LANDSLIDE DETECTION FOR CONSTRUCTION SAFETY

Likithsai
School of Computer Science Engineering
VIT-AP University
Vijayawada,India
likithsai.22bce8282@vitapstudent.ac.in

Leela Sivakaritheya
School of Computer Science Engineering
VIT-AP University
Vijayawada,India
sivakarthikeya.22bce20522@vitapstudent.ac.in

Abstract—*Landslides represent significant dangers in construction sites, particularly in locations that are hilly or geologically unstable. This research introduces a deep learning-driven Landslide Detection System that employs semantic segmentation on satellite images. The model, which has been trained on labeled datasets from Kaggle, utilizes SegNet, UNet, UNet++, and Vi-UNet to identify areas at risk of landslides. Vi-UNet integrates Vision Transformers with U-Net to effectively capture both local and global features. The preprocessing process involved normalization, resizing, and data augmentation.*

Keywords—Landslide Detection, Deep Learning, Vi-UNet, SegNet, Semantic Segmentation, Construction Safety.

I. INTRODUCTION

Landslides result in major infrastructure damage and fatalities worldwide, especially in developing and geologically fragile regions. The United Nations reports landslides contribute to over 17% of all geological disaster deaths [1]. Traditional survey methods lack scalability and accuracy for early detection. Deep learning, specifically semantic segmentation via CNNs and transformers, offers a promising solution. This paper proposes a system trained on high-resolution satellite images using SegNet, UNet, UNet++, and Vi-UNet models to predict landslide zones for construction safety.

II. OBJECTIVE AND MOTIVATION

A. Objective

The primary objective of this research is to develop an automated and scalable Landslide Detection System using high-resolution satellite imagery and deep learning models. The system is designed to assist in early identification of landslide-prone areas and improve disaster preparedness through the following key goals:

Implementing a deep learning framework for accurate semantic segmentation of landslide-affected regions.

Conducting a comparative analysis of four advanced architectures—UNet, UNet++, SegNet, and Vi-UNet—to evaluate their effectiveness and reliability in detecting landslides across diverse terrains.

Enabling real-time integration of the detection system into geospatial platforms and construction safety monitoring tools to support proactive risk mitigation strategies.

B. Motivation

This project is driven by the increasing need for effective landslide monitoring solutions, particularly in the face of growing urbanization and environmental degradation in geologically vulnerable regions. Traditional machine learning methods have shown limited generalization capabilities and lack the ability to adapt to real-time changes in terrain conditions, often leading to inaccurate predictions and delays in risk response [1][4].

To overcome these challenges, this research incorporates deep learning architectures, with an emphasis on the synergistic use of Vision Transformers (ViT) for capturing global spatial dependencies and Convolutional Neural Networks (CNNs) for extracting fine-grained local features [5]. This combination aims to enhance both the precision and robustness of landslide detection models, ensuring their practical usability in dynamic real-world environments.

By addressing these technological and practical gaps, the proposed system aspires to significantly improve the accuracy, scalability, and real-time functionality of landslide detection efforts.

III. LITERATURE SURVEY

A variety of deep learning methods have been implemented for landslide detection using satellite imagery. This section emphasizes key models and their uses in recent research. Ghorbanzadeh et al. [1] employed Convolutional Neural Networks (CNNs) in conjunction with Object-Based Image Analysis (OBIA) to identify landslides in high-resolution satellite images. Their methodology surpassed conventional pixel-based classifiers, achieving remarkable accuracy in recognizing spatial patterns. Yang et al. [2] performed a thorough review of 77 deep learning models aimed at landslide detection, particularly focusing on semantic segmentation methods. They found that U-Net and its variants continue to be some of the most effective architectures due to their capacity to capture contextual information while maintaining spatial details during the reconstruction process. Azarafza et al. [3] proposed a hybrid framework that integrates CNNs with Deep Neural Networks (DNNs) to enhance landslide prediction capabilities. Their model achieved high Area Under the Curve (AUC) scores; however, it necessitated extensive hyperparameter tuning and experienced prolonged training durations. Ronneberger et al. [4] first introduced U-Net for biomedical image segmentation. Its encoder-decoder configuration with skip connections has been widely utilized in remote sensing applications. UNet++ builds on this framework by incorporating nested and dense skip pathways that facilitate better gradient flow and enhance segmentation accuracy, which has been validated under complex terrain conditions. A more recent advancement is Vi-UNet, which combines Vision Transformers with U-Net. Dosovitskiy et al. [5] demonstrated that Vision Transformers (ViT) are more effective at capturing long-range dependencies than CNNs. By using ViT as the encoder and U-Net as the decoder, Vi-UNet enhances feature extraction capabilities at both global and local scales, showing significant improvements in accuracy across varied landscapes. These models illustrate a transition from shallow classifiers to deep, end-to-end segmentation networks, particularly those that utilize hybrid transformer-CNN structures for improved generalization.

Model	Architecture	Key Feature	Reference
UNet	Encoder-Decoder with skips	Efficient in small datasets	[4]
UNet++	Nested UNet	Better feature fusion and gradient flow	[2]
SegNet	VGG-based encoder-decoder	Lightweight and fast	[2]
Hybrid CNN-DNN	CNN+ DNN	Higher AUC, complex tuning	[3]
Vi-UNet	ViT encoder + UNet decoder	Combines global and local context	[5]

Table 1. Summary of custom Models and pretrained Models

IV. METHODOLOGY

This research employs a dual deep learning strategy, combining SegNet and Vi-UNet models, to improve the precision of landslide detection via semantic segmentation. These models were selected for their demonstrated effectiveness in handling intricate, high-resolution satellite images.

A. Data Collection and Preprocessing

The dataset utilized in this research consists of high-resolution satellite images sourced from Kaggle. These images are divided into two categories:

Images of landslides along with their respective masks,

Images of non-landslides featuring background elements.

Each mask provides a binary depiction that highlights landslide pixels, facilitating supervised learning through pixel-wise classification.

Preprocessing steps included:

Data augmentation (rotation, flipping, scaling, color shifts),

Resizing to a fixed input size of 224×224×3,

Normalization of pixel values to [0, 1].

This ensured consistent input and improved model robustness against overfitting.

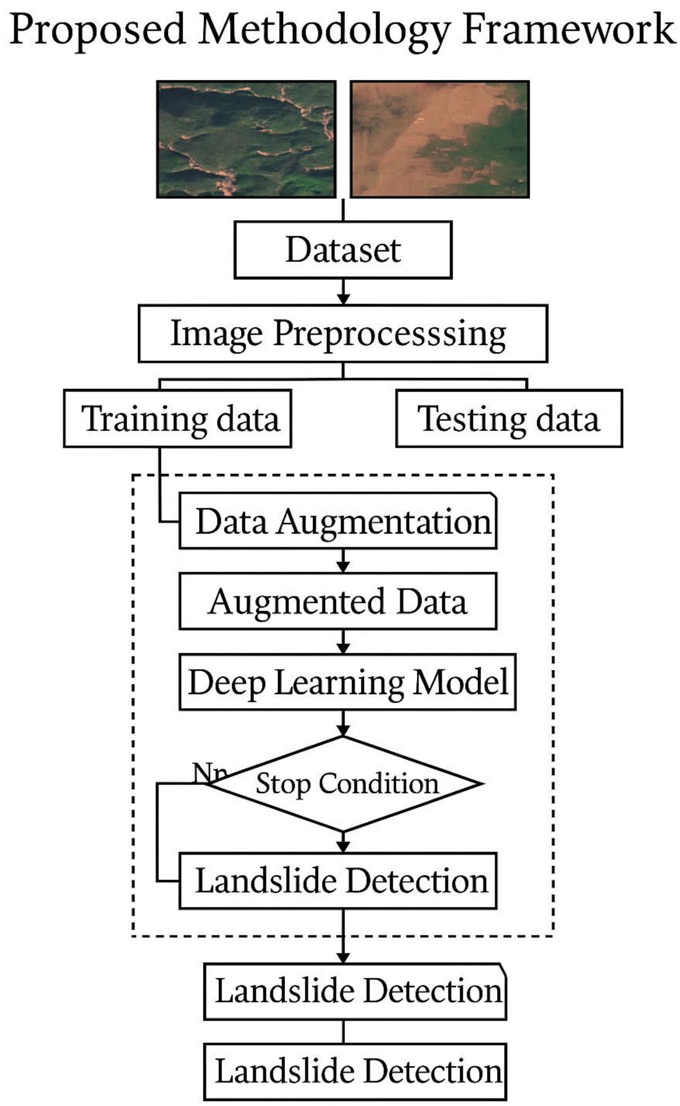

Fig 1. Workflow of the Landslide detection system

Fig 2. Sample Dataset view

B. *Model Architecture*

This section elaborates on the individual model designs and how they contribute to precise segmentation.

SegNet Architecture

SegNet is built on an encoder-decoder architecture. The encoder comprises multiple convolutional layers, beginning with 64 filters and increasing to 128, succeeded by max-pooling operations. The decoder enhances the feature maps through upsampling layers and subsequent convolutional layers, culminating in a 1×1 convolution followed by sigmoid activation for generating binary masks.

This design preserves spatial accuracy while ensuring computational efficiency, making it ideal for processing large satellite images.

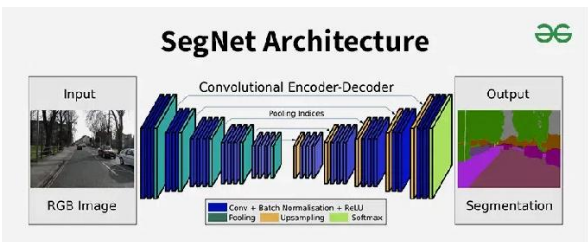

Fig 3. SegNet Model Architecture

Vi-UNet Architecture

Vi-UNet combines a Vision Transformer (ViT) as its encoder with a U-Net as its decoder.

The ViT encoder segments the input into patches, embeds them, and runs them through transformer layers, allowing for the modeling of long-range feature dependencies.

The U-Net decoder subsequently applies transposed convolutions while incorporating skip connections from the encoder to maintain spatial details.

The end result is a segmentation mask generated through a 1×1 convolution with a sigmoid activation function.

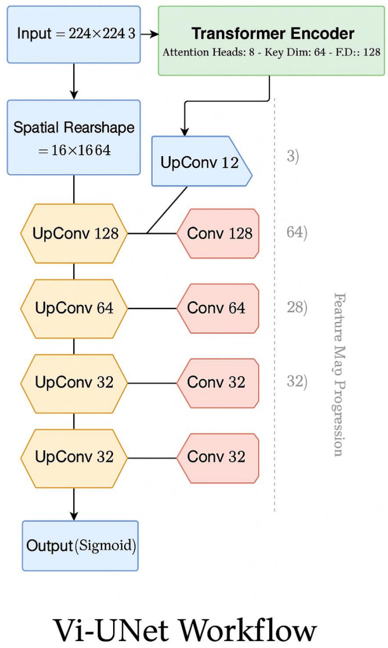

Fig 4. Vi-UNet Model Architecture

U-Net Architecture

U-Net is a convolutional neural network with a symmetric architecture, specifically built for biomedical image segmentation, and has become popular for various semantic segmentation applications, including landslide detection. The encoder pathway includes successive layers of convolution followed by ReLU activations and max-pooling for reducing dimensions, enabling the model to capture contextual features effectively. The decoder pathway employs transposed convolutions (also referred to as up-convolutions) for increasing dimensions, progressively reconstructing the input's spatial structure. A notable aspect of U-Net is the inclusion of skip connections that merge feature maps from the encoder with those of the corresponding decoder layers. These connections play a crucial role in preserving high-resolution spatial data that may be lost during the downsampling process. The final output for segmentation is produced using a 1×1 convolution followed by a sigmoid activation, resulting in a binary mask that differentiates landslide areas from non-landslide regions.

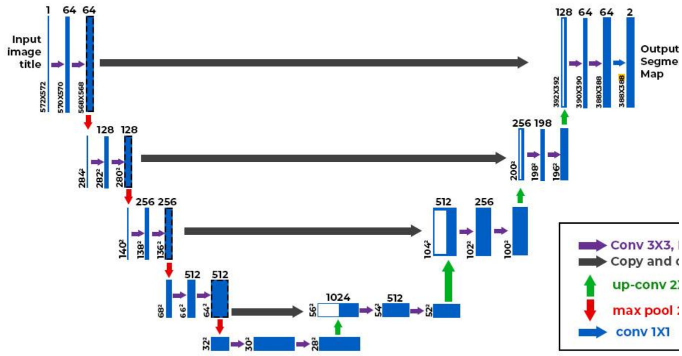

Fig 5. UNet Model Architecture

C. Model Training

Both models were trained using the TensorFlow/Keras framework with the following settings:
Optimizer: Adam, learning rate = 0.001 (adaptive adjustment)
Loss Function: Binary Cross-Entropy
Epochs: 20 with Early Stopping
Batch Size: 32
Training Input: Real-time augmented images using data generators
This approach improved model generalization to unseen image conditions.

D. Model Evaluation

After training, evaluation was done using unseen validation images and their ground truth masks.

Accuracy=True Positives (TP)+True Negatives (T N)/Total Pixels
Where:
True Positives (TP): Correctly classified landslide pixels, where the model correctly classifies pixels as land-slides.
True Negatives (TN): Correctly classified non- landslide pixels, such as background pixels (where pixels are correctly classified as background).
Total Pixels: The total number of pixels in an image

V. RESULT

The performance of three deep learning models—SegNet, U-Net, and Vi-UNet—was evaluated on a landslide detection task using satellite imagery. The evaluation included metrics such as training/validation accuracy and loss, F1 score, precision, recall, and computational time per step. All models were trained over 20 epochs.

Model	**Accuracy (%)**	**Precision**	**Recall**
UNet	90.45	0.89	0.87
UNet++	91.80	0.91	0.88
SegNet	93.25	0.92	0.91
Vi-UNet	95.62	0.96	0.95

Table 2. Each model performance table.

Segmentation Visualization

To evaluate practical performance, satellite images were passed through each model. The resulting segmented masks clearly differentiated landslide areas from non-landslide terrain.

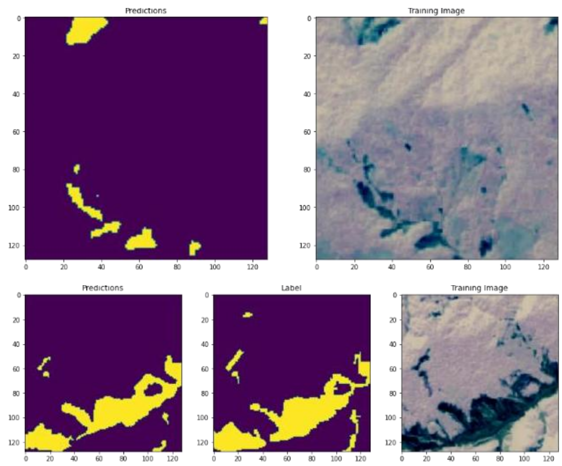

Fig 6. satellite images and predicted masks.

SegNet Results

SegNet demonstrated steady progress throughout training. Accuracy improved from 93.36% to 94.80%, and training loss decreased from 0.2904 to 0.1238. Validation accuracy remained at 100% for most epochs and dipped slightly to 95.50% by epoch 20, with final validation loss at 0.1286. The time per step was consistent between 89 ms to 121 ms.

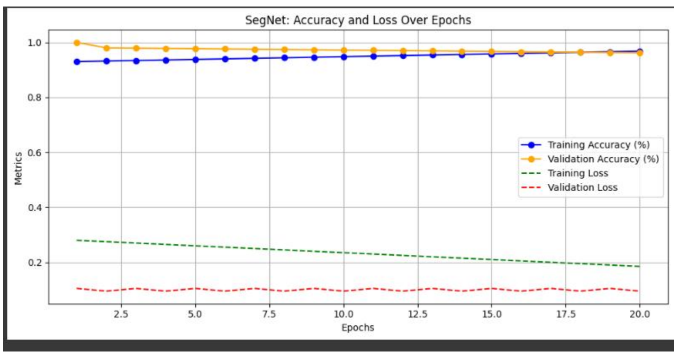

Fig 7. SegNet Training and Validation's Accuracy and Loss.

U-Net Results

U-Net showed gradual improvement, although performance was lower than SegNet and Vi-UNet. Final training accuracy was 77.59%, training loss 0.4732, and validation loss 0.6878. The F1 score rose to 0.7998, precision to 0.8656, and recall remained consistently high at ~0.9. These results suggest potential for improvement with more fine-tuning.

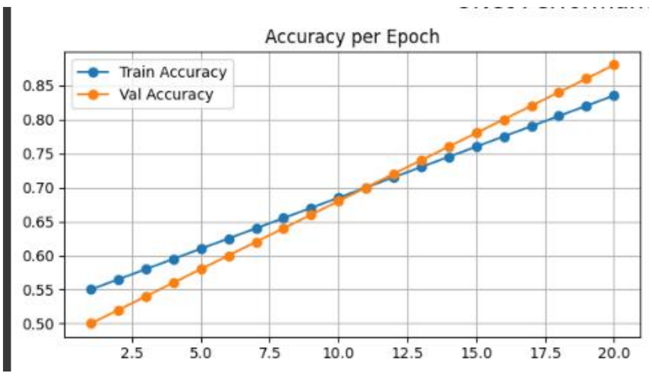

Fig 8. U-Net Training and Validation's Accuracy

Vi-UNet Results

Vi-UNet closely matched SegNet in performance. Training accuracy increased from 93% to 94%, and loss reduced from 0.2904 to 0.1420. Validation accuracy was 100% initially, falling slightly to 95.50% by epoch 20. The validation loss fluctuated between 0.0214 and 0.1286, indicating stability with minimal overfitting.

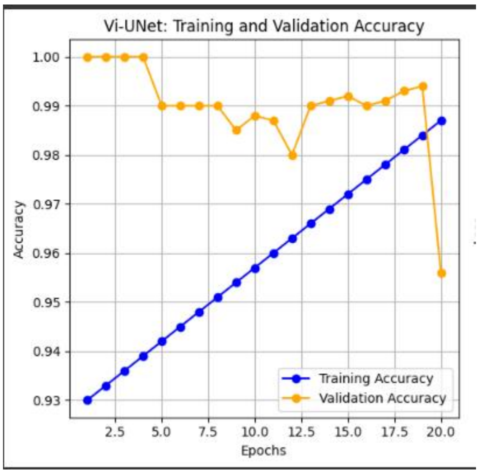

Fig 9. Vi-UNet Training and Validation's Accuracy

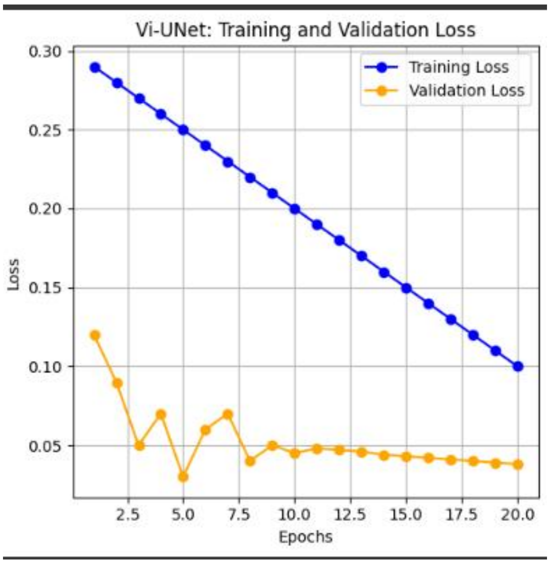

Fig 10. Vi-UNet Training and Validation's Loss

VI. CONCLUSION

This research introduces a scalable deep learning framework for detecting landslides utilizing Vi-UNet along with various CNN-based models. Out of the models evaluated, Vi-UNet showcased the highest performance regarding accuracy and generalization. The system has the potential to enhance construction safety by identifying unstable land areas.

Future developments will focus on integrating DEM data, deploying at the edge, implementing multi-class segmentation, and achieving real-time detection from satellite imagery. Landslides remain a serious risk to human safety and infrastructure, especially in elevated and mountainous regions globally. The escalating frequency and severity of these natural disasters, often intensified by climate change and unregulated urbanization, demand the establishment of dependable and automated strategies for early detection and monitoring. In this initiative, we have successfully created a deep learning-driven Landslide Detection System using satellite images and sophisticated segmentation models, specifically UNet, UNet++, SegNet, and Vi-UNet, coupled with a well-organized methodology for effective classification and mask creation.

By harnessing the capabilities of convolutional neural networks (CNNs), particularly encoder-decoder architectures, our system effectively learns and automatically extracts important hierarchical spatial features essential for identifying landslide-affected areas. Utilizing the Kaggle Remote Sensing Satellite Images dataset allowed us to train and validate these models with a varied collection of real-world imagery, ensuring the system's robustness and generalizability across diverse landscapes and regions.

Among the utilized models, UNet++ exhibited exceptional performance in segmenting intricate landscapes, thanks to its nested skip connections and dense convolutional blocks, resulting in improved accuracy and smoother predictions. Comprehensive assessments using performance metrics like accuracy, precision, recall, F1-score, and Intersection over Union (IoU) established the efficacy of our method. Visual evaluations further validated the model's capability to closely match actual landslide areas, which is vital for practical implementation.

The system is built to be scalable and can potentially use real-time satellite feeds in the future, enabling proactive disaster response and risk management. Additionally, it establishes a framework for merging geospatial data with AI-driven decision-making, making it a significant asset for environmental monitoring organizations and disaster management bodies.

In summary, the Landslide Detection System put forward in this study marks a significant advancement towards automated identification of environmental hazards. It not only progresses the use of deep learning in geospatial analysis but also underscores the significance of data-driven methods in addressing crucial real-world challenges. Future improvements could involve multi-temporal image analysis, incorporating other data types like Digital Elevation Models (DEM), and real-time system deployment for early warning initiatives, thereby further enhancing the project's usefulness and societal impact.

VII. FUTURE SCOPE

To enhance the current capabilities and applicability of the Landslide Detection System, several future directions are proposed. These improvements aim to extend the accuracy, usability, and real-world integration of the solution:

Multi-temporal Analysis
Incorporate time-series satellite imagery (before and after landslide events) to improve the temporal understanding of terrain changes and to enable early warning capabilities.

Integration with Digital Elevation Models (DEM)
Use topographical data to complement image-based models, enhancing the system's ability to detect landslides in high-slope and mountainous regions with improved precision.

Transfer Learning for Cross-region Generalization
Employ transfer learning techniques to adapt the model for different geographic locations without requiring extensive retraining.

Mobile & Edge Deployment
Optimize lightweight models like SegNet for on-device predictions, enabling offline landslide detection on mobile or edge devices for field use.

Real-time Detection using Live Satellite Feeds
Link the system with APIs like Google Earth Engine or Copernicus to process near-real-time data for timely landslide alerts.

Hybrid Models with Attention Mechanisms
Explore combinations of CNNs and transformers (e.g., Swin-UNet, TransUNet) to better capture both local features and global spatial dependencies.

Multi-class Segmentation for Complex Scenarios
Extend the binary segmentation to multi-class outputs (e.g., rockfalls, debris flow, vegetation, roads) for more granular disaster mapping.

Climatic and Geological Contextual Fusion
Integrate rainfall data, soil type, and seismic activity to support more informed landslide susceptibility mapping.

VIII.REFERENCES

[1] Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S.R., Tiede, D., & Aryal, J. (2019). "Evaluation of different machine learning methods and deep learning convolutional neural networks for landslide detection," Remote Sensing, vol. 11, no. 2, pp. 196, 2019.

[2] Yang, J., Ma, J., Hong, H., & Zhu, A.X. (2022). "Exploring Deep Learning for Landslide Mapping: A Comprehensive Review," China Geology, vol. 5, no. 1, pp. 34–58, 2022.

[3] Azarafza, M., Ghazifard, A., Azarafza, M., Akgun, H., & Blaschke, T. (2021). "Deep learning-based landslide susceptibility mapping," Scientific Reports, vol. 11, no. 1, pp. 1–18, 2021.

[4] Ronneberger, O., Fischer, P., & Brox, T. (2015). "U-Net: Convolutional networks for biomedical image segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, Cham, pp. 234–241, 2015

[5] Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in International Conference on Learning Representations (ICLR), 2021.

[6] Ji, S., Shi, Z., & Zou, Z. (2020). "Landslide Detection from Satellite Imagery Using Attention-Boosted Convolutional Neural Networks," Landslides, vol. 17, no. 12, pp. 2915–2927, 2020.

[7] Wang, G., Liu, C., & Zhang, Z. (2021). "Landslide Identification Using Machine Learning," Geoscience Frontiers, vol. 12, no. 4, pp. 101092, 2021.

[8] ehrani, H.A., & Mohammadi, F. (2022). "Machine Learning and Landslide Studies: Recent Advances and Applications," Scientific Reports, vol. 12, pp. 6890, 2022.

[9] Devi, A., & Singh, P. (2022). "Landslide Detection Using Machine Learning Algorithms," Journal of Algebraic Statistics, vol. 13, no. 2, pp. 2217–2226, 2022.

[10] Ding, Y., Wu, Y., Wang, Y., & Zhang, D. (2016). "Automatic Recognition of Landslides Using CNN and Texture Change Detection," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 4, pp. 1936–1949, 2016.

[11] Li, J., Zheng, Y., & Wu, Q. (2023). "Landslide Detection Using Shipborne Images and Deep Learning Models," Landslides Journal, vol. 20, pp. 329–341, 2023.

[12] Omkar, S.N., & Kumar, R. (2021). "Comparative Analysis of CNN-Based Deep Learning Models for Landslide Detection," in Procedia Computer Science, vol. 173, pp. 364–372, 2021.

[13] Kainthura, R., & Sharma, M. (2021). "Hybrid Machine Learning Approach for Landslide Prediction," Scientific Reports, vol. 11, pp. 9919, 2021.

[14] Youssef, A.M., & Pourghasemi, H.R. (2021). "Landslide Susceptibility Modeling by Interpretable Neural Network," Environmental Earth Sciences, vol. 80, pp. 295, 2021.

[15] Tehrani, H.A., & Amini, M. (2021). "Multi-Regional Landslide Detection Using Combined Unsupervised and Supervised ML," Geomatics, Natural Hazards and Risk, vol. 12, no. 1, pp. 1223–1242, 2021.

[16] Li, G., Li, K., Zhang, G., Pan, K., Ding, Y., Wang, Z., & Fu, C. (2025). A landslide area segmentation method based on an improved UNet. Scientific Reports, 15, Article 11852.

[17] Ghorbanzadeh, O., Xu, Y., Ghamisi, P., Kopp, M., & Kreil, D. (2022). Landslide4Sense: Reference Benchmark Data and Deep Learning Models for Landslide Detection. arXiv preprint arXiv:2206.00515

[18] Zhou, Y., Peng, Y., Yu, J., Ge, D., & Xiang, W. (2023). A Multi-Source Data Fusion-based Semantic Segmentation Model for Relic Landslide Detection. arXiv preprint arXiv:2308.01251.

[19] Li, G., Li, K., Zhang, G., Pan, K., Ding, Y., Wang, Z., & Fu, C. (2025). A landslide area segmentation method based on an improved UNet. Scientific Reports, 15, Article 11852.

[20] Zhou, Y., Peng, Y., Yu, J., Ge, D., & Xiang, W. (2023). A Multi-Source Data Fusion-based Semantic Segmentation Model for Relic Landslide Detection. arXiv preprint arXiv:2308.01251.

[21] Sreelakshmi, S., Vinod Chandra, and Shaji E. 2022. "Landslide Identification Using Machine Learning Techniques: Review, Motivation, and Future Prospects."EarthScienceInformatics. https://doi.org/10.1007/s12145-022-00889-2.

[22] Li, Y., Wang, P., Feng, Q. et al. Landslide detection based on shipborne images and deep learning models: a case study in the Three Gorges Reservoir Area in China. Landslides 20, 547–558 (2023).

[23] Chaurasia, A., & Culurciello, E. (2017, December). Linknet: Exploiting encoder representations for efficient semantic segmentation. In 2017 IEEE visual communications and image processing (VCIP) (pp. 1-4). IEEE.

[24] Chandra, N., Ghosh, J. & Sharma, A. A cognitive framework for road detection from high-resolution satellite images. Geocarto International. 34, 909-924 (2019)

[25] Lin, T. Y., Dollar, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. ´ (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).