

Optimizing the Topographical Bottleneck: An ASPP-Guided Attention ResUNet for Multispectral Semantic Segmentation

Ritesh Kumar*

riteshchandel@gmail.com

Assistant Professor
(AI&DS)

Rakshit Dabral **

rakshiddabral1@gmail.com

** Scholar, Btech (AI&DS) 4th Year

** Department of Artificial Intelligence and Data Science
Dr. Akhilesh Das Gupta Institute of Professional Studies

Abstract

The accurate detection of heterogeneous terrain features, such as landslides, remains a significant challenge in remote sensing due to complex topographical variations and extreme class imbalance. In this paper, we propose a novel multi-scale deep learning framework powered by a customized 9-channel topographical feature stack. This stack fuses globally normalized Sentinel-2 multispectral optical bands (Red, Green, Blue) and vegetation/moisture indices (NDVI, NDMI) with Digital Elevation Model (DEM) data (Slope, Elevation). Crucially, we explicitly model terrain-vegetation dynamics through engineered interconnection features (Slope-NDVI and Slope-NDMI) to capture physical precursors of terrain failure. We feed this spatial data into an optimized Attention ResUNet architecture enhanced with Atrous Spatial Pyramid Pooling (ASPP), enabling the simultaneous capture of localized pixel signatures and wide-area contextual topography. To combat spatial overfitting and enforce context-invariant learning, we implement a simple data augmentation strategy adapted for multi-channel physical data, optimized via a hybrid BCE-Dice loss function. Empirical evaluations demonstrate that our proposed model achieves a strong balance of precision and recall, yielding a peak F1 score of 0.750 and an Intersection over Union (IoU) of 0.597. Furthermore, post-inference probability threshold calibration reveals optimal boundary extraction at a 0.40 confidence threshold, significantly recovering false negatives along ambiguous topographical edges. This research demonstrates that integrating multi-scale spatial pooling with engineered topo-spectral interaction features provides a highly robust solution for automated terrain segmentation.

Index Terms:- Semantic Segmentation, Attention ResUNet, Atrous Spatial Pyramid Pooling, Multi-spectral Imagery, Digital Elevation Model, Landslide Detection

Abbreviations:-

ASPP: Atrous Spatial Pyramid Pooling

BCE: Binary Cross Entropy

ResNet : Residual Network

DEM: Digital Elevation Model

NDMI: Normalized Difference Moisture Index

NDVI: Normalized Difference Vegetation Index

1. INTRODUCTION

According to the literature [1] landslide also known as landslips, rock-slips or rock-slides is a form of mass wasting that include a wide range of ground movements. Landslide occur in a variety of environments, characterized by either steep or gentle slope gradients. Gravity is the primary driving force for a landslide to occur but in many cases heavy rainfall, earthquake, a slope cut to build road etc can lead to a landslide. It occurs when the shear stress exerted on a slope exceeds the shear strength of its underlying materials, triggering a critical shift from geological stability to instability. This complex mechanical failure is driven by a compounding mix of natural and synthetic factors. Landslide can have a huge impact on the community and wildlife near the affected site. For example, On 23 August 2025, Uttarakhand was hit by landslide leading to 6 fatalities [2].

To address these issues a common approach involves the creation of manual data creation of landslide data by going to the site manually and conducting detailed field reconnaissance. This often require selective field checking to verify terrain features, alongside recording the specific details like the state of activity, movement type, and material thickness. In more complex areas, it can even involve subsurface exploratory work or detailed drilling programs for physical sampling. While this hands-on method yields highly accurate information, it is labor-intensive, potentially hazardous in active zones, and struggles to capture real-time changes during severe weather [3].

Thanks to advancements in Geo-spatial Technologies, they have significantly enhanced the accessibility and prevalence of remote sensing and GIS-based mapping methods. Tools like LiDAR, satellite radar (InSAR), and aerial photography now allow experts to safely analyze terrain, detect ground displacement, and model hazards over vast areas without the physical limitations of traditional field surveys. Furthermore the recent advances

in machine learning and deep learning has introduced innovative methodologies for addressing these problems. Where system can automatically do all these manual labour tedious task.

In recent years, research has focused on traditional models like Random Forest, AdaBoostM1, Bagging [4]. However, there is a shift in research trend toward deep learning utilizing models such as U-net [5] [6], which are powered by computer vision often leads to better result as compared to manual task. Most of the deep learning project available online focuses on a specific region for their project leading to issues while cross validating. Hence to over come this issue we decided to use Landslide4Sense dataset.

Landslide4Sense dataset has three splits, training/test/validation consisting of 3799, 800, and 245 image patches, respectively [7]. Each image patch is a composite of 14 bands that include:

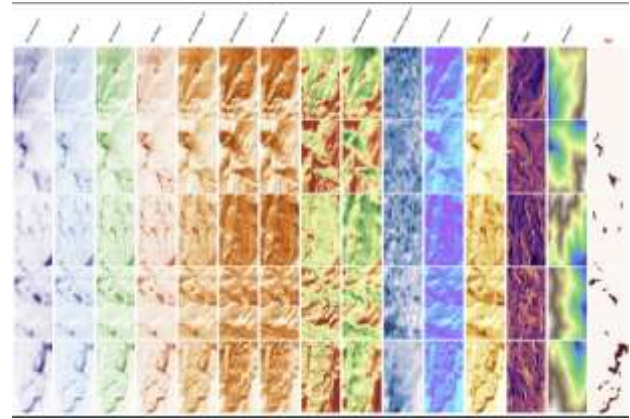
Table I
BANDS INFORMATION IN LANDSLIDE4SENSE DATASET[8]

Band	Description
Band 1	Atmospheric correction (aerosol scattering)
Band 2	Vegetation senescing/browning, soil background, aerosol scattering
Band 3	Green peak, total chlorophyll in vegetation
Band 4	Maximum chlorophyll absorption
Band 5	Red edge position, atmospheric/fluorescence baseline correction
Band 6	Red edge position, aerosol load retrieval
Band 7	Leaf area index
Band 8	NIR
Band 9	Water vapor absorption, atmospheric correction
Band 10	Thin cirrus detection for atmospheric correction
Band 11	Lignin/starch, forest biomass, snow/ice/cloud separation
Band 12	Mediterranean vegetation, clay soil erosion, live vs. dead biomass distinction
Band 13	Slope
Band 14	Digital Elevation Model (DEM)

Inspired by advances in deep learning methods and to address the above mentioned issues, we propose a

deep-learning based model for landslide detection and segmentation task by exploring remote sensing images in this paper. We mainly contribute:

- We conducted multiple experiments to indicate the role of custom input features, network architectures, loss function etc.
- Introduction of a 9-channel topo spectral stack incorporating engineered interaction features
- Development of ASPP-ARUnet, which integrates ASPP and Attention mechanism to resolve multi-scale topographical features
- A comprehensive ablation study demonstrating the limitation of standard attention in high-imbalance remote sensing tasks.


Fig.1. Random Samples with all the bands and mask
Table II
DATASET STATISTICS

Dataset	Images Count	Images Size
Landslide4Sense [8]	3044	128x128x14

1.1 Challenges

The development of a multi-scale semantic segmentation pipeline that fuses 9-channel topo-spectral data for introduces several notable challenges. The foremost difficulty lies in managing the extreme unbalanced class. This inherent asymmetry initially biases the model toward conservative background predictions, degrading the recall and masking the network's true predictive capabilities under standard binary classification thresholds. Traditional computer vision augmentations like color jitters or pixel blending, mathematically invalidates the physical indices like NDVI and NDMI, necessitating the implementation of a strict soft augmentation strategy to preserve the structural and spectral integrity of satellite-terrain pairs. Additionally, fusing highly heterogeneous data—combining optical reflectance with physical elevation and engineered interaction features—introduced the risk of overspecialization, demanding careful architectural scaling (such as reducing baseline filters) to prevent the network from memorizing irrelevant environmental noise.

Together, these obstacles highlight the central tension between multi-modal data complexity, spatial generalization, and metric alignment in remote sensing, emphasizing the critical need for domain-specific augmentation, multi-scale context pooling and calibrated evaluation.

1.3 Applications

The real-world applications of this multi-scale top-spectral segmentation model extend far beyond academic research, offering critical, data-driven tools for global disaster management and infrastructure planning. Engineers can and urban planner can utilize these models to detect high risk zones, ensuring the new infrastructure-such as highway, dam or residential development is constructed safely away from vulnerable slopes. Beyond immediate disaster mitigation, this architecture serve as a powerful for insurance industry to assess regional catastrophic risk, as well as ecological organization can track long-term soil degradation, deforestation impact and terrain stability over time.

II. LITERATURE REVIEW

[9] **Rekha R. Nair, Tina Babu, Tripti Singh & Afnaan K**, in their paper “Unet and Semantic Segmentation Based Landslide Detection System” highlighted the advancement in Deep learning based semantic segmentation model for satellite imagery analysis. Utilizing Landslide4Sense data for the benchmark testing of their model.

[10] **Syed Mujtaba Hussaine, Linlong Mu, Yimin Lu, Syed Sajid Hussain**, in their paper “Landslide Image Segmentation with Attention Residual U-Net : A Hybrid Deep Learning Model”, presented an Attention Residual Unet (ARU-Net) model, an advanced deep learning framework that combines residual network and attention mechanism within the basic UNet model to enhance the training stability. This model was able to capture and resolve the issue with other models where the diverse characteristic of satellite images often lead to false detection and increased computational cost limiting models accuracy. The core innovation done was the implementation of residual block which allowed the model to train properly even for the dense architecture by adding a skip connection inside the convolution block and other was the attention block which allowed the model to extract features with great precision.

[11] **Meng Tang, Yuelin He, Muhammed Aslam, Edore Akpokodje and Syeda Fizzah Jilani**, in their paper “Enhanced U-Net++ for Improved Semantic Segmentation in Landslide Detection”, showcased how traditional method lack the ability to extract meaningful feature from dataset in landslide segmentation. They also showed how conventional architectures including U-Net, Swim-Net and Unet++ exhibit limitation in global contextual comprehension, edge refinement, parameter optimization etc. To fix all the mentioned issues they proposed a model integrating VGG convolutional operations, a U-Net++ framework and attention mechanism. There proposed model consisted of three

principle components. A modified VGG backbone network , UNet++ dense architecture and an attention mechanism block.

[12] **Ayush Srivastava, Priyanshu Bansal, Aayush Kumar Jha & Brindha Subburaj**, in their paper “Elevation Aware U-Net Based Landslide Segmentation Leveraging DEM and Saliency Map”, showcased how an pretrain CNN model can be used as a backbone to extract terrain Variation spatially as well as on the basis of elevation. This model was evaluated using F1 score, dice coefficient and IoU.

III.METHODOLOGY

As the paper focuses on landslide detection and segmentation task, we first collect published and benchmark remote sensing dataset which should consist of landslide event. For this project we decided to go with Landslide4Sense [8]. Given this is our selected dataset, we define two major task for these dataset which include dataset splitting, augmentation and custom feature creation.

3.1 Remote Sensing Image Dataset

Landslide4Sense dataset [8] highlights landslide appearance worldwide from 2015-2021. This dataset consist of 3844 multi-spectral images each of which is a combination of 14 bands. In particular band from B1 to B12 (12 Bands) are from Sentinel-2 which are the comprehensively described in Table I. B13 band is from slope data and B14 is taken from ALOS PALSAR satellite. Each band is of same size 128x128 and the mask data presents a binary image of size 128x128. Mask is entirely black with white pixels highlighting the landslide affected area. Some of the images with entire 14 bands are shown in Fig.1. The number of image containing the landslide is 58% of all the images in dataset. However, only 2.3% of the pixel actually highlights the landslide region. This showcases the class imbalance in the dataset [13].

Table III

IMAGE STATISTICS

Landslide Image Ratio	Landslide Pixel Ratio
58%	2.3%

To access the Landslide4Sense dataset, it is available under the name Landslide4Sense in Kaggle [14]. The data is provided in Hierarchical Data Format Version 5 (HDF5) [15].

3.2 Task Definition

While historical approaches to landslide analysis often bifurcate the problem into either image-level classification (e.g., using the Bijie dataset) or segmentation (e.g., using the LandSlide4Sense dataset), this study is exclusively focused on advancing the rigorous demands of pixel-wise semantic segmentation.

Specifically, we utilize the LandSlide4Sense dataset to model complex topographical failures, effectively performing detection and precise boundary delineation simultaneously.

We define the landslide segmentation task as a dense, pixel-wise binary classification problem. Let the input be defined as a multispectral and elevation feature stack X :

$$X \in \mathbb{R}^{H \times W \times C}$$

Where H and W represents the spatial dimensions of the remote sensing patch (128x128) pixels and C represents the number of engineered channels ($C=9$). This 9-channel stack fuses normalized Sentinel-2 optical bands, vegetation/moisture indices (NDVI, NDMI) and explicitly engineered terrain-vegetation interactions dynamics (Slope-NDVI and Slope-NDMI).

The objective of the proposed multi-scale Attention ResUnet architecture is to learn a mapping function $f(X; \theta)$, parameterized by the network weight θ , that predicts a corresponding continuous probability map $\hat{Y}$:

$$f(X; \theta) \rightarrow \hat{Y} \in [0, 1]^{H \times W}$$

To extract the final discrete landslide boundaries, a calibrated probability threshold τ (empirically tuned to $\tau=0.40$) is applied to the continuous output $\hat{Y}$ to generate the final binary prediction mask Y_{pred} :

$$Y_{pred}^{(i,j)} = \begin{cases} 1, & \text{if } \hat{Y}^{(i,j)} \geq \tau \\ 0, & \text{otherwise} \end{cases}$$

Where a pixel value of 1 designates the presence of a landslide (the positive target class) and 0 designates stable terrain or background elements. By formulating the task through this dense prediction paradigm on a heterogeneous feature stack, the model is forced to evaluate the localized topo-spectral signature of every individual pixel against its broader topographical context.

3.3 Data Splitting

As LandSlide4Sense dataset [14] has not published its label for test set. We split the training data into two parts, using 80:20 ratio, for training and testing.

3.4 Evaluation Metrics

To evaluate the predictive performance of the proposed model on the landslide segmentation task, we report a comprehensive suite of pixel level metrics. Because landslide footprints constitute a small fraction of the total topographical area, relying solely on Overall Pixel Accuracy can be misleading due to extreme class imbalance. Therefore, we primarily evaluate the model's classification reliability using Precision, Recall and F-1 Score

Furthermore, to quantify the precise spatial overlap and morphological accuracy of the predicted landslide boundaries against the ground truth annotation we utilize

two region based metrics: the Jaccard index and the Dice Coefficient.

3.5 Experimental Setup

We construct our proposed deep neural network with TensorFlow framework. All the network are trained for 100 epochs with callbacks such as Early-stopping, Checkpointer and Reduced Learning Rate on Nvidia Tesla T4 GPU 32GB. All evaluating models in this paper use Adam [16] for the optimization in the training process.

IV. BASELINE MODEL

We first propose a baseline model for landslide segmentation task and evaluate it on LandSlide4Sense dataset. The proposed pipeline consist of three main components: feature engineering and data augmentation, the U-Net backbone network architecture and the network head with chosen loss function for segmentation task.

4.1 Feature Engineering and Data Augmentation

Rather than passing raw satellite bands directly, the input images are per-processed into a 9-channel tensor (128x128x9) to enhance feature representation. These 9 channels include normalized RGB bands, calculated vegetation and moisture indices (NDVI and NDMI), normalized topological data (Slope and Elevation), and interaction terms combining slope with NDVI and NDMI.

Table IV

NEW BAND DATA IS GENERATED FROM 14 ORIGINAL BANDS IN LANDSLIDE4SENSE DATASET

New Band Data	Formula Used
NDVI	$(B8-B4)/(B8+B4+1e-6)$
NDMI	$(B8-B11)/(B8+B11+1e-6)$
Slope_NDVI	$(B13/4.04+1e-6)*(1-NDVI)$
Slope_NDMI	$(B13/4.04+1e-6)*NDMI$

During the training phase, data augmentation is applied dynamically at the batch level via a custom data generator. To improve spatial in-variance, each input and corresponding mask undergoes synchronized spatial transformations, including random horizontal and vertical flips, and random 90-degree rotations.

4.2 The U-net Based Backbone Architecture

The backbone utilizes a standard U-Net consisting of down-sampling and an up-sampling path. The downsampling part captures context using repeated blocks of two 3x3 convolutional layers (using ReLU activation), followed by a Spatial Dropout layer (ranging from 0.1 to 0.3 to prevent overfitting), and a 2x2 MaxPooling layer. The feature channels progressively double from 16 to 256. Whereas the upsampling path enables precise localization. It utilizes 2x2 transposed convolutions for up-sampling, concatenates the resulting feature maps with the corresponding skip connections from the down-sampling path, and applies blocks of two 3x3 convolutions with Dropout. The feature channels progressively halves back down to 16.

4.3 The Head and Loss Function

Given the output feature map from the final up-sampling block, a 1x1 convolutional layer with a Sigmoid activation function is applied to generate the final predicted mask of 128x128x1. To address the inherent class imbalance between background and the landslide pixels, the model is optimized using the negative Jaccard Coefficient (Intersection Over Union) loss function, which directly penalizes the mismatch between the predicted and ground truth landslide regions:

$$Loss_{Jaccard}(y, \hat{y}) = -\frac{\sum(y_i \cdot \hat{y}_i) + 1}{\sum y_i + \sum \hat{y}_i - \sum(y_i \cdot \hat{y}_i) + 1}$$

Where y_i represents the ground truth and $\hat{y}_i$ represents the predicted probability.

4.5 Evaluation of Baseline Model

We evaluated the proposed Baseline Model on the LandSlide4Sense dataset for the semantic segmentation task. After the training process, the validation split (128x128x9) inputs are fed into the baseline model to obtain the predicted mask. The prediction are evaluated against the ground truth using strict pixel-level metrics. As shown in Table V, the proposed baseline model demonstrated strong performance, achieving an F1 Score of 72.31% and a Jaccard coefficient of 56.05%.

Table V

EVALUATE THE PROPOSED BASELINE MODEL ON LANDSLIDE4SENSE DATASET FOR SEGMENTATION TASK

Network	F1 Score	Jaccard (IoU)
U-Net (Baseline)	72.31%	56.05%

V. INTEGRATION AND EVALUATION OF ATTENTION U-NET

To further refine the segmentation boundaries, we extended the baseline model by integrating an Attention Mechanism into the U-Net architecture. While the standard U-Net relies on direct skip connections to concatenate high-resolution features from the encoder with up-sampled features from decoder, this can sometimes introduce irrelevant background noise into the spatial representation.

To explore this, the Attention U-Net [17] introduces soft attention gate to the skip connections. As defined in our architecture, the attention block takes two inputs: the spatial feature map from the skip connection path and a gating signal from the expansive path. The Attention U-Net model was trained using the exact same hyperparameters, feature engineering pipeline and negative Jaccard Coefficient loss function as the baseline to ensure fair comparison. Upon evaluating the test split, the Attention U-Net demonstrated the significant shift in predictive behaviour. As detailed in Table VI, the inclusion of the attention gates yielded much higher Recall of 85.19% (Compared to baselines 72.19%). This indicates that the attention mechanism made the model highly sensitive, successfully detecting a vast majority of the true landslide pixels and minimizing the false negatives.

However, this increased sensitivity came at the cost of strict localization. The model achieved a Precision of 54.50% and a Jaccard coefficient (mIoU) of 5.23%. This trade-off suggests that while the attention gates aggressively highlighted landslide-like topographies, they also introduced a higher rate of false positives by over-activating on background features with similar spectral signatures. Consequently, the overall F1 score dropped to 66.32%, indicating that for this specific dataset configuration, the baseline U-Net provides a more balanced and strict boundary delineation without further hyperparameter tuning of the attention modules.

Table VI

PERFORMANCE COMPARISON OF SEGMENTATION MODELS ON LANDSLIDE4SENSE DATASET

Network	Accuracy	Precision	Recall
U-Net (Baseline)	98.67%	72.67%	72.19%
Attention U-Net	97.92%	54.50%	85.19%

VI. PROPOSE ASPP-ARUNET FOR LANDSLIDE SEGMENTATION AND DETECTION

Based on the extensive structural refinements and feature engineering conducted above, we indicate that

leveraging a combination of targeted deep-learning techniques significantly improves segmentation performance in complex topographical environments. Specifically, the integration of a weighted combined loss function (0.2 Binary Cross-Entropy and 0.8 Dice Loss), 9 strategically engineered input channels (combining normalized multispectral bands with derived topographical indices such as Slope-NDVI and Slope-NDMI), and a robust, synchronized spatial augmentation pipeline proves highly effective. The implemented network architecture leverages a combination of advanced deep learning components—namely, Residual blocks, Attention Gates, and Atrous Spatial Pyramid Pooling (ASPP)—alongside dynamic threshold-based post-processing to achieve robust segmentation.

Given these advanced techniques, we propose a novel network architecture specifically optimized for the landslide segmentation task, referred to as ASPP-ARUNet (ASPP-Guided Attention Residual U-Net). The proposed model synthesizes three critical components to address the unique challenges of satellite imagery:

- **Residual Convolutional Blocks:** Replacing standard convolutions with Res-Conv blocks throughout the encoder and decoder to mitigate vanishing gradients and ensure the propagation of fine-grained spatial features.
- **ASPP Bottleneck:** The implementation of Atrous Spatial Pyramid Pooling at the network's bottleneck. By utilizing parallel dilated convolutions (with rates of 1, 6, and 12), the model captures multi-scale contextual information—crucial for identifying landslides of varying sizes—without degrading the spatial resolution of the feature maps.
- **Soft Attention Gates:** Applied within the expansive path, these gates utilize a 1x1 gating signal to actively suppress irrelevant background features (such as clouds or stable bare ground) while amplifying the salient features of landslide boundaries before concatenation.

To adapt to the severe class imbalance inherent in the LandSlide4Sense dataset, the proposed ASPP-ARUNet relies heavily on spatial overlap metrics rather than standard accuracy. Evaluated on the LandSlide4Sense dataset, the proposed ASPP-ARUNet provides a highly robust, end-to-end framework for high-fidelity boundary segmentation. As shown in Table VII

Table VII
PERFORMANCE COMPARISON LANDSLIDE4SENSE DATASET

Network	F1	Precision	Recall	IoU

U-Net (Baseline)	72.31%	72.67%	72.19%	56.05%
Attention U-Net	66.32%	54.50%	85.19%	5.23%
Our Proposed ASPP-ARUNET	75.06%	75.85%	74.54%	59.75%

Some segmentation results obtained from ASPP-ARNet model on LandSlide4Sense dataset are shown in Fig. 2

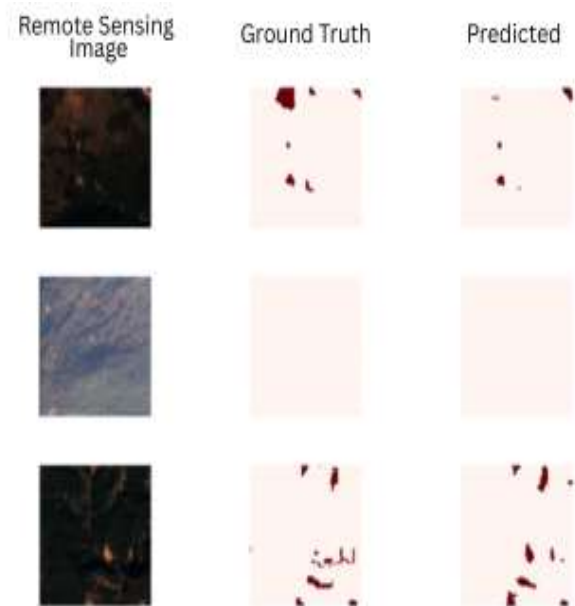


Fig.2. Segmentation Result Obtained from ASPP-ARUNet

VII.CONCLUSION

We have presented a deep learning approach for landslide segmentation from remote sensing imagery. By evaluating the effect of various improvements of feature engineering, network architecture, loss function and post processing. We finally construct the ASPP-ARUNet, which bases on Attention Residual U-net architecture. The extensive experiments prove that our proposed ASPP-ARUNet is robust on LandSlide4Sense dataset, which shows potential to apply remote sensing image based landslide analysis system.

References

- [1] Wikipedia, "https://en.wikipedia.org/wiki/Landslide"
- [2] 5 killed, 11 missing as fresh cloudbursts, landslides hit uttarakhand, "https://www.ndtv.com/india-news/5-killed-11-

missing-as-fresh-cloudbursts-landslides-hit-uttarakhand-19184621".

[3] Introduction to Landslide Evaluation Tools- Mapping, Remote Sensing, and Monitoring of Landslides, "https://pubs.usgs.gov/circ/1325/pdf/Sections/AppendixB.pdf"

[4] Predicting Landslide Using Machine Learning Techniques, "https://www.researchgate.net/publication/382302854_Predicting_Landslide_Using_Machine_Learning_Techniques"

[5] Hesheng Chen, Yi He, Lifeng Zhang, Wang Yang, Yaoxiang Liu, Binghai Gao, "A Multi-Input Channel U-Net Landslide Detection Method Fusing SAR Multisource Remote Sensing Data ", December 05-2032

[6] Yingxu Song, Yujia Zou, Yuan Li, Yueshun He, Weicheng Wu, Ruiqing Niu and ShuaiXu, "Enhancing Landslide Detection with SBConv-Optimized U-Net Architecture Based on Multisource Remote Sensing Data", June 12-2024

[7] Landslide4Sense: Multi-sensor landslide detection competition & benchmark dataset, "https://github.com/iarai/Landslide4Sense-2022"

[8] Omid Ghorbanzadeh, Yonghao Xu, Hengwei Zhao, Junjue Wang, Yanfei Zhong, Dong Zhao, Qi Zang, Shuang Wang, Fahong Zhang, Yilei Shi, Xiao Xiang Zhu, Lin Bai, Weile Li, Weihang Peng, and Pedram Ghamisi, "The outcome of the 2022 landslide4sense competition: Advanced landslide detection from multisource satellite imagery," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 9927–9942, 2022[9] Rekha R.Nair, Tina Babu, Tripti Singh, Afnaan K, "Unet and Semantic Segmentation Based Landslide Detection System," International Conference on Computing for Sustainable Global Development (INDIACom), August 21-2025

[10] Syed Mujtaba Hussaine, Linlong Mu, Yimin Lu & Syed Sajid Hussain, "Landslide Image Segmentation with Attention Residual U-Net: A Hybrid Deep Learning Model",Procedia Computer Science, vol. 258, pp. 2029-2039, 2025

[11] Meng Tang, Yuelin He, Muhammed Aslam, Edore Akpokodje & Syeda Fizzah Jilani, "Enhanced U-Net++ for Improved Semantic Segmentation in Landslide Detection", in Sensing in Harsh Environments: Power, Communication and Material Challenges, vol. 25, April 23-2025

[12] Ayush Srivastava, Priyanshu Bansal, Aayush Kumar Jha and Brindha Subburaj, "Elevation Aware U-Net Based Landslide Segmentation Leveraging DEM and Saliency Map", in 2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC) , July 26-2025

[13] Lam Pham , Cam Le, Hieu Tang, Khang Truong, Truong Nguyen, Jasmin Lampert, Alexander Schindler, Martin Boyer and Son Phan , "RMAU-Net: A Residual-Multihead-Attention U-Net Architecture for Landslide Segmentation and Detection from Remote Sensing Images", Computer Vision and Pattern Recognition , July 15-2025

[14] Landslide4Sense Dataset Kaggle, "https://www.kaggle.com/datasets/tekbahadurkshetri/landslide4sense"

[15] HDF5 Wikipedia, "https://en.wikipedia.org/wiki/Hierarchical_Data_Format"

[16] P. K. Diederik and B. Jimmy, "Adam: A method for stochastic optimization," CoRR, vol. abs/1412.6980, 2015

[17] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori , Steven

McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker and Daniel Ruecket, "Attention U-Net: Learning Where to Look for the Pancreas", Computer vision and pattern recognition, April 11,2018