iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** III **Month of publication:** March 2025

DOI: https://doi.org/10.22214/ijraset.2025.67275

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

Evaluating Local-Global Landslide Features Combining CNN and Transformer for Landslide Susceptibility Mapping

Ashlin James[1], Mr. Ashish L[2]
[1]*MCA Scholor,* [2]*Assistant Professor, Department of MCA, Nehru College of Engineering and Research Center, Pambady*

Abstract: ***Landslide susceptibility mapping (LSM) is essential for assessing landslide risk and preventing geological hazards. Despite the advances in deep learning, convolutional neural networks (CNNs) and transformer models still face challenges in achieving optimal mapping accuracy and effectively extracting multilevel landslide features. This study introduces CTLGNet, a CNN-transformer local-global feature extraction network, combining the strengths of both models to capture both local and global landslide features. We applied CTLGNet to LSM in the Three Gorges Reservoir and Jiuzhaigou, using nine landslide conditioning factors to construct the dataset. The dataset was randomly split into training, validation, and test sets (6:2:2 ratio). CTLGNet was compared to CNN, ResNet, DenseNet, ViT, and FrIT using various evaluation metrics. The results showed that CTLGNet outperforms all other models in terms of landslide prediction, with AUC values of 0.9817 and 0.9693 for the two regions. Although its Recall was slightly lower than some models, CTLGNet effectively extracts both local and global landslide features, achieving precise landslide localization and detail capture. Overall, CTLGNet excels in multilevel feature extraction and demonstrates strong potential for widespread LSM applications.***

Keywords: Convolutional neural network (CNN), landslide local-global features, landslide susceptibility mapping (LSM), transformer.

I. INTRODUCTION

Landslides are one of the most devastating natural disasters, threatening human life, property, and sustainable development. Landslide susceptibility mapping (LSM), which estimates the spatial likelihood of landslides based on geological, environmental, and historical data, provides crucial information for disaster mitigation and planning[1]. Many academics have evaluated different machine learning (ML) algorithms in LSM tasks during the past 20 years due to the rapid development of artificial intelligence (AI)[2],[3]. Recent advancements in artificial intelligence (AI) have led to the application of machine learning (ML) algorithms for LSM, enabling the automatic detection of patterns in landslide data and modeling complex relationships[4]. However, ML methods often require extensive feature engineering and are prone to issues like overfitting and underfitting. Deep learning (DL), a subset of ML, has shown significant success in imagerelated tasks[5],[6]. DL models, trained on large datasets in parallel computing environments, can automatically capture complex patterns in landslide and conditioning factors (LCFs), making them particularly effective for LSM[7],[8],[9]. Popular DL models, such as convolutional neural networks (CNNs) and transformers, each have distinct strengths and weaknesses[10]. CNNs excel at extracting local features (LLFs) from multisource remote sensing data but struggle with processing global contextual information due to their limited receptive field[11].

CNN-based models were able to effectively extract landslide local features (LLFs) from data because of the spatial translation invariance and minimal inductive bias of convolutions Remote sensing photos from multiple sources [13]. However, CNN-based models' small receptive field limits their capacity to handle large amounts of contextual data, which makes it difficult to capture landslide global features (LGFs) [14]. Transformer is an encoder-decoder sequence transformation model that captures global contextual information by expanding the receptive field through the use of the self-attention (SA) mechanism. Vision Transformer (ViT) is a noteworthy application that effectively uses the pure transformer backbone for image classification tasks, resulting in competitive performance in LSM tasks [15]. Compared to CNN, the SA mechanism enables the transformer to extract LGFs more accurately by calculating the correlations and weights between each pixel and every other pixel in the image [16]. Transformer-based models, however, usually need to be trained on vast amounts of data and have comparatively poor LLF extraction capabilities [17]. Transformers, particularly the vision transformer (ViT), use the self-attention mechanism to capture global features (LGFs) but are less effective at extracting local features and require large-scale training data.

Both LGFs and LLFs are crucial for LSM. LGFs provide key information about the size, location, and distribution of landslides, while LLFs capture specific characteristics like edges, textures, and shapes. Combining the strengths of CNNs and transformers can improve the extraction of both types of features, enhancing LSM accuracy. Hybrid CNN-transformer models have been successfully used in tasks like image classification and semantic segmentation, but their effectiveness in LSM has not been extensively studied. To address this gap, we propose a hybrid CNN-transformer model, CTLGNet, designed specifically for LSM tasks. CTLGNet combines CNNs for LLF extraction and transformers for LGF modeling, enabling comprehensive feature extraction and improved landslide prediction. This paper presents a comparative analysis of CTLGNet with individual models, evaluating their capabilities in feature extraction and computational efficiency, to enhance the accuracy and reliability of LSM.

II. STUDY AREA AND DATA

Two landslide-prone areas were chosen as study areas to confirm the universality and dependability of the suggested model.

1) Site A: located in the western Hubei Province of China within the Three Gorges Reservoir area, including Zigui and Badong. It has a subtropical monsoon climate with hot summers and cold winters, experiencing heavy rainfall, especially from June to September, averaging 1100 mm annually. The region's rugged topography and unstable lithology, combined with developed faults, make it prone to landslides. As a key area for water conservation, assessing landslide risk in Site A is crucial.

2 Site B: located in Jiuzhaigou County, Aba Tibetan and Qiang Autonomous Prefecture, Sichuan Province, China. It has a highland humid climate with less precipitation and is characterized by steep topography, deep valleys, and high peaks where the Sichuan Basin meets the Tibetan Plateau. The region's complex geology and active tectonic movement contribute to geological risks. A 7.0 magnitude earthquake in 2017 triggered numerous landslides, posing significant danger to locals and visitors. Therefore, accurate landslide susceptibility mapping (LSM) in Site B is crucial for risk assessment and mitigation.

A. Landslide Inventory Map

The landslide inventory maps for Site A and Site B were created using historical landslide data, remote sensing imagery, and on-site field survey data for Site A, and past landslide documentation and Google Earth imagery for Site B. Site A covers 23.40 km² and contains 202 landslides, mainly along the Yangtze River's reservoir banks. Site B spans 9.51 km² and has nearly 4,000 landslides. These maps show landslide distribution and generate labels for model training in landslide susceptibility mapping.

B. Landslide Conditioning Factors

LCF selection is essential for LSM, as landslides occur in areas with similar geological and environmental conditions to past events. Nine LCFs were chosen: elevation, aspect, slope, lithology, distance from fault (DTF), distance from river (DTR), precipitation, land use/land cover (LULC), and NDVI. The digital elevation model generated the first three LCFs, while lithology and fault data were obtained through vectorization from the National Geological Archive's 1:2,000,000 scale geological map database.TABLE 1 shows detailed information on LCFs in Sites A and B.

LCFs	Site A	Site B
Elevation(m)	[80, 2000]	[1603, 4787]
Aspect	Flat, North, Northeast, East, Southeast, South, Southwest, West, Northwest	
Slope(°)	[0, 78.42]	[0, 78.00]
Lithology	Soft rock, Soft and hard, Hard rock	Quaternary, Neogene, Triassic, Permian, Carboniferous, Devonian
DTF(m)	[0, 8753.58]	[0, 7728.60]
DTR(m)	[0, 4706.22]	[0, 6423.22]
Precipitation (mm)	[274.6, 320.9]	[104.0, 133.5]
LULC	Cropland, Forest, Shrub, Grassland, Water, Snow/Ice, Barren, Impervious, Wetland	
NDVI	[-1, 1]	[-1, 1]

Table 1

III. METHODOLOGY

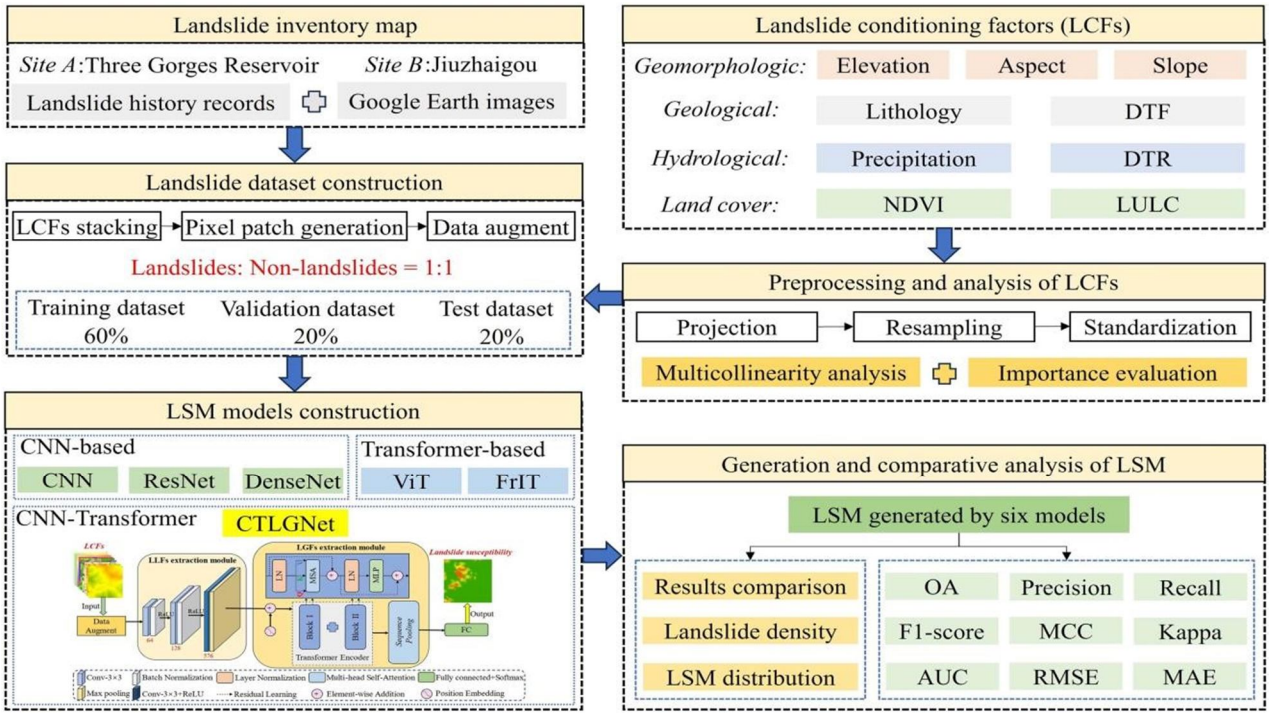


Fig 1

We designate nine geo-environmental parameters as LCFs and analyze their multicollinearity and significance. Patch blocks from images are created to form landslide datasets, which are then split into training, validation, and test sets (6:2:2 ratio). LSM is generated using CNN, ResNet, DenseNet, ViT, FrIT, and CTLGNet. Finally, the models' generality and accuracy are compared. The comprehensive flowchart is shown in Fig. 1.

A. *Preprocessing of Data*

Data preprocessing involved coordinate system unification, resampling, and picture normalization. We standardized the coordinate systems and used bilinear interpolation for resampling to unify spatial resolution. This ensured consistent row and column counts for LCFs. The Z-score method was applied to normalize the LCF images, reducing magnitude differences between variables and speeding up model convergence[20].

B. *LCFs Selection*

Since low importance or significant multicollinearity can have a detrimental effect on model stability and efficiency, it is imperative to ensure the independence and relevance of LCFs throughout the LSM process. In order to generate the landslide datasets, multicollinearity analysis and importance rating are necessary[21].

1) *Multicollinearity Analysis:* We used the Variance Inflation Factor (VIF) and tolerance (TOL) to do a multicollinearity analysis to evaluate the correlation between LCFs. VIF's mathematical expression is as follows, and it aids in quantifying multicollinearity;

$$\text{VIF} = \frac{1}{1 - R_i^2}$$

2) *Evaluation of Importance:* We used the random forest model to compute the Gini index for Sites A and B to assess the significance of each LCF in the occurrence of landslides. When the total of all LCF Gini indices equals 1, a higher Gini coefficient denotes greater relevance.

C. Building a Landslide Dataset

Following preprocessing and LCF selection, the LCFs are stacked into a 3-D matrix (H, W, C), where T is the number of LCFs, W is the width, and H is the height. The center of a patch block, which is enlarged to 72 × 72 × 9, is chosen at random from a landslide pixel. Ten thousand landslide patch blocks and an equal number of non-landslide blocks are produced for Sites A and B. A 6:2:2 ratio is used to separate the dataset into training, validation, and test sets. While keeping the block size for model training constant, data augmentation methods such as random horizontal flipping, rotation, and scaling are applied to the training set to improve sample diversity.

D. CNN-Based Models

1) *CNN:* CNN: The LeNet-5 CNN model, which was first used for digit recognition in the 1990s, is useful for extracting LLFs from input data since it has local perception, parameter sharing, and translation invariance. CNNs have shown great success in landslide susceptibility mapping (LSM) and are frequently used in computer vision (CV). In this work, we used 2-D convolutional kernels to build a ten-layer CNN network for LSM. The architecture efficiently extracts high-dimensional landslide features and produces good LSM results thanks to its two convolutional layers (3 × 3 kernel), two max-pooling layers, two ReLU activation layers, a batch normalization (BN) layer, and a fully connected layer.
2) *ResNet:* ResNet addresses network degradation and gradient vanishing issues by introducing residual learning, where input feature maps are added to the output after convolution operations through skip connections. This allows the network to learn residuals directly. We used ResNet18 for the landslide susceptibility mapping (LSM) model, with an architecture of 18 layers. The residual blocks in ResNet18 help train deeper networks efficiently, ensuring better propagation and retention of landslide features.
3) *DenseNet:* DenseNet addresses gradient vanishing and feature reuse by introducing dense and skip connections, where each layer connects to all previous layers, ensuring rich information flow. Its structure includes multiple dense blocks, each with convolutional layers and a transition layer to control feature map dimensionality and reduce complexity. In this study, we used DenseNet-BC (with bottleneck and compression) to enhance LSM results by reducing feature map dimensionality and improving performance.

E. Transformer-Based Model

1) *VIT:* Introduced by Google in 2020, extends the transformer model to computer vision by treating images as 1-D sequences. It divides the image into patch blocks, flattens them into vectors, and uses an embedding layer to map them to a high-dimensional space. The vectors then pass through multiple encoder layers, including a self-attention (SA) layer, to extract global features. For LSM, we used the original ViT structure with four attention heads and eight transformer layers, effectively capturing landslide global features and focusing on key areas of the image.
2) *FrIT:* Introduced to address the limitations of CNN and ViT by capturing both global and local contextual features[24]. Unlike ViT, which uses the self-attention (SA) mechanism, FrIT utilizes 2-D fractional Fourier transform layers to extract global context. The rest of its structure is similar to ViT. We adapted FrIT for the LSM task and compared its performance with the model proposed in our study.

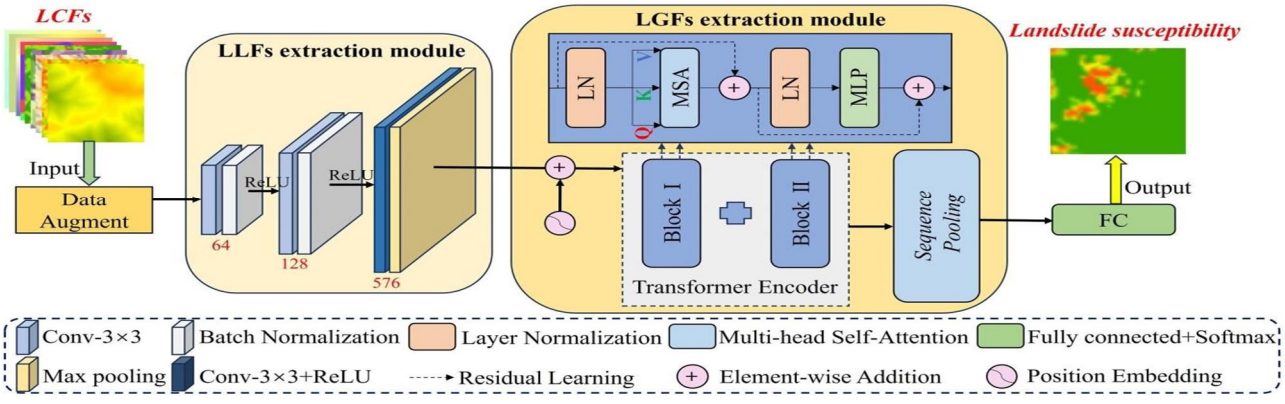


Fig 2

F. CTLGNet

It is a hybrid model designed to improve landslide spatial prediction accuracy by simultaneously extracting local (LLFs) and global features (LGFs). It combines a convolutional architecture for LLF extraction and an improved transformer architecture for LGF extraction, as shown in Fig. 2.

1) The LLFs extraction module in CTLGNet consists of three blocks: two convolution layers (3×3 kernel) with BN and ReLU to capture low- and high-level LLFs, and a final block with a convolutional layer, ReLU, and max pooling to extract higher-level LLFs. This structure ensures the effective extraction of LLFs from landslide areas of various sizes. $X \in R^{H*W*C}$ represents the input image

X_p = MaxPool(Conv(Conv(Conv(X))))

X is transformed into a 1-D sequence $Xp \in R1\times(P^2 \cdot C)$, where p is the resolution of each image patch, enabling the model to effectively capture a broader range of high-dimensional and critical LLFs.

2) LGFs Extraction Module: The LGFs extraction module uses an improved transformer after patch and position embedding. Each transformer encoder includes multihead self-attention (MSA), layer normalization (LN), multilayer perceptron (MLP) with GELU activation, and skip connections. The encoder input is XL.

X'_{L+1}= MLP(LN(X_L))+ X_L

$X_{L+1} = \text{MSA}\left(\text{LN}\left(X'_{L+1}\right)\right) + X'_{L+1}$

In CTLGNet, we replace the class token with sequence pooling (SeqPool), which captures relevant information from different input image regions. SeqPool is an attention-based method that balances the sequential embeddings in the transformer encoder's output. This enhances the model's ability to handle spatially sparse data, improving LSM performance.

G. Model Evaluation Metrics

This study uses several evaluation metrics to assess model performance, including overall accuracy (OA), precision, recall, F1-score, Matthews correlation coefficient (MCC), Kappa coefficient, root mean square error (RMSE), mean absolute error (MAE), area under the curve (AUC), and receiver operating characteristic (ROC). Landslide density (LD) is a measure of landslide distribution within LSM results, classified into five categories: very low (VL), low (L), moderate (M), high (H), and very high (VH). LD is calculated as the ratio of landslide area to the total zone area within the susceptibility zones.

H. Experimental Environment and Hyperparameter Settings

Microsoft Visual Studio 2021, PyCharm 2021, and Python 3.8 were used in the software's development, and the Linux-Ubuntu 18.04.05 LTS operating system was used for testing. DL was constructed using the Keras 2.4.3 framework. Using the stochastic gradient descent optimizer, a learning rate of 0.001, and a batch size of 16, hyperparameters were adjusted by trial and error. The output layer selected cross-entropy loss with label smoothing and employed the "Softmax" activation function. 100 training epochs were conducted, and the best model was chosen by tracking validation accuracy.

IV. RESULTS AND ANALYSIS

A. Analysis of the LCFs

The multicollinearity analysis of LCFs in Sites A and B (Fig. 3) shows low multicollinearity. In Site A, all LCFs have VIF values below 2, with elevation having the highest VIF of 1.933. In Site B, all LCFs have VIF values below 5, with elevation and precipitation having the highest VIFs of 4.560 and 4.649, respectively. Fig. 4 shows that all LCFs influence landslide occurrence to varying degrees, with DTF having low importance in both sites. NDVI is less important in Site A but significant in Site B, while slope is highly important in Site A but less so in Site B. Due to the lack of strong multicollinearity and the importance of all LCFs, we will use them to obtain LSM results.

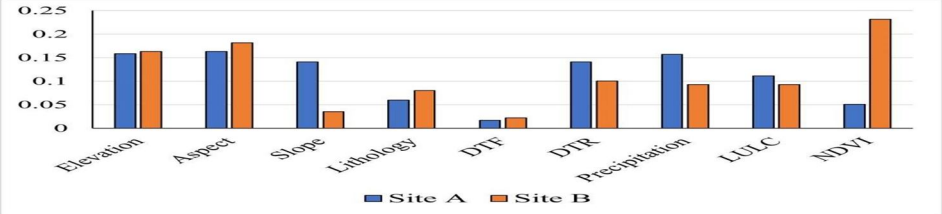


Fig 4

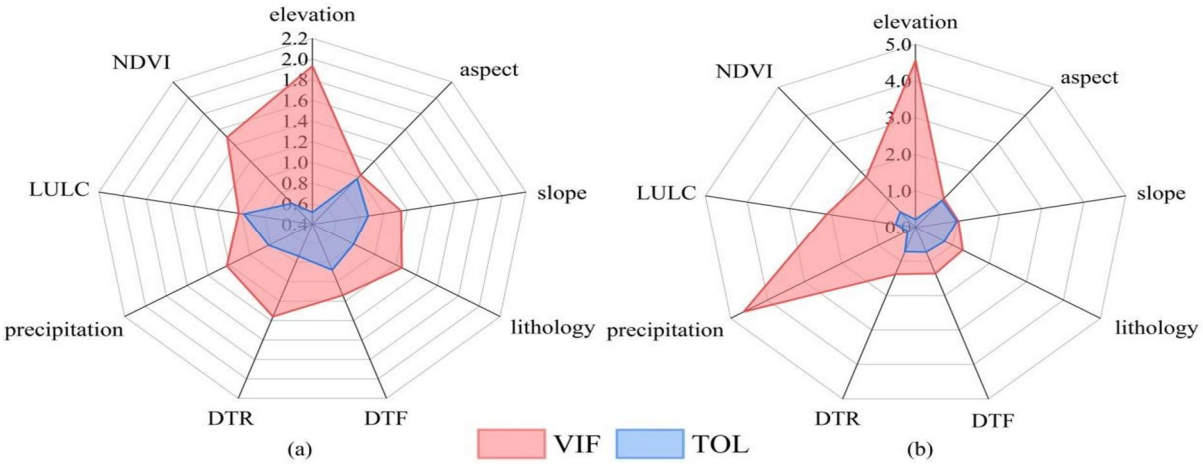


Fig 3

B. Comparison of LSM Results

Using six network architectures, we classified landslide susceptibility in Sites A and B into five categories: VL, L, M, H, and VH. Most areas were classified as VL and L, with H and VH zones closely aligning with historical landslides. The CTLGNet model showed the best alignment with historical landslides. In Site A, VL and L zones are mainly along the Yangtze River, while in Site B, they are in the western and central regions. In Site A, 87.21% of the area is low landslide-prone, and CTLGNet accurately predicted 97.71% of historical landslides. The VH zones accounted for 6.27% of the area but represented 88.95% of predicted susceptibility. LD values increased with susceptibility levels, with CTLGNet showing the highest LD values, indicating its superior predictive power over other models.

C. Comparative Evaluation of Model Performance

The dataset was split into training, validation, and test sets to evaluate the performance of six models. ROC curves and AUC values (Fig. 5) showed that CTLGNet outperformed all models, achieving the highest AUC values for both Site A (0.9817) and Site B (0.9693). Table 2 compares model metrics, with CTLGNet leading in most, including accuracy, precision, F1-score, MCC, and Kappa. In Site A, CTLGNet achieved an OA of 95.45%, precision of 93.55%, F1-score of 95.46%, and MCC of 0.910. In Site B, it achieved an OA of 93.48%, precision of 91.14%, and recall of 96.62%. These results confirm CTLGNet as the top-performing model for landslide susceptibility mapping.

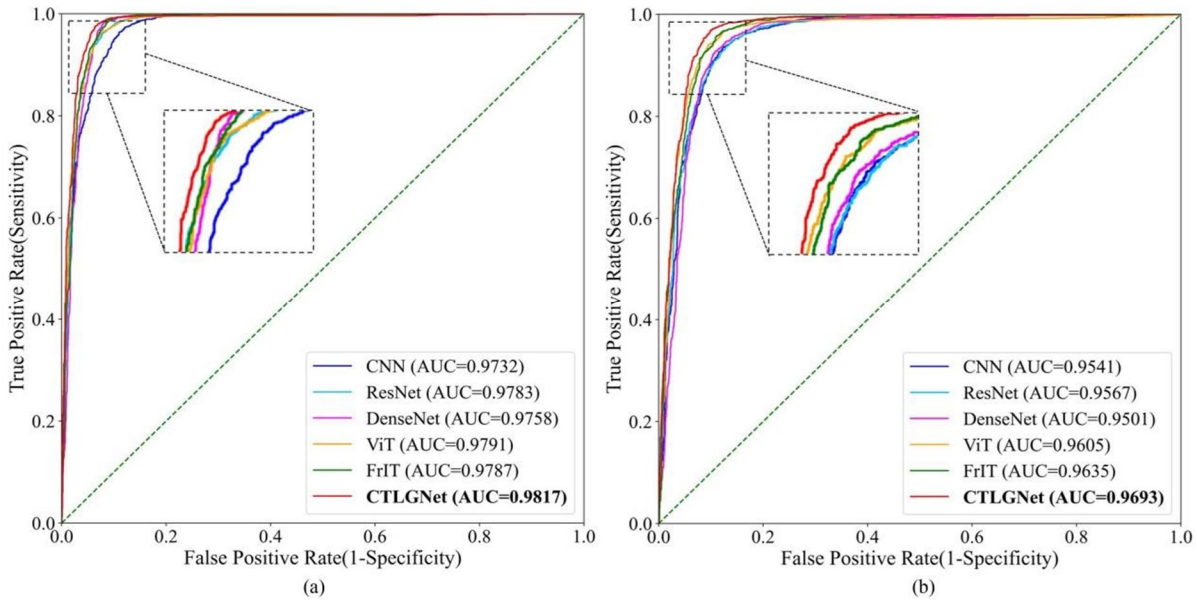


fig 5

MODEL EVALUATION METRICS OF EACH MODEL IN SITES A AND B

Site	Models	OA	Precision	Recall	F1-score	MCC	Kappa	AUC	RMSE	MAE
Site A	CNN	92.500%	88.956%	96.743%	92.686%	0.853	0.850	0.9732	0.274	0.075
	ResNet	93.800%	90.362%	97.812%	93.939%	0.879	0.876	0.9783	0.249	0.062
	DenseNet	92.625%	91.079%	**98.727%**	94.749%	0.896	0.893	0.9758	0.232	0.054
	ViT	94.050%	91.614%	96.743%	94.108%	0.882	0.881	0.9791	0.244	0.060
	FrIT	94.700%	91.481%	98.371%	94.801%	0.894	0.896	0.9787	0.230	0.053
	CTLGNet	**95.450%**	**93.552%**	97.456%	**95.464%**	**0.910**	**0.909**	**0.9817**	**0.213**	**0.045**
Site B	CNN	90.550%	88.213%	93.597%	90.825%	0.812	0.811	0.9541	0.307	0.095
	ResNet	90.075%	86.376%	95.147%	90.549%	0.806	0.802	0.9567	0.315	0.099
	DenseNet	90.850%	88.387%	94.047%	91.129%	0.819	0.817	0.9501	0.303	0.092
	ViT	92.050%	90.196%	94.347%	92.225%	0.842	0.841	0.9605	0.282	0.080
	FrIT	92.225%	88.858%	96.548%	92.543%	0.844	0.847	0.9635	0.278	0.078
	CTLGNet	**93.475%**	**91.136%**	**96.623%**	**93.799%**	**0.871**	**0.869**	**0.9693**	**0.255**	**0.065**

Table 2

V. CONCLUSION

In this article, we propose CTLGNet, a model that incorporates both LLFs and LGFs for landslide susceptibility mapping (LSM). It was applied in the Three Gorges Reservoir area and Jiuzhaigou, using historical landslide data and nine LCFs. CTLGNet's performance was evaluated against five models: CNN, ResNet, DenseNet, ViT, and FrIT.The results show that CTLGNet provides accurate LSM, with the VH and H susceptibility zones closely matching historical landslide locations. It outperforms other models in all evaluation metrics except Recall, achieving AUC values of 0.9817 and 0.9693 for the two regions. Additionally, CTLGNet produces the highest mean landslide susceptibility values and the lowest MAD and SD within historical landslide areas, indicating superior localization and detail extraction. It also has the lowest number of parameters and FLOPs among transformer-based models, making it more computationally efficient. In conclusion, CTLGNet demonstrates excellent predictive power and generalization, making it highly promising for a wide range of LSM applications.

VI. CHALLENGES AND FUTURE WORK

The accuracy and credibility of landslide susceptibility mapping (LSM) are heavily influenced by the quality of data sources. With advancements in remote sensing technology, multimodal remote sensing data can provide complementary information, but integrating such data to create a high-quality landslide dataset remains a challenge. In addition, training samples are essential for building LSM models. While this study used basic data augmentation techniques like flipping and rotation, more advanced methods, proven effective in image classification, could further improve LSM accuracy. Exploring the potential of these sophisticated techniques to enhance landslide feature detection is an important area for future research. Finally, the chosen model architecture significantly impacts LSM performance. Different deep learning models may vary in how they extract and process landslide features. While this study used a sequential CNN-transformer structure, future research could explore parallel or hierarchical architectures to improve LSM model performance.

REFERENCES

P. T. T. Ngo et al., "Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran," Geosci. Front., vol. 12, no. 2, pp. 505–519, Aug. 2020

[1] A. Merghadi et al., "Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance," Earth-Sci. Rev., vol. 207, Aug. 2020, Art. no. 103225.

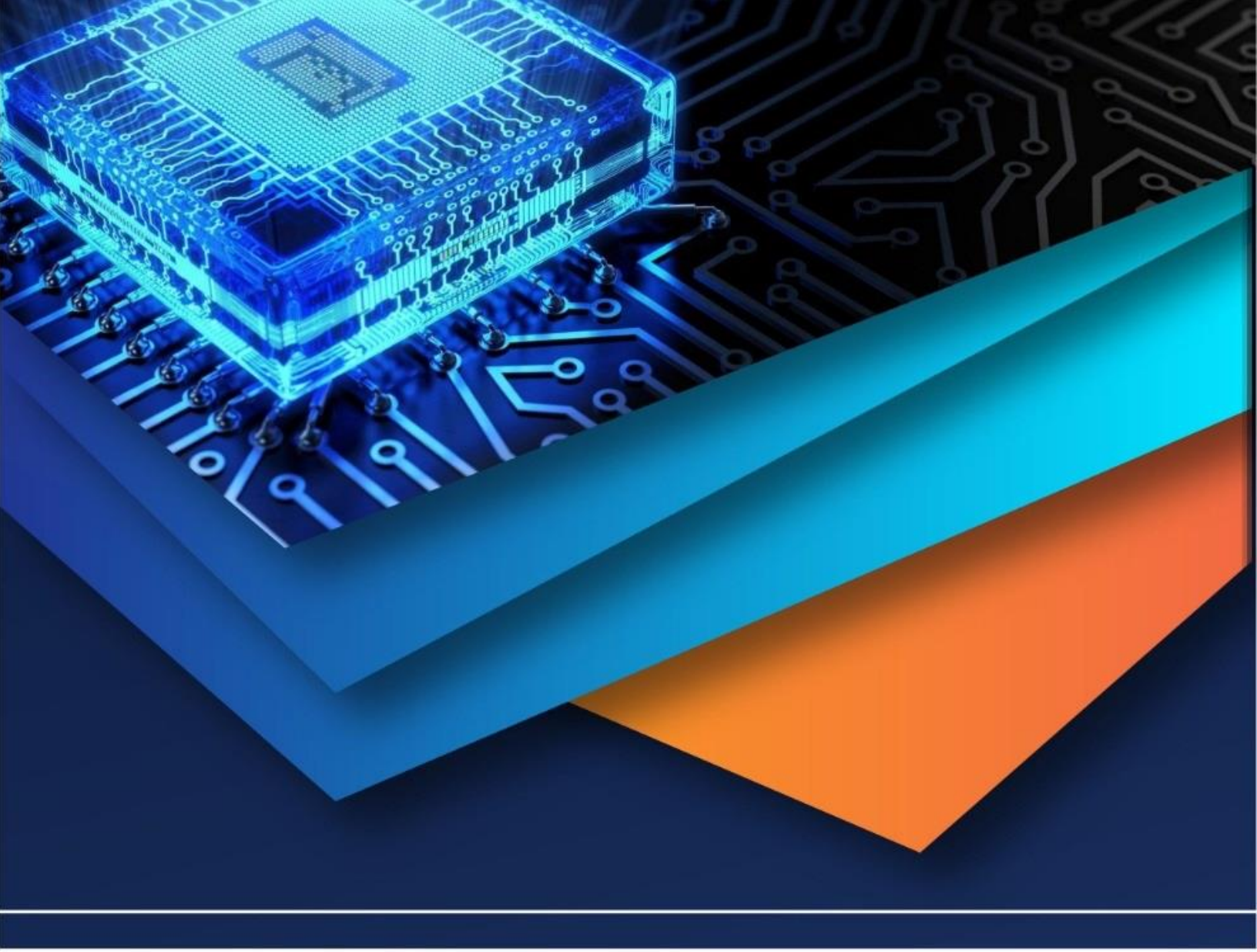
[2] R. Wei, C. Ye, T. Sui, Y. Ge, Y. Li, and J. Li, "Combining spatial response features and machine learning classifiers for landslide susceptibility mapping," Int. J. Appl. Earth Observ. Geoinformation, vol. 107, Mar. 2022, Art. no. 102681.

[3] Y. Achour and H. R. Pourghasemi, "How do machine learning techniques help in increasing accuracy of landslide susceptibility maps?" Geosci. Front., vol. 11, no. 3, pp. 871–883, May 2020.

[4] L. Zhang and L. Zhang, "Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities," IEEE Geosci. Remote Sens. Mag., vol. 10, no. 2, pp. 270–294, Jun. 2022.

[5] L. Zhang, L. Song, B. Du, and Y. Zhang, "Nonlocal low-rank tensor completion for visual data," IEEE Trans. Cybern., vol. 51, no. 2, pp. 673–685, Feb. 2021.

[6] W. Huang et al., "Landslide susceptibility mapping and dynamic response along the Sichuan-Tibet transportation corridor using deep learning algorithms," CATENA, vol. 222, Mar. 2023, Art. no. 106866.
[7] C. Chen and L. Fan, "An attribution deep learning interpretation model for landslide susceptibility mapping in the Three Gorges Reservoir area," IEEE Trans. Geosci. Remote Sens., vol. 61, Oct. 2023, Art. no. 3000515.
[8] T. Chen, Q. Wang, Z. Zhao, G. Liu, J. Dou, and A. Plaza, "LCFSTE: Landslide conditioning factors and swing transformer ensemble for landslide susceptibility assessment," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 17, pp. 6444–6454, Mar. 2024.
[9] L. Wang et al., "UNetFormer: A UNet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery," ISPRS J. Photogrammetry Remote Sens., vol. 190, pp. 196–214, Aug. 2022.
[10] L. Lv, T. Chen, J. Dou, and A. Plaza, "A hybrid ensemble-based deep learning framework for landslide susceptibility mapping," Int. J. Appl. Earth Observ. Geoinformation, vol. 108, Apr. 2022, Art. no. 102713.
[11] Y. Ge, G. Liu, H. Tang, B. Zhao, and C. Xiong, "Comparative analysis of five convolutional neural networks for landslide susceptibility assessment," Bull. Eng. Geol. Environ., vol. 82, no. 10, Sep. 2023, Art. no. 337.
[12] A. M. Youssef, B. Pradhan, A. Dikshit, M. M. Al-Katheri, S. S. Matar, and A. M. Mahdi, "Landslide susceptibility mapping using CNN-1D and 2D deep learning algorithms: Comparison of their performance at AsirRegion, KSA," Bull. Eng. Geol. Environ., vol. 81, no. 4, Mar. 2022, Art. no. 165.
[13] X. Liu, Y. Wu, W. Liang, Y. Cao, and M. Li, "High-resolution SAR image classification using global-local network structure based on vision transformer and CNN," IEEE Geosci. Remote Sens. Lett., vol. 19, Feb. 2022, Art. no. 4505405.
[14] D. Wang et al., "Evaluation of deep learning algorithms for landslide susceptibility mapping in an alpinegorge area: A case study in Jiuzhaigou County," J. Mountain Sci., vol. 20, no. 2, pp. 484–500, Feb. 2023.
[15] A. Jamali, S. K. Roy, and P. Ghamisi, "WetMapFormer: A unified deep CNN and vision transformer for complex wetland mapping," Int. J. Appl. Earth Observ. Geoinformation, vol. 120, Jun. 2023, Art. no. 103333.
[16] A.A.Aleissaeeetal.,"Transformers in remote sensing: A survey," Remote Sens., vol. 15, no. 7, pp. 1860– 1860, Mar. 2023.
[17] J. Ma, M. Li, X. Tang, X. Zhang, F. Liu, and L. Jiao, "Homo–heterogenous transformer learning framework for RS scene classification," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 15, pp. 2223– 2239, Mar. 2022.
[18] Y. Feng, H. Xu, J. Jiang, H. Liu, and J.Zheng, "ICIF-Net:Intra-scalecrossinteraction and inter-scale feature fusion network for bi-temporal remote sensing image change detection," IEEE Trans. Geosci. Remote Sens., vol. 60, Apr. 2022, Art. no. 4410213.
[19] X. Gao, T. Chen, R. Niu, and A. Plaza, "Recognition and mapping of landslide using a fully convolutional DenseNet and influencing factors," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 7881–7894, Aug. 2021.
[20] T. Chen et al., "BisDeNet: A new lightweight deep learning-based framework for efficient landslide detection," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 17, pp. 3648–3663, Jan. 2024.
[21] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, "Vision transformers for remote sensing image classification," Remote Sens., vol. 13, no. 3, p. 516, Feb. 2021.
[22] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," ACM Comput. Surv., vol. 54, no. 10, pp. 1–41, Jan. 2022.
[23] X. Zhao et al., "Fractional Fourier image transformer for multimodal remote sensing data classification," trans.NeuralNetw.Learn.Syst., vol. 35, no. 2, pp. 2314–2326, Feb. 2024.
[24] B. Gao et al., "Landslide risk evaluation in Shenzhen based on stacking ensemble learning and InSAR," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 16, pp. 1–18, Jul. 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY