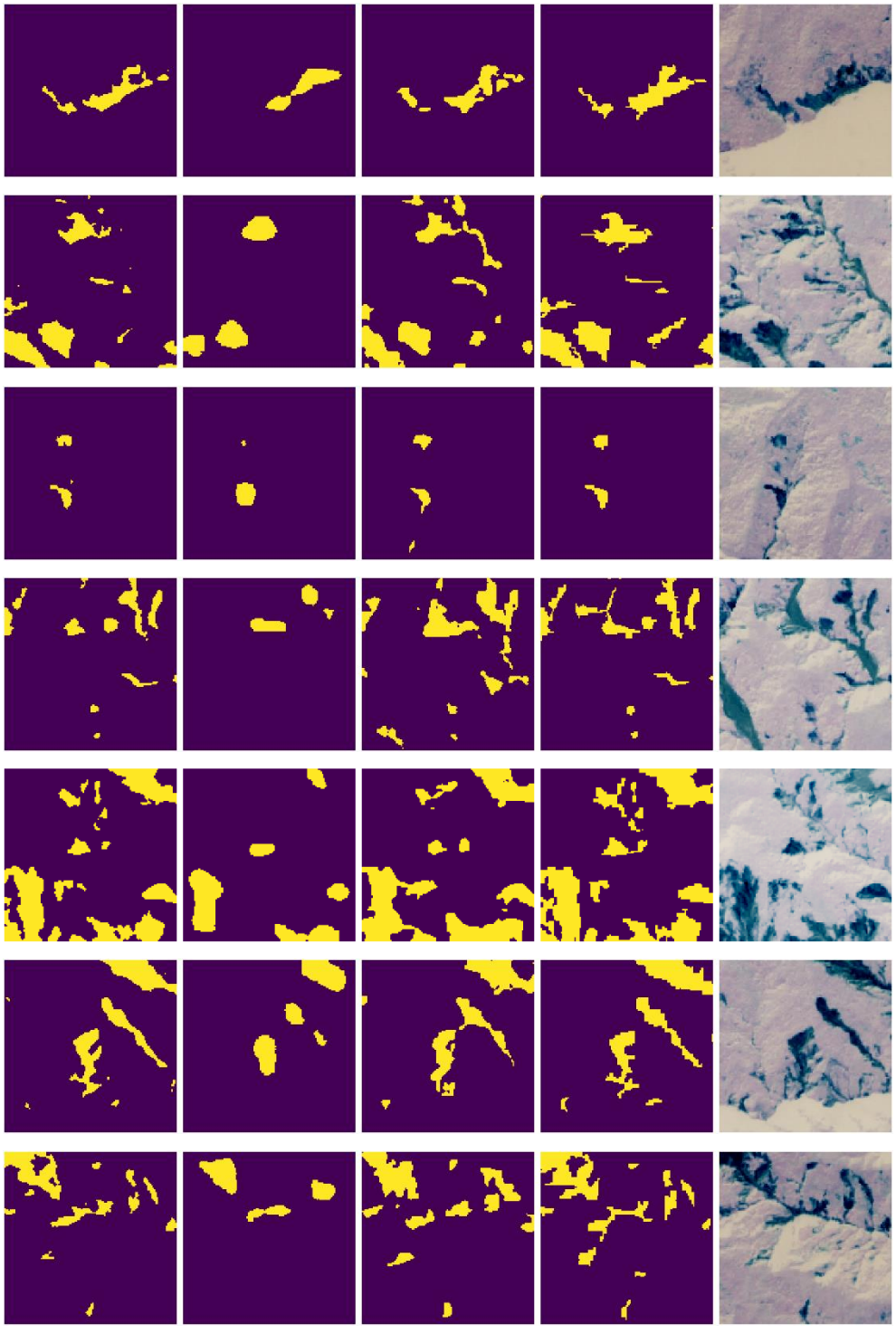


Fig. 9. Landslide Prediction Comparison of UNet-CBAM, DP-FCN, and SegNet

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

Table 1 presents the performance comparison of DP-FCN, SegNet, UNet, and UNet-CBAM across several evaluation metrics, including Accuracy, Intersection over Union (IoU), Dice Coefficient, Specificity, Balanced Accuracy, and Matthews Correlation Coefficient (MCC). Among the models, UNet-CBAM achieves the highest accuracy and balanced accuracy, indicating its superior overall performance in landslide detection. Although SegNet and UNet demonstrate similar performance in terms of specificity, UNet-CBAM exhibits a higher MCC (0.737), representing a 9.3% increase over UNet, reflecting its better capacity for handling imbalanced datasets and distinguishing true positive and negative cases. Additionally, UNet and UNet-CBAM outperform DP-FCN in terms of Dice Coefficient and IoU, with UNet-CBAM showing a Dice Coefficient of 0.650, marginally higher than UNet's 0.646. Despite slight differences in certain metrics, the UNet-CBAM consistently outperforms the other architectures, especially in terms of balanced accuracy and MCC, underscoring the efficacy of the integrated attention mechanism in enhancing model sensitivity and robustness in complex landslide detection tasks.

Table 1. Comparison of Model Performance for Landslide Detection

	Accuracy	IoU	Dice Coefficient	Specificity	Balanced Accuracy	MCC
DP-FCN	0.979	0.090	0.337	0.993	0.678	0.439
SegNet	0.984	0.228	0.602	0.993	0.773	0.610
UNet	0.986	0.262	0.646	0.995	0.804	0.674
UNet-CBAM	0.989	0.246	0.650	0.995	0.840	0.737

Figure 10 presents the confusion matrices for four landslide detection models, with true labels on the y-axis and predicted labels on the x-axis. For UNet-CBAM (subplot a) and UNet (subplot b), the true negatives (0,0) are approximately 4.8×10^6, while for DP-FCN (subplot c) and SegNet (subplot d), the value is around 4.7×10^6. In subplot (a) UNet-CBAM, the false negatives (1,0) are 24159, false positives (0,1) are 44347, and true positives (1,1) are 93206, reflecting its strong detection performance. In subplot (b) U-Net, the values for false negatives (1,0), false positives (0,1), and true positives (1,1) are 23584, 50000, and 87683, respectively, indicating a slight decline in performance compared to UNet-CBAM. In subplot (c) DP-FCN, the confusion matrix reveals a considerable decrease in performance, with false negatives (1,0) at 43489, false positives (0,1) at 74793, and true positives (1,1) at 62760, indicating a higher rate of missed detections. Lastly, in subplot (d) SegNet, the values are 44907 for false negatives (1,0), 54839 for false positives (0,1), and 82714 for true positives (1,1), placing SegNet's performance between that of U-Net and DP-FCN. Overall, UNet-CBAM demonstrates the most optimal balance between false positives and false negatives, achieving the highest true positive rate and superior overall detection accuracy compared to the other models.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

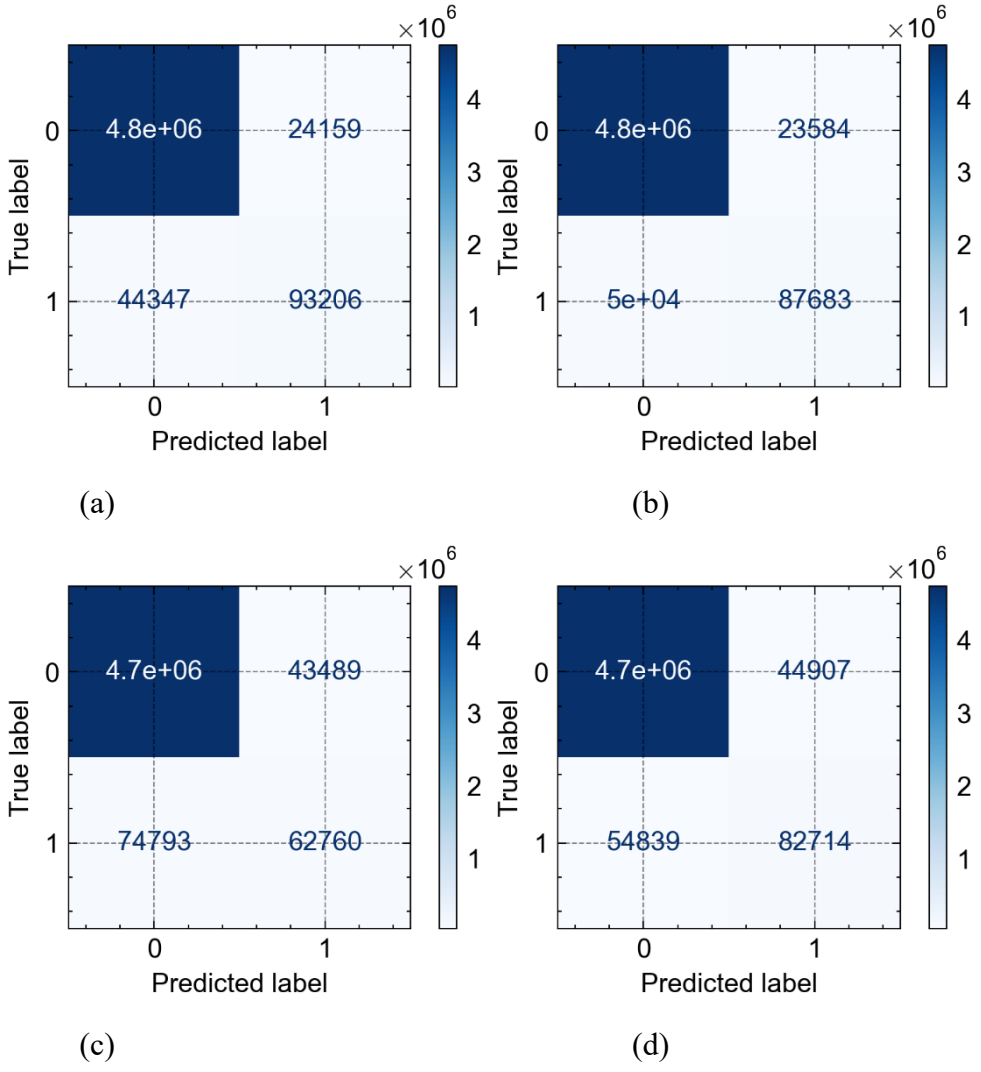


Fig. 10. Comparative confusion matrices illustrating landslide detection accuracy across four models

Figure 11 compares the ROC curves for the UNet-CBAM, DP-FCN, and SegNet, with the false positive rate on the x-axis and the true positive rate on the y-axis. The UNet-CBAM achieves stability the fastest, reaching its maximum true positive rate around 0.05 and maintaining it with an AUC of 0.99, indicating high accuracy. In contrast, the DP-FCN is the slowest to stabilize, reaching its peak true positive rate near 0.4, with an AUC of 0.96. SegNet exhibits rapid growth around 0.05 but then increases more gradually, with an AUC of 0.94. Overall, the UNet-CBAM outperforms the others in terms of both stability and overall accuracy.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

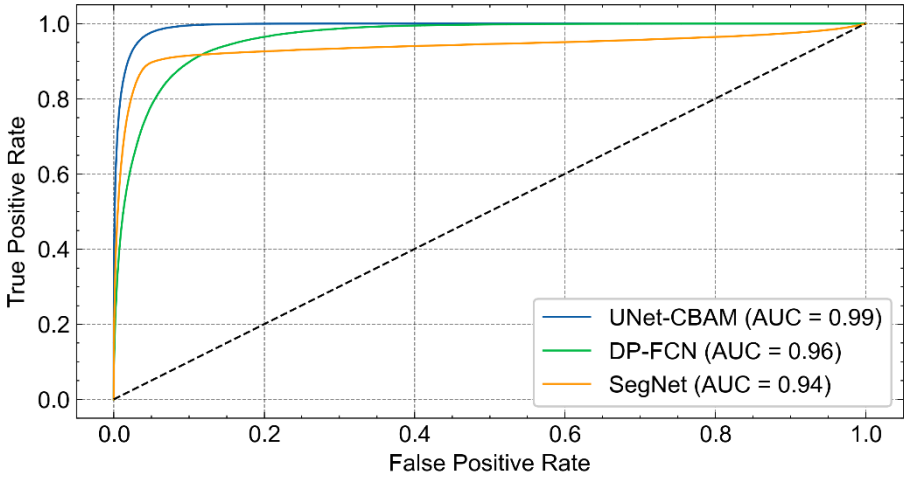


Fig. 11. ROC Curve Comparison for Landslide Detection

4 Discussion

This study presents the UNet-CBAM as an effective approach for landslide detection in remote sensing imagery, achieving high levels of accuracy and robustness across multiple evaluation metrics. Specifically, the model outperforms traditional architectures such as UNet, DP-FCN, and SegNet, demonstrated by superior performance on metrics like accuracy, IoU, Dice Coefficient, and Matthews Correlation Coefficient. In particular, the high AUC (0.99) of the ROC curve and the model's ability to quickly reach a stable true positive rate indicate its strong predictive capability in distinguishing landslide-prone regions. These results suggest that integrating the CBAM attention mechanism into UNet allows for more accurate feature extraction and spatial relevance, enhancing model performance even in complex terrain.

Compared with previous studies, which mainly used baseline segmentation models like UNet, DP-FCN, and SegNet for landslide detection, our findings reveal the added benefit of attention mechanisms in remote sensing applications. Past literature highlights the importance of pixel-wise feature extraction, particularly with models like SegNet, which performs well on basic segmentation but lacks advanced attention refinement, resulting in comparatively lower performance on complex images. While DP-FCN show limitations in detecting minor details of landslide areas, especially in noisy regions, the proposed model addresses this by incorporating spatial and channel-wise attention, achieving significantly better alignment with ground truth labels. This outcome supports recent advances that emphasize attention mechanisms' role in enhancing model accuracy for high-stakes environmental monitoring applications.

The enhanced performance of the proposed model likely arises from the model’s ability to focus on critical regions within images, a capability facilitated by the CBAM module’s dual attention mechanisms. By capturing both spatial and channel-wise relationships, the proposed model prioritizes the most relevant features for landslide identification, improving its detection capability in difficult regions, such as areas with high vegetation or steep terrain. This attention-driven focus enables the

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

proposed model to recognize subtle changes indicative of landslides, such as variations in vegetation index, slope, and brightness. For practical applications, this suggests that the model could be highly valuable in early warning systems, where precise detection of landslide-prone areas could contribute to timely disaster response and risk management.

While the proposed model demonstrates strong performance, there are several limitations. The model's reliance on high-quality annotated data restricts its applicability in regions where labeled landslide data is scarce or inaccessible. Additionally, despite achieving high accuracy, the model occasionally misclassifies small non-landslide areas, especially in the lower-left region of images, as landslide-prone. This limitation may stem from insufficient representation of diverse landslide types and surrounding conditions in the training dataset. Furthermore, the model's complexity introduces computational demands that may challenge its deployment in real-time or resource-constrained settings, potentially affecting its scalability.

Future work could focus on expanding the dataset to include more varied landslide scenarios, enhancing the model's generalizability to different terrains and environmental conditions. Additionally, integrating multi-temporal data and additional sensors, such as GNSS or environmental measurements, could improve the model's robustness and accuracy. Exploring lightweight versions of UNet-CBAM or developing optimization techniques for real-time applications would be beneficial for practical deployment in early-warning systems. Lastly, adapting the model to semi-supervised or unsupervised frameworks could reduce its dependence on annotated data, allowing it to be effectively applied in regions with limited data availability.

5 Conclusion

This study aimed to develop and evaluate an enhanced deep learning model, UNet-CBAM, for landslide detection in remote sensing imagery, addressing the need for higher accuracy and reliability in identifying landslide-prone regions. By integrating a CBAM into the standard UNet architecture, the research sought to improve spatial and channel-wise feature extraction for better performance in complex terrain environments.

The proposed model demonstrated superior performance across multiple metrics, achieving high accuracy, IoU, Dice Coefficient, and an MCC, all outperforming traditional models like UNet, DP-FCN, and SegNet. Particularly, the model's ROC curve stability and high AUC illustrate its robust capacity for true identification of landslide areas. This study confirms that attention mechanisms significantly enhance the model's ability to capture fine-grained features in remote sensing images, improving landslide area alignment with ground-truth labels.

This study makes a significant contribution to the field by innovatively integrating CBAM with UNet to enhance landslide detection, establishing a new benchmark for the application of attention mechanisms in environmental monitoring. The improved segmentation accuracy, particularly in areas with dense vegetation or complex topography, highlights the advantages of the proposed model over traditional segmentation methods. This approach refines remote sensing techniques for geohazard detection and provides a foundation for incorporating attention mechanisms into broader environmental monitoring applications.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

The findings of this research have significant practical implications for early warning and risk management systems in geohazard-prone regions. By offering a reliable, high-accuracy model, the proposed approach can be effectively integrated into landslide monitoring systems, enabling timely identification of high-risk areas and supporting informed decision-making for disaster prevention. The model's robustness makes it well-suited for large-scale monitoring in regions where landslide risks pose threats to communities, infrastructure, and ecological stability.

Future research will focus on optimizing the proposed model for real-time applications and exploring semi-supervised or unsupervised learning techniques to reduce data dependency. Additionally, incorporating other data sources, such as multi-temporal or environmental data, could further enhance the model's robustness, leading to more comprehensive and proactive geohazard monitoring. This study paves the way for future developments, with the potential to extend its application to a broader range of environmental monitoring challenges.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant No. 42404017), Natural Science Basic Research Program of Shaanxi(Program No. 2025JC-YBQN-403), Hunan Provincial Natural Science Foundation of China (Grant No. 2024JJ8377)

Code/Data availability: Code and data cannot be shared publicly but may be requested from the corresponding author upon reasonable justification.

Author contributions: HaiYang Li conceptualized the study, developed the methodology, performed formal analysis, and supervised the project. Jing Wang contributed to data curation, investigation, and manuscript review. Shuguang Wu assisted with conceptualization, investigation, and project administration. Yawei Wang focused on validation, formal analysis, and visualization. Yehemin Gao contributed to methodology, data curation, and manuscript review. Guigen Nie handled funding acquisition, project administration, and supervision.

The authors declare that they have no conflict of interest.

References

Aghdam, E. K., Azad, R., Zarvani, M., and Merhof, D.: Attention Swin U-Net: Cross-contextual attention mechanism for skin lesion segmentation, in: Proceedings of the 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), IEEE, 1–5, doi:10.1109/ISBI53787.2023.10230011, 2023.

Alam, S. A., Maiti, R., and Mandal, S.: Comparative assessment of landslide risk modelling by bivariate model in East Sikkim Himalaya, India, in: Agriculture and Climatic Issues in South Asia, CRC Press, 229–266, doi:10.1201/9781003377825-20, 2023.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

Antara, I. M. O. G., Shimizu, N., Osawa, T., and Nuarsa, I. W.: An application of SegNet for detecting landslide areas by using fully polarimetric SAR data, Ecotrophic, 13, 215–226, doi:10.24843/EJES.2019.v13.i02.p09, 2019.

Baheti, B., Innani, S., Gajre, S., and Talbar, S.: Eff-UNet: A novel architecture for semantic segmentation in unstructured environment, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020, 358–359, doi:10.1109/CVPRW50498.2020.00115, 2020.

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., and Wang, M.: Swin-Unet: Unet-like pure transformer for medical image segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV), Springer, 205–218, doi:10.1007/978-3-031-19818-4_12, 2022.

Chen, B., Liu, Y., Zhang, Z., Lu, G., and Kong, A. W. K.: TransAttUnet: Multi-level attention-guided U-Net with transformer for medical image segmentation, IEEE Trans. Emerg. Top. Comput. Intell., 7, 1–12, doi:10.1109/TETCI.2023.3309626, 2023.

Chen, C., Zhang, C., Wang, J., Li, D., Li, Y., and Hong, J.: Semantic segmentation of mechanical assembly using selective kernel convolution UNet with fully connected conditional random field, Measurement, 209, 112499, doi:10.1016/j.measurement.2023.112499, 2023.

Chen, H., He, Y., Zhang, L., Yao, S., Yang, W., Fang, Y., Liu, Y., and Gao, B.: A landslide extraction method of channel attention mechanism U-Net network based on Sentinel-2A remote sensing images, Int. J. Digit. Earth, 16, 552–577, doi:10.1080/17538947.2023.2188659, 2023.

Dai, F. C., Lee, C. F., and Ngai, Y. Y.: Landslide risk assessment and management: an overview, Eng. Geol., 64, 65–87, doi:10.1016/S0013-7952(01)00093-X, 2002.

Dang, K. B., Nguyen, C. Q., Tran, Q. C., Nguyen, H., Nguyen, T. T., Nguyen, D. A., Tran, T. H., Bui, P. T., Giang, T. L., and Lenh, T. A.: Comparison between U-shaped structural deep learning models to detect landslide traces, Sci. Total Environ., 912, 169113, doi:10.1016/j.scitotenv.2023.169113, 2024.

Delgado-Moreno, D., and Gao, Y.: Forest degradation estimation through trend analysis of annual time series NDVI, NDMI and NDFI (2010–2020) using Landsat images, in: Advances in Geospatial Data Science: Selected Papers from the 2021 International Conference on Geospatial Information Sciences, Springer, 149–159, doi:10.1007/978-3-030-98096-2_11, 2022.

Devara, M., Maurya, V. K., and Dwivedi, R.: Landslide extraction using a novel empirical method and binary semantic segmentation U-NET framework using Sentinel-2 imagery, Remote Sens. Lett., 15, 326–338, doi:10.1080/2150704X.2024.2320178, 2024.

Dong, Z., An, S., Zhang, J., Yu, J., Li, J., and Xu, D.: L-Unet: A landslide extraction model using multi-scale feature fusion and attention mechanism, Remote Sens., 14, 2552, doi:10.3390/rs14112552, 2022.

Fang, C., Fan, X., Zhong, H., Lombardo, L., Tanyas, H., and Wang, X.: A novel historical landslide detection approach based on LiDAR and lightweight attention U-Net, Remote Sens., 14, 4357, doi:10.3390/rs14174357, 2022.

Fang, K., Tang, H., Li, C., Su, X., An, P., and Sun, S.: Centrifuge modelling of landslides and landslide hazard mitigation: A review, Geosci. Front., 14, 101493, doi:10.1016/j.gsf.2022.101493, 2023.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

Fu, B., Li, Y., Han, Z., Fang, Z., Chen, N., Hu, G., and Wang, W.: RIPF-Unet for regional landslides detection: a novel deep learning model boosted by reversed image pyramid features, Nat. Hazards, 119, 701–719, doi:10.1007/s11069-023-06145-0, 2023.

Ghorbanzadeh, O., Crivellari, A., Ghamisi, P., Shahabi, H., and Blaschke, T.: A comprehensive transferability evaluation of U-Net and ResU-Net for landslide detection from Sentinel-2 data (case study areas from Taiwan, China, and Japan), Sci. Rep., 11, 14629, doi:10.1038/s41598-021-94190-9, 2021.

Ghorbanzadeh, O., Shahabi, H., Crivellari, A., Homayouni, S., Blaschke, T., and Ghamisi, P.: Landslide detection using deep learning and object-based image analysis, Landslides, 19, 929–939, doi:10.1007/s10346-021-01843-x, 2022.

Huang, G., Du, S., and Wang, D.: GNSS techniques for real-time monitoring of landslides: A review, Satellite Navig., 4, 5, doi:10.1186/s43020-023-00095-5, 2023.

Huang, G., Zhu, J., Li, J., Wang, Z., Cheng, L., Liu, L., Li, H., and Zhou, J.: Channel-attention U-Net: Channel attention mechanism for semantic segmentation of esophagus and esophageal cancer, IEEE Access, 8, 122798–122810, doi:10.1109/ACCESS.2020.3006564, 2020.

Huang, L., Miron, A., Hone, K., and Li, Y.: Segmenting medical images: From UNet to Res-UNet and nnUNet, in: Proceedings of the 2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS), Guadalajara, Mexico, 26–28 June 2024, 483–489, doi:10.1109/CBMS61543.2024.00086, 2024.

Levi, R., Marzullo, A., Savini, G., Savevski, V., and Politi, L. S.: An Unet boosting training strategy for the BraTS-ISBI 2024 Goat Challenge, in: Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), Athens, Greece, 27–30 May 2024, 1–4, doi:10.1109/ISBI61543.2024.00001, 2024.

Li, Q., Song, D., Yuan, C., and Nie, W.: An image recognition method for the deformation area of open-pit rock slopes under variable rainfall, Measurement, 188, 110544, doi:10.1016/j.measurement.2021.110544, 2022.

Liu, F., and Wang, L.: UNet-based model for crack detection integrating visual explanations, Constr. Build. Mater., 322, 126265, doi:10.1016/j.conbuildmat.2021.126265, 2022.

Liu, X., Peng, Y., Lu, Z., Li, W., Yu, J., Ge, D., and Xiang, W.: Feature-fusion segmentation network for landslide detection using high-resolution remote sensing images and digital elevation model data, IEEE Trans. Geosci. Remote Sens., 61, 4500314, doi:10.1109/TGRS.2022.3233637, 2023.

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

Lv, Z., Liu, T., Kong, X., Shi, C., and Benediktsson, J. A.: Landslide inventory mapping with bitemporal aerial remote sensing images based on the dual-path fully convolutional network, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 13, 4575–4584, doi:10.1109/JSTARS.2020.2980895, 2020.

Matas-Granados, L., Pizarro, M., Cayuela, L., Domingo, D., Gómez, D., and García, M.: Long-term monitoring of NDVI changes by remote sensing to assess the vulnerability of threatened plants, Biol. Conserv., 265, 109428, doi:10.1016/j.biocon.2021.109428, 2022.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

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

Niu, C., Gao, O., Lu, W., Liu, W., and Lai, T.: Reg-SA–UNet++: A lightweight landslide detection network based on single-temporal images captured postlandslide, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 15, 9746–9759, doi:10.1109/JSTARS.2022.3219897, 2022.

Ofli, F., Imran, M., Qazi, U., Roch, J., Pennington, C., Banks, V., and Bossu, R.: Landslide detection in real-time social media image streams, Neural Comput. Appl., 35, 17809–17819, doi:10.1007/s00521-023-08648-0, 2023.

Pellarin, T., Mialon, A., Biron, R., Coulaud, C., Gibon, F., Kerr, Y., Lafaysse, M., Mercier, B., Morin, S., and Redor, I.: Three years of L-band brightness temperature measurements in a mountainous area: Topography, vegetation and snowmelt issues, Remote Sens. Environ., 180, 85–98, doi:10.1016/j.rse.2016.02.045, 2016.

Shi, W., Zhang, M., Ke, H., Fang, X., Zhan, Z., and Chen, S.: Landslide recognition by deep convolutional neural network and change detection, IEEE Trans. Geosci. Remote Sens., 59, 4654–4672, doi:10.1109/TGRS.2020.3022141, 2020.

Sujatha, E. R., Sudarsan, J. S., and Nithiyanantham, S.: A review on sustainable reinforcing techniques to stabilize slopes against landslides, Int. J. Environ. Sci. Technol., 20, 13873–13882, doi:10.1007/s13762-023-04832-w, 2023.

Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., and Lacasse, S.: Machine learning and landslide studies: recent advances and applications, Nat. Hazards, 114, 1197–1245, doi:10.1007/s11069-022-05423-7, 2022.

Tolooshams, B., Giri, R., Song, A. H., Isik, U., and Krishnaswamy, A.: Channel-attention dense U-Net for multichannel speech enhancement, in: Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 836–840, doi:10.1109/ICASSP40776.2020.9054450, 2020.

Vaca, P., Moscoso, M., and Vaca-Cárdenas, M.: Analysis of spectral indices of images captured from an unmanned aerial vehicle at the ANA Moyocancha altitude station, J. Surv. Fish. Sci., 10, 2979–2990, doi:10.52547/journal.10.3.1050, 2023.

Vega, J., and Hidalgo, C.: Evaluation of U-Net transfer learning model for semantic segmentation of landslides in the Colombian tropical mountain region, MATEC Web Conf., 2024, 19002, doi:10.1051/matecconf/202419002, 2024.

Wang, H., Cao, P., Yang, J., and Zaiane, O.: Narrowing the semantic gaps in U-Net with learnable skip connections: The case of medical image segmentation, Neural Netw., 178, 106546, doi:10.1016/j.neunet.2024.106546, 2024.

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

Wei, R., Ye, C., Sui, T., Zhang, H., Ge, Y., and Li, Y.: A feature enhancement framework for landslide detection, Int. J. Appl. Earth Obs. Geoinf., 124, 103521, doi:10.1016/j.jag.2023.103521, 2023.

Yu, H., Ma, Y., Wang, L., Zhai, Y., and Wang, X.: A landslide intelligent detection method based on CNN and RSG_R, in: Proceedings of the 2017 IEEE International Conference on Mechatronics and Automation (ICMA), IEEE, 40–44, doi:10.1109/ICMA.2017.8015914, 2017.

https://doi.org/10.5194/egusphere-2025-2348
Preprint. Discussion started: 28 May 2025

Zhang, L., Shen, J., and Zhu, B.: A research on an improved Unet-based concrete crack detection algorithm, Struct. Health Monit., 20, 1864–1879, doi:10.1177/1475921720940068, 2021.

Zhang, L., Zhang, R., Dou, J., Hou, S., Xiang, Z., Wang, H., Yang, P., and Liu, X.: Advancing reservoir landslide stability assessment via TS-InSAR and airborne LiDAR observations in the Daping landslide group, Three Gorges Reservoir Area, China, Landslides, 21, 169–188, doi:10.1007/s10346-024-02337-2, 2024.

Zhang, Y., Kong, J., Long, S., Zhu, Y., and He, F.: Convolutional block attention module U-Net: A method to improve attention mechanism and U-Net for remote sensing images, J. Appl. Remote Sens., 16, 026516, doi:10.1117/1.JRS.16.026516, 2022.

Zhao, F., Wu, Z., Wang, L., Lin, W., Gilmore, J. H., Xia, S., Shen, D., and Li, G.: Spherical deformable U-Net: Application to cortical surface parcellation and development prediction, IEEE Trans. Med. Imaging, 40, 1217–1228, doi:10.1109/TMI.2021.3059293, 2021.

Zhou, T., Canu, S., and Ruan, S.: Automatic COVID-19 CT segmentation using U-Net integrated spatial and channel attention mechanism, Int. J. Imaging Syst. Technol., 31, 16–27, doi:10.1002/ima.22527, 2021.