# Exploring the Capability of TransUNet Model for Landslide Classification

Shang-Nien Tsai [1], Fuan Tsai.[1, 2*]

[1] Department of Civil Engineering, National Central University, 300, Zhongda Rd., Zhongli, Taoyuan 32001, Taiwan

[2] Center for Space and Remote Sensing Research, National Central University, 300, Zhongda Rd., Zhongli, Taoyuan 32001, Taiwan

ftsai@csrsr.ncu.edu.tw

## 1. Introduction

Landslides are severe natural disasters due to their destructiveness and unpredictability, significantly impacting human lives and property. In Taiwan, notable events like the April 2024 Hualien Earthquake and Typhoon Morakot in August 2009 highlight the devastating impact of landslides. The 2024 Hualien Earthquake triggered numerous landslides across affected regions, while Typhoon Morakot caused extensive landslides and destroyed Xiaolin Village. Therefore, prompt and precise identification is essential for disaster prevention and management. (Li & Chen, 2024)

Deep learning techniques, such as Convolutional Neural Networks (CNNs), have demonstrated exceptional performance in feature extraction, significantly reducing the time required to identify relevant landslide features. (Yang, Xu, & Li, 2022) However, CNNs are limited to primarily extracting local features and have weaker capabilities for capturing global context. To overcome this limitation, the Transformer architecture has been introduced, excelling at capturing global contextual information and enhancing the model's ability to handle long-range dependencies. (Yang, Xu, & Li, 2022)

This study employs the TransUNet model, which combines CNN and Transformer structures, to evaluate its efficacy in landslide classification. Originally designed for medical image segmentation, TransUNet is investigated here for its potential in remote sensing applications, specifically for detecting and classifying landslide-affected areas in Taiwan's Laonong River Basin.

## 2. MATERIALS AND METHODS

### 2.1 Study area:

Taiwan's heavy rainfall, complex terrain, and frequent earthquakes make it particularly susceptible to landslides. Therefore, this study focuses on the Laonong

River Basin in Taiwan, as shown in the figure below. The study area includes three zones: the First Area and Second Area are designated for training and validation with two-class and five-class labels, respectively, while the Test Area is used to evaluate the model's performance after training.

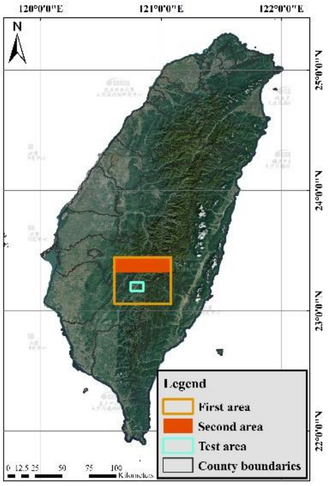



Figure 1: Study area.

**2.2 Dataset:**

This study applies the TransUNet model for landslide classification using satellite imagery. The dataset includes pre- and post-event Spot7 images (Nov 19, 2016, and Nov 14, 2022) with 6-meter multispectral and 1.5-meter panchromatic resolutions, covering B, G, R, and NIR bands. Additional data includes 5-meter (2016) and 6-meter (2022) DSMs, slope data, NDVI, and GLCM indices. Labels are derived from Taiwan's landslide inventory maps for 2016 and 2022.

**2.3 TransUNet model**

TransUNet combines the strengths of Convolutional Neural Networks (CNNs) and Transformers, making it particularly effective in feature extraction. In its architecture, CNN layers are first used to extract local features, providing detailed spatial information. These features are then passed through a series of Transformer layers, which capture global contextual information by modeling long-range dependencies. This combination allows TransUNet to leverage both local and global features, leading to superior segmentation performance. Additionally, the model employs multiple upsampling and feature concatenation layers, enhancing its ability to retain fine details and provide precise segmentation outputs.

**2.4 Procedure**

**2.4.1 Data Pre-processing**

Using Spot7 images from November 19, 2016 (pre-event) and November 14, 2022 (post-event), TOAR (Top of Atmosphere Reflectance) values were computed to standardize reflectance and reduce the influence of solar irradiance variations across different periods. In the image processing stage, multispectral and panchromatic images were merged and translated, followed by resampling to a 1.5-meter resolution. Pansharpening was applied to enhance image details, resulting in high-resolution pre- and post-event images. Additionally, Digital Surface Models (DSM) from 2016 and 2022 were resampled to a 1.5-meter resolution to match the image data, and slope information was derived from these DSMs to capture terrain variability. For labeling, two classification schemes were defined: a two-class scheme with labels "landslide" and "non-landslide," and a five-class scheme with more refined categories, including "unchanged," "non-landslide," "old landslide," "new landslide," and "vegetation reclaim." Feature extraction involved calculating NDVI and GLCM indices from the pre- and post-event images, followed by normalizing the data into 16 bands (R, G, B, NIR, DSM, slope, NDVI, and GLCM for each time period).

### 2.4.2 Dataset Preparation

The processed data were divided into training, validation, and testing sets. Each dataset was cropped into 224x224 tiles. Data from the training areas were used for both training and validation (further split into 80% for training and 20% for validation), while data from the testing areas were reserved exclusively for testing. The images and labels were stored in .npz files, structured for efficient model input with 16-band imagery and either 2-class or 5-class labels.

### 2.4.3 Model Training

The TransUNet model was trained in two types of training schemes. The first scheme used a two-class labels to verify the model's ability to identify landslide-affected areas. The second scheme used five-class labels to evaluate the model's capability in categorizing different types of landslide-affected areas.

### 2.4.4 Model Testing and Evaluation

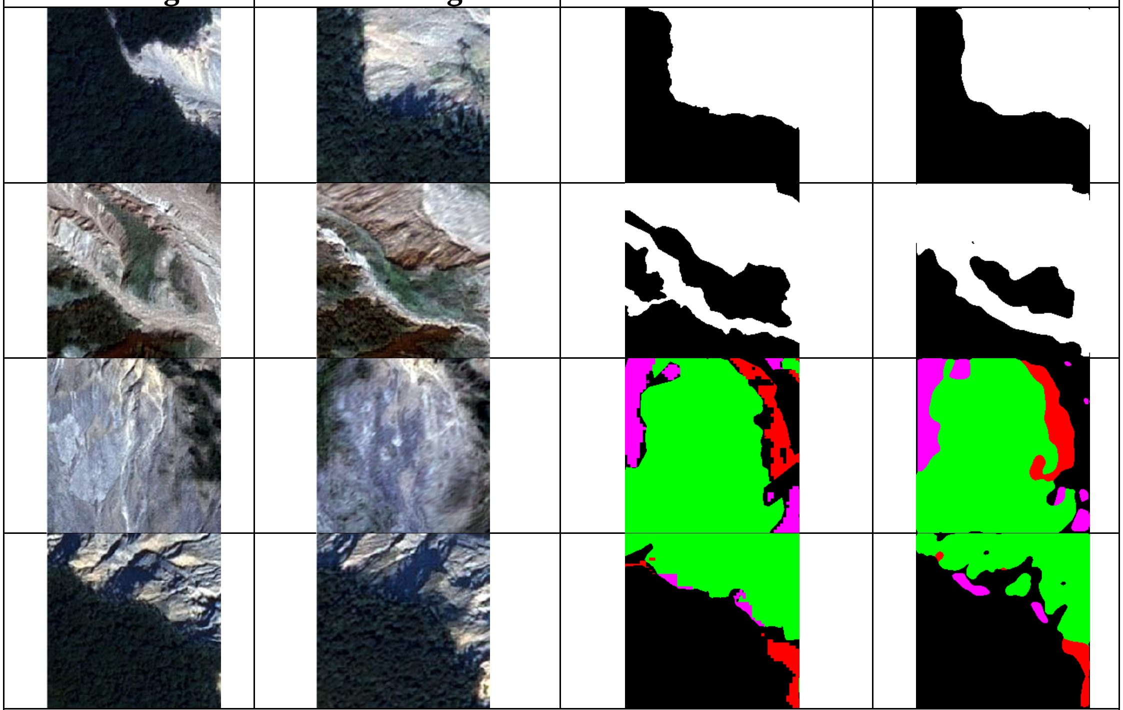
The trained models were evaluated using the testing dataset to assess their accuracy. The results were validated against ground truth labels, and precision, recall, and F1-score metrics were calculated to measure the performance of the TransUNet model in classifying landslide-affected areas.

## 3. Results and Discussion

Preliminary findings indicate that TransUNet effectively detects landslide-affected areas and can classify landslide types. In the two-class labeling scheme, the testing results, as shown in the top two rows of Table 1, demonstrate the model's capability to identify the location and contours of landslides, achieving an Average Precision of 0.916578, Recall of 0.831822, and F1-score of 0.872145. In the five-class labeling scheme, the testing results, as shown in the bottom two rows of Table 1, illustrate the model's ability to classify different types of landslide-affected areas, with an Average Precision of 0.764655, Average Recall of 0.831303, and Average F1-score of 0.796587. However, some misclassifications still occur in smaller areas, indicating room for improvement in detecting finer details.

Table 1: 2-class and 5-class labels testing results

| Pre image | Post image | Label | Prediction |
|---|---|---|---|





The testing results show that while the model can classify landslide-affected areas in the five-class scheme, challenges persist in smaller areas. Categories with fewer samples, like "non-landslide," are difficult for the model to classify accurately. Dividing labels into five categories also reduces sample sizes per class, lowering training effectiveness compared to the two-class scheme. Increasing sample sizes for each category will be crucial for improving multi-class classification accuracy in future work.

## 4. Conclusion and Recommendation

This study demonstrates the potential of the TransUNet model for classifying landslide-affected areas in Taiwan's Laonong River Basin using satellite imagery. The two-class labeling scheme performed well in identifying landslide locations and boundaries, achieving high precision, recall, and F1-score. However, the five-class scheme faced challenges, particularly in smaller or less frequent categories like "non-landslide," likely due to reduced sample sizes.

To improve multi-class classification accuracy, future work will explore a two-stage training approach, using initial classification outputs as additional inputs to increase sample. Integrating other deep learning models will also be considered to enhance precision in complex categories. These efforts aim to improve model generalization and support more accurate landslide mapping, aiding disaster management in landslide-prone areas.

## References

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., ... & Zhou, Y. (2021). TransUNet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306.

Huang, P.-J. (2021). Landslide Detection with Machine Learning Analysis of Satellite Images [Master’s thesis, National Central University]. Taiwan Thesis Knowledge Added Value System. https://hdl.handle.net/11296/uu227q

Li, X., & Chen, T. (2024, July). Landslide recognition based on CNN and Transformer feature fusion. In IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium (pp. 8246-8249). IEEE.

Yang, Z., Xu, C., & Li, L. (2022). Landslide detection based on ResU-Net with transformer and CBAM embedded: Two examples with geologically different environments. Remote Sensing, 14(12), 2885.

Yao, W., Bai, J., Liao, W., Chen, Y., Liu, M., & Xie, Y. (2024). From CNN to transformer: A review of medical image segmentation models. Journal of Imaging Informatics in Medicine, 1-19.

Yu, M., & Qin, F. (2023). Research on the applicability of Transformer model in remote-sensing image segmentation. Applied Sciences, 13(4), 2261.