Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-2222

Digital assessment and prediction of rainfall-induced landslides using convolutional neural networks (CNN)

Sangseom Jeong, Sangmin Lee
Department of Civil Engineering, Yonsei University, Republic of Korea, soj9081@yonsei.ac.kr; inyu1225@gmail.com

Yongmin Kim, Yuan Shen Chua
University of Glasgow, Singapore, yongmin.kim@glasgow.ac.uk; 2942017C@student.gla.ac.uk

Homin Park
Singapore University of Social Sciences, Singapore, hominpark@suss.edu.sg

ABSTRACT: This study employs a U-Net architecture, known for its effectiveness in image segmentation and spatial feature extraction. By treating input data as spatially rich features, the U-Net model can efficiently predict instability of slope with factors of safety (FS), enabling nearly instantaneous evaluations without the computational burden associated with conventional techniques. A comprehensive set of simulated landslide scenarios was assessed by regional-scaled analyses to ensure that the model encountered a variety of slope geometries, soil properties, and triggering factors such as rainfall. The results demonstrated that the proposed U-Net-based solution achieves high accuracy in FS predictions, matching or exceeding the reliability of traditional methods while significantly reducing computation time. This innovative approach has the potential to transform landslide risk management by delivering real-time FS assessments, thereby enabling more proactive and adequate decision-making in slope stability monitoring and disaster prevention efforts.

KEYWORDS: Deep learning, Landslide, CNN (Convolutional Neural Network), U-Net

1 INTRODUCTION

The factor of safety (FS) is a critical metric for evaluating slope stability, with accurate predictions becoming increasingly essential due to escalating threats to human lives, infrastructure, and the environment (Kim et al., 2022). Timely and precise estimation of FS facilitates the early identification of vulnerable regions, allowing for the implementation of effective mitigation strategies. Traditional methods for assessing FS, including deterministic and probabilistic approaches, are often computationally demanding and unsuitable for real-time applications, particularly in areas susceptible to rapid landslide events. To overcome these challenges, this study proposes an innovative convolutional neural network-based methodology for the automatic and rapid identification of rainfall-induced landslides characterized by FS values approaching unity. The proposed model employs a U-Net architecture, known for its robustness in image segmentation tasks and effective spatial feature extraction capabilities (Azarafza et al., 2021; Collini et al., 2022; Du et al., 2020). This architecture processes input data as spatially distributed features, significantly enhancing the efficiency of FS predictions while minimizing computational burdens.

In this study, extensive numerical simulations were conducted to generate diverse datasets representing various slope geometries, soil properties, and rainfall intensities. The developed CNN-based approach enables swift analysis of large datasets, facilitating the rapid and reliable detection of shallow rainfall-triggered landslides within the targeted region. Furthermore, the versatility of the developed model indicates its potential applicability to other geographic areas exhibiting similar landslide susceptibility conditions.

2 METHODOLOGY

2.1 *Conceptual Framework*

To develop an accurate landslide prediction model, logically structured datasets derived from physics-based slope stability analyses were utilized as training data due to the lack of available natural or satellite imagery data. These datasets were generated by systematically varying key geotechnical parameters, including cohesion (c), internal friction angle (ϕ), hydraulic conductivity (k), and rainfall intensity, to evaluate their combined effects on slope instability. Rasterized digital elevation models (DEMs) and geospatial maps were employed to compute grid-based FS. The computed FS values were subsequently structured into multi-dimensional tensors, preserving the physical integrity of the data and ensuring suitability for training deep learning models.

The proposed deep learning approach employed semantic segmentation, facilitating pixel-level classification directly aligned with the grid-based FS values derived from numerical analyses (Kim et al., 2014). This technique enables the model to detect localized slope instability by learning spatial features at the cell level. Similar to the RGB channels used in conventional image processing, individual geotechnical parameters were assigned separate input channels, enhancing the model's ability to interpret complex spatial-geotechnical interactions effectively.

2.2 *U-Net model*

Semantic segmentation is a computer vision technique that identifies the exact shape and location of individual objects within an image at a pixel level. Compared to image classification and object detection, semantic segmentation delivers higher spatial resolution, making it particularly suitable for geographic data analysis tasks such as landslide prediction. The U-Net is a prominent architecture used in semantic segmentation.

The U-Net architecture features a symmetric structure consisting of an encoder and a decoder, as illustrated in Figure 1. The encoder progressively reduces the spatial dimensions of the input, creating condensed feature maps. The decoder then expands these feature maps back to the original resolution, enabling precise pixel-level predictions. A key characteristic of the U-Net architecture is its utilization of skip connections, which directly connect high-resolution encoder feature maps to

corresponding stages in the decoder. These skip connections effectively preserve spatial and boundary information that is typically lost during downsampling, thus improving object boundary clarity (Ronneberger et al., 2015).

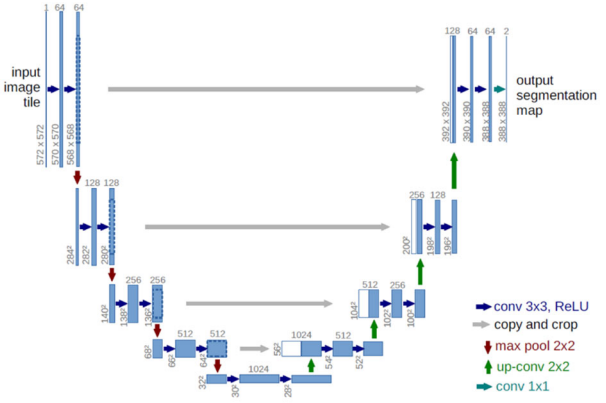


Figure 1. Architecture of U-Net mode (Ronneberger et al., 2015)

The U-Net-based prediction model follows a systematic four-stage workflow: data generation, preprocessing, training, and prediction. Initially, comprehensive parametric simulations were conducted using LANDFLOW 3.0 (LANDFLOW Program version 3) to generate a diverse range of landslide scenarios. Next, the datasets underwent preprocessing involving min-max normalization and zero-padding to ensure spatial uniformity, followed by restructuring into multi-channel input tensors. Additionally, data augmentation techniques were selectively applied to enhance the model's generalization capabilities. The third stage involved training the U-Net model using preprocessed multi-channel inputs and binary landslide masks, classifying each grid cell based on a FS threshold of unity. Hyperparameters such as learning rate and dropout ratio were systematically optimized to achieve robust model training. In the final stage, the trained model's performance was evaluated through standard metrics including Intersection over Union (IoU), Dice Score, Precision, and Recall. Predicted results were visually compared with original FS maps to validate the spatial accuracy of model predictions. The detailed workflow of the U-Net model development process is depicted in Figure 2.

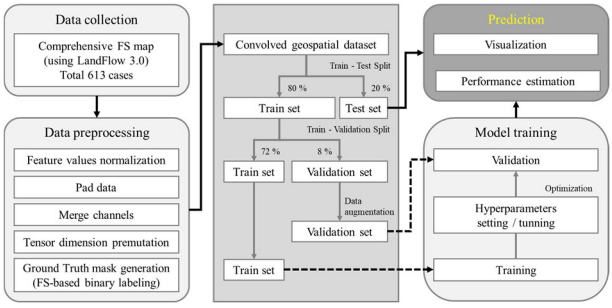


Figure 2. Flow chart of the U-Net prediction model

2.3 *Training data generation*

Mt. Ansan, selected as the study area, provided high-resolution topographic and soil depth data representative of typical mountainous terrain (Jeong et al., 2022), as illustrated in Figure 3. FS maps were generated using LANDFLOW 3.0, incorporating diverse geotechnical parameters specific to Mt. Ansan, Seoul. The analysis integrated terrain data, index soil properties, shear strength parameters, hydraulic parameters, and rainfall conditions. The resulting FS maps, visually depicted in Figure 4, were outputted in a grid-based text format, detailing stability conditions for individual cells.

To simplify the analysis, rainfall intensity was maintained constant for a duration of 24 hours. Key parameters, including cohesion, internal friction angle, and hydraulic conductivity, were defined based on field investigation data from mountainous regions in Seoul, while other parameters were kept constant to minimize additional variability (Seoul Metropolitan Government, 2013).

The number of failure cells was recorded as an indirect indicator of landslide susceptibility. The Mt. Ansan grid consisted of 436 × 489 cells, totaling 213,204 cells, with 85,909 classified as mountainous. Initially, 388 cases were prepared in the first-stage dataset, incrementally ranging from scenarios with minimal or no landslide occurrences (one to ten failed cells). This dataset primarily focused on assessing the impacts of cohesion (c) and internal friction angle (ϕ) under constant rainfall intensities of 4.4 mm/hr and 5.2 mm/hr. The second-stage dataset, comprising an additional 225 cases, incorporated varied rainfall intensities to comprehensively evaluate the effects of hydraulic conductivity. Overall, a total of 613 cases were analyzed.

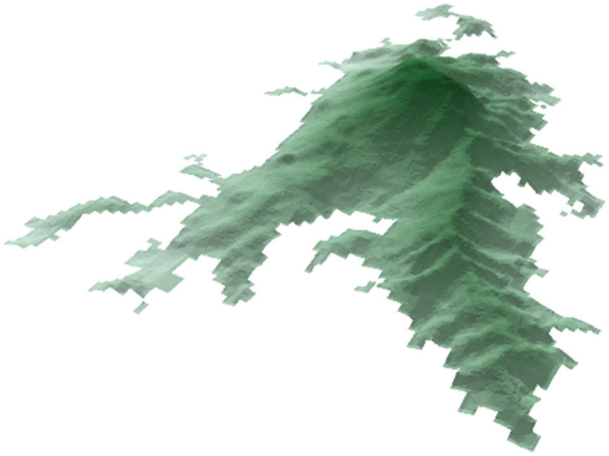


Figure 3. 3D visualized DEM of the study area (Mt. Ansan in Korea)

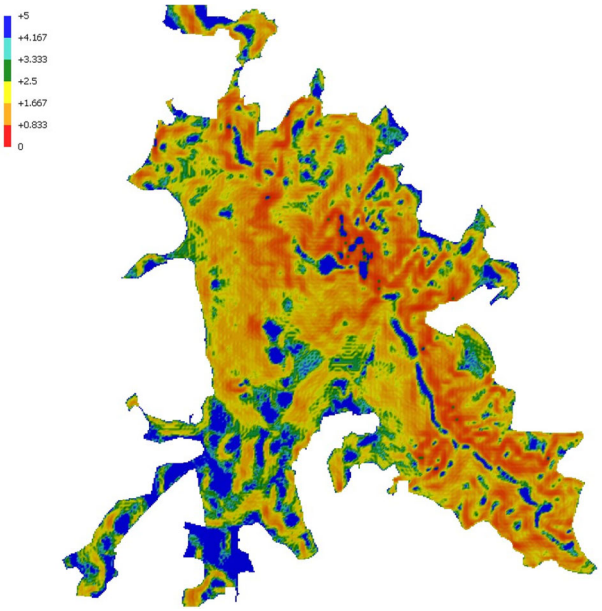


Figure 4. FS map calculated by LANDFLOW 3.0

2.4 *Data preprocessing*

2.4.1 *Normalization*

Variations in the range of input features during deep learning model training can lead to slow convergence, reduced

performance, or excessive bias toward certain features. To address these issues and enhance training stability, normalization is essential. Several normalization methods exist, including Z-score normalization, min-max normalization, logarithmic transformation, and robust normalization. In this study, min-max normalization was specifically selected due to the clearly defined physical ranges and finite, bounded values of input parameters such as terrain, soil characteristics, and rainfall data. Min-max normalization scales data linearly to a range between 0 and 1, facilitating optimal interaction with activation functions, accelerating convergence, and enhancing predictive accuracy.

Rainfall intensity normalization employed the 24-hour, 50-year return period rainfall intensity derived from Seoul's Intensity-Duration-Frequency (IDF) curve as proposed by Heo et al. (2000). Elevation and soil depth normalization leveraged observed maximum and minimum values from the Mt. Ansan Digital Elevation Model (DEM) and soil depth datasets, ensuring accurate and physically meaningful scaling. Consequently, consistent normalization of all input channels ensured balanced contribution of each feature during model training, enhancing overall stability and reliability, as shown in Equations 1 and 2.

$$I(t,T) = \frac{366.6 + 134.1 \ln \frac{T}{t^{0.02385}}}{1.254 + 0.05396 \ln \frac{\sqrt{T}}{t} + \sqrt{t}} \quad (t \le 360\ min) \tag{1}$$

$$I(t,T) = \frac{190.1 + 84.88 \ln \frac{T}{t^{-0.01775}}}{-2.228 + 0.5258 \ln \frac{\sqrt{T}}{t} + \sqrt{t}} \quad (t > 360\ min) \tag{2}$$

2.4.2 *Data padding*

Maintaining consistent spatial input dimensions is crucial for deep learning-based landslide prediction models, especially when employing convolutional neural networks (CNNs). CNNs inherently require uniform spatial dimensions to facilitate effective convolution operations. Therefore, padding was applied to standardize all input data to a uniform dimension of 489×489 pixels. Padding involves adding artificial values (typically zeros) around original data, which serves multiple critical functions: it prevents information loss at the boundaries during convolution operations, ensures dimensional consistency across diverse input datasets, and enables seamless integration with data augmentation processes.

Considering that data augmentation methods such as rotation, flipping, and cropping were applied, consistent input sizes post-augmentation became a necessity. Padding regions were explicitly set to zero, thereby preventing misinterpretation by the CNN model and maintaining numerical integrity. This practice preserved the spatial consistency and boundary information and mitigated distortions introduced by data augmentation, securing enhanced training stability and efficiency.

2.4.3 *Channel merging*

Following normalization and padding, each feature existed initially as an independent data channel: elevation, soil depth, cohesion (c), internal friction angle (ϕ), hydraulic conductivity (k), and rainfall intensity. These individual feature channels were subsequently merged into unified multi-channel tensors to leverage the CNN's capability for simultaneous multi-feature processing. The channel merging process was conceptually similar to RGB image processing, where individual channels (Red, Green, Blue) collectively represent a single image. However, in contrast to RGB channels, each channel in this study explicitly represented geotechnical and hydrological parameters with distinct physical interpretations. This multi-channel structure facilitated the CNN model's learning of both individual feature characteristics and their complex spatial interdependencies, significantly enhancing the model's predictive performance.

Additionally, tensor dimensions were carefully rearranged to meet CNN input format requirements, structured as (channels, height, width). This permutation ensured compatibility with established deep learning frameworks, further contributing to training robustness.

2.4.4 *Ground truth mask generation*

Accurate binary ground truth masks were generated from the FS results computed by LANDFLOW 3.0 for each simulated scenario, serving as essential target data for CNN training, validation, and evaluation. Cells with FS values below 1.0 were labeled as landslide occurrences (assigned a value of 1), while cells with FS values equal to or exceeding 1.0 were labeled as stable (assigned a value of 0). This classification scheme aligns with widely accepted standards in slope stability analysis.

The binary masks facilitated clear and precise pixel-level classification, aligning perfectly with semantic segmentation tasks inherent to CNN-based methodologies. Moreover, since standard performance metrics such as Intersection over Union (IoU), Dice Score, Precision, and Recall inherently rely on binary classifications, this approach ensured robust and meaningful model performance assessments. The binary masks were consistently applied throughout all phases of model training and evaluation, maintaining methodological consistency and accuracy.

2.4.5 *Data augmentation*

Data augmentation techniques are pivotal for enhancing model generalization and effectively learning diverse data distributions, particularly when dealing with limited datasets. In this study, conservative and physically meaningful augmentation strategies were selectively applied to ensure numerical consistency and preserve the interpretability of geotechnical parameters. Specifically, horizontal and vertical flips were implemented to alter spatial arrangements without impacting numerical integrity. Rotations were restricted to discrete angles of 90°, 180°, and 270°, effectively diversifying spatial orientations without distorting underlying numerical grid relationships. In contrast, excessive manipulations such as random scaling, distortion, or noise addition were deliberately avoided, as these could compromise the physical validity of the data, as shown in Figure 5.

Through carefully chosen augmentation methods, the original dataset size was effectively doubled. This substantial increase in dataset diversity significantly improved the training robustness and generalization capabilities of the CNN model, allowing it to adapt effectively to various real-world conditions.

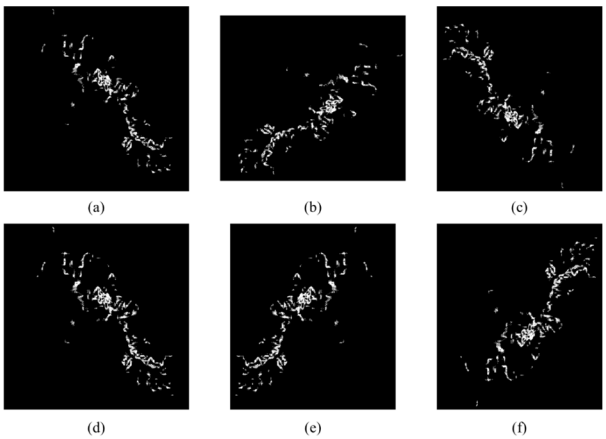


Figure 5. Augmented binary mask (a) original image 1; (b) rotation 90deg; (c) rotation 180deg; (d) original image 2; (e) horizontal flip; (f) vertical flip

2.5 *Model training*

For optimization during training, the Adam optimizer was selected due to its proven capability in rapidly adjusting network weights by considering both the mean and variance of past gradient computations. Adam's adaptive moment estimation ensures efficient navigation through the complex optimization landscape, thereby offering rapid convergence and robust training stability.

The loss function implemented in this study was a combination of Binary Cross-Entropy (BCE) loss and Dice loss. BCE loss quantifies the probabilistic differences between the predicted and actual binary ground truth values on a pixel-wise basis, making it particularly suitable for binary classification problems such as landslide detection. Dice loss, on the other hand, directly optimizes spatial overlap between predicted segmentation masks and ground truth masks, effectively addressing challenges posed by imbalanced class distributions in pixel-level segmentation tasks. The integration of these two loss functions thus enabled simultaneous consideration of pixel-level probabilistic accuracy and spatial agreement, significantly enhancing the model's predictive reliability and accuracy.

Model training was executed within a GPU-accelerated computational environment, employing mini-batch learning techniques to efficiently manage memory utilization and computational speed. Throughout the training process, performance metrics were continuously monitored and evaluated using a dedicated validation dataset, ensuring the model's robust performance and generalization capability.

2.6 *Hyperparameter optimization*

Hyperparameter tuning was conducted to enhance the predictive performance and stability of the CNN-based U-Net model. In deep learning, hyperparameters play a critical role as they directly influence the convergence rate, generalization capability, and the model's resistance to overfitting. This study primarily focused on tuning two key hyperparameters: the learning rate and the dropout ratio.

The learning rate, which governs the step size during the optimization process by determining how significantly model weights are updated in each iteration, was adjusted experimentally. A reduced learning rate was tested to achieve more stable and precise convergence, while an increased learning rate was explored to accelerate convergence during the early training stages. These adjustments aimed to strike a balance between convergence speed and training stability.

In parallel, dropout—a regularization strategy that randomly deactivates a fraction of neurons during training—was applied with a probability of 0.2. This widely adopted setting in deep learning literature mitigates the risk of overfitting by limiting the model's reliance on specific neurons and promoting redundancy in feature learning.

Experiments were conducted under three configurations: (1) learning rate adjustment alone, (2) dropout application alone, and (3) combined adjustment of learning rate and dropout. These configurations enabled a comprehensive assessment of their individual and synergistic effects on model performance.

2.7 *Evaluation metrics*

Model performance was quantitatively assessed using primary metrics such as Intersection over Union (IoU), Dice Score, and Accuracy, along with reference metrics such as Precision and Recall. These metrics mathematically quantify the alignment or relationship between the predicted outcomes and the ground truth, crucial for evaluating the practical predictive effectiveness of the model, particularly in class-imbalanced scenarios such as landslide occurrences.

The Intersection over Union (IoU) metric is defined as the ratio of the intersection area between predicted regions and actual regions to their union area. It provides a pixel-level measure of the overlap between predictions and ground truth, with values closer to 1 indicating highly precise predictions. IoU is formally calculated as presented in Equation (3):

$$IoU = \frac{TP}{TP + FP + FN} \tag{3}$$

where TP (True Positive): Correct prediction of landslide occurrence, FP (False Positive): Incorrect prediction indicating landslide occurrence where none actually occurred, FN (False Negative): Missed prediction of an actual landslide occurrence.

The Dice Score, also referred to as the F1 score, assesses the similarity between two sets and is notably effective for evaluating models on imbalanced datasets by emphasizing minority class predictions more strongly than IoU. The Dice Score is calculated as shown in Equation (4):

$$Dice\ Score = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

Accuracy represents the proportion of correctly predicted cells (comprising both True Positive (TP) and True Negative (TN) cases) relative to the total number of cells. However, accuracy alone can misleadingly indicate high performance in landslide prediction models due to the dominance of True Negative instances. Hence, accuracy was utilized primarily as a supplementary metric. The accuracy metric is computed according to Equation (5):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Precision refers to the proportion of predicted landslide regions that correctly overlap with the actual landslide occurrences. It decreases when the number of False Positives (FP) increases. Precision is calculated as shown in Equation (6):

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

Recall represents the proportion of actual landslide occurrences that are correctly identified by the model. It decreases when the number of False Negatives (FN) increases. Recall is computed according to Equation (7):

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

These metrics together comprehensively evaluate model performance, ensuring a robust and balanced assessment of predictive capability across various operational scenarios.

3 RESULT AND DISCUSSION

In this study, the model's training performance was primarily evaluated using the Validation Dice Score. To ensure consistent comparisons across various experimental setups, the batch size was uniformly set at 8, the learning rate fixed at 1e-4, and the number of training epochs maintained at 75. The training procedure utilized a two-stage dataset configuration, enabling detailed analysis of performance variations related to different data compositions, feature selections, and the impact of data augmentation.

The first-stage dataset predominantly comprised scenarios exploring threshold conditions of landslide occurrence under relatively low rainfall intensities, with controlled variations in cohesion and internal friction angle. The second-stage dataset significantly enhanced training complexity by integrating additional variability in rainfall intensity and hydraulic conductivity (k).

Input features were methodically structured to incrementally increase complexity: initially incorporating basic terrain and environmental characteristics (elevation, soil depth, rainfall intensity), and subsequently adding key geotechnical parameters—cohesion (c), internal friction angle (ϕ), and hydraulic conductivity (k). Additionally, experiments investigated the impact of applying data augmentation techniques. Performance metrics including Validation Dice Score, Train Dice Score, and Train Loss were meticulously documented to quantitatively assess the learning efficiency and efficacy of different experimental configurations.

After hyperparameter optimization, the U-Net model underwent rigorous evaluation using an independent test dataset. Test results revealed a mean Intersection over Union (mIoU) of 0.31, indicating moderate overall predictive capability. However, the average accuracy, computed from the confusion matrices, reached 0.989, demonstrating extremely high overall classification correctness. Detailed performance distributions and statistics related to failure cell counts are illustrated in Figure 6.

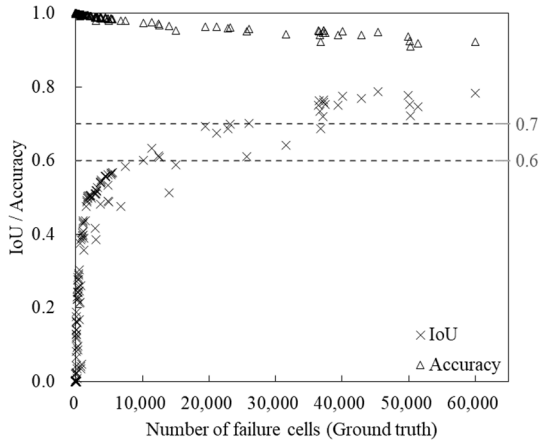


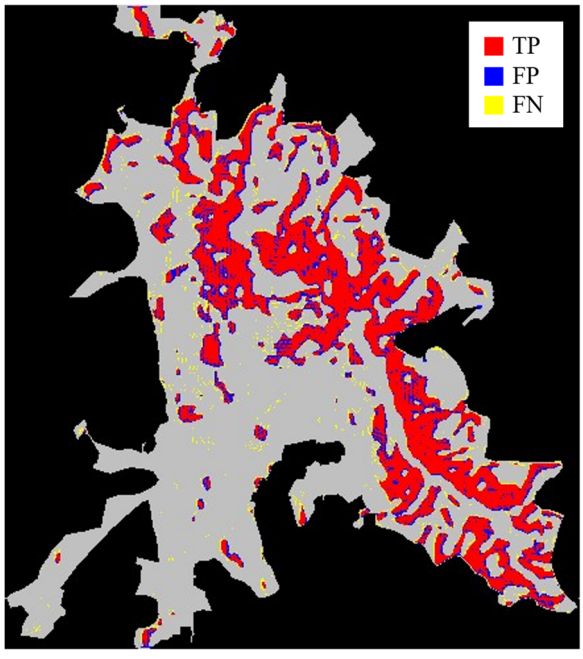
Figure 6. Distribution of test set data and prediction accuracy

Analysis of a randomly selected test dataset's predictions revealed distinct performance patterns correlated with the number of failure cells. Specifically, cases with failure cells exceeding 20,000 generally achieved IoU scores of 0.7 or higher, indicating a strong predictive capability. Conversely, cases featuring fewer than 7,500 failure cells typically yielded IoU scores below 0.6, with cases having fewer than 3,000 failure cells consistently demonstrating significantly poorer performance. This observation indicates that the Dice Score, employed as the primary training metric, experienced marked deterioration as the number of failure cells decreased.

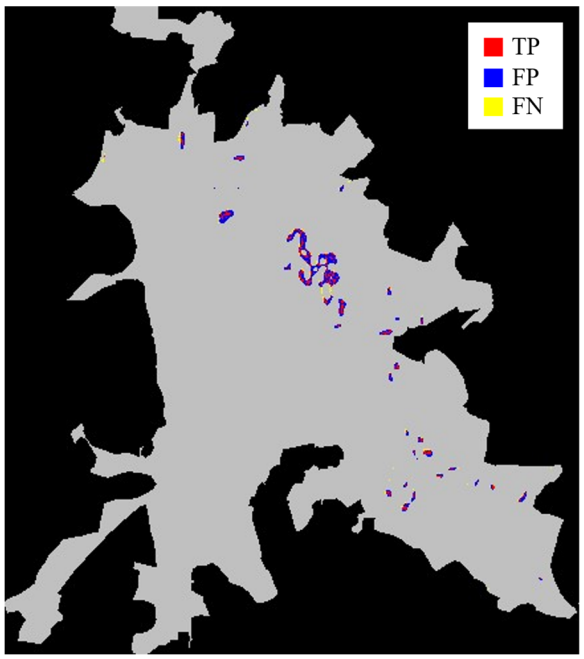
This observed phenomenon is fundamentally associated with the inherent mathematical sensitivity of binary classification metrics such as IoU and Dice Score to True Positive (TP) counts. When TP cell counts are minimal, the denominators of the metrics become disproportionately small, resulting in sharply reduced scores. Consequently, even minor prediction inaccuracies can disproportionately degrade performance metrics, potentially causing an underestimation of model performance in scenarios with limited landslide occurrences.

In contrast, overall accuracy—which measures the proportion of correctly predicted pixels (both True Positive and True Negative) to the total pixel count—remained exceptionally high (consistently exceeding 0.95) across all scenarios. This reflects the dominance of True Negative pixels within the dataset, where most regions remained stable (non-landslide areas). However, given the context of landslide prediction, reliance solely on accuracy metrics can be misleading. It is therefore critical to prioritize spatially sensitive metrics such as IoU and Dice Score, which provide a more nuanced assessment of prediction quality in landslide-prone regions.

Ultimately, the CNN model exhibited high predictive reliability in scenarios involving extensive landslide occurrences characterized by large failure areas but demonstrated notable performance deterioration for scenarios involving limited landslide extents. Representative prediction results are presented in Figure 7.



(a) 8.5mm/hr×24hr, 14kPa, 26.6°, 1.3e-6m/s



(b) 4.7mm/hr×24hr, 15kPa, 23°, 4.4e-6m/s

Figure 7. Overlayed image based on prediction results of test set

Figure 7(a) illustrates a scenario with a rainfall intensity of 8.5 mm/hr, a cohesion of 14 kPa, an internal friction angle of 26.6°, and a hydraulic conductivity of 1.3e-6 m/s, achieving an IoU of around 0.7. This case also resulted in a Dice Score of 0.82, an Accuracy of 0.96, a Precision of 0.81, and a Recall of 0.83, indicating a well-balanced predictive performance. The model predictions successfully delineated the majority of actual failure regions (True Positives), though minor False Negative distributions were observed adjacent to correctly predicted areas.

Figure 7(b) illustrates a contrasting scenario with a lower rainfall intensity of 4.7 mm/hr, cohesion of 15 kPa, an internal friction angle of 23°, and hydraulic conductivity of 4.4e-6 m/s, resulting in an IoU of approximately 0.3. The corresponding Dice Score was 0.45, with an Accuracy of 0.99, a Precision of 0.39, and a Recall of 0.54, reflecting a tendency toward overprediction and reduced sensitivity to actual failure zones. In this case, the limited failure area posed significant predictive challenges, leading to numerous False Positives, which highlights the difficulties inherent in accurately predicting small-scale landslide occurrences.

4 CONCLUSIONS

The primary objective of this study was to develop a Convolutional Neural Network (CNN)-based prediction framework to evaluate landslide susceptibility using geotechnical parameters. The methodology uniquely leveraged geotechnical hazard maps, specifically FS maps, as training data inputs, thereby bridging the gap between geotechnical engineering and deep learning approaches in the field of computer vision.

A robust CNN-based landslide susceptibility prediction model was successfully developed utilizing the U-Net architecture, widely recognized for its efficacy in semantic segmentation tasks. The model demonstrated strong predictive capability, accurately identifying irregular and spatially dispersed landslide characteristics at the pixel level. This highlights the feasibility and effectiveness of integrating direct geotechnical inputs into CNN frameworks.

The study rigorously assessed various conditions during model training, including sequential addition of geotechnical parameters, the application of data augmentation techniques, and strategic dataset composition. Each condition distinctly contributed to enhancements in model learning performance and overall predictive reliability.

Finally, the prediction outcomes underscored the model's strength in accurately forecasting landslide occurrences when failure areas exceeded specific thresholds, notably around 5% to 10% of the slope area. Conversely, performance notably decreased for scenarios involving minimal and sparse landslide occurrences, as reflected in lower quantitative metrics, such as the Dice score and Intersection over Union (IoU). Consequently, future advancements should prioritize improving model prediction capabilities for smaller-scale landslide events, which is essential for refining the overall effectiveness of CNN-based landslide susceptibility frameworks.

5 ACKNOWLEDGEMENT

This work was supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2018-NR031077).

6 REFERENCES

Kim, J., Lee, K., Jeong, S., and Kim, G. 2014. GIS-based prediction method of landslide susceptibility using a rainfall infiltration-groundwater flow model. *Engineering geology*, 182, 63-78.

Kim, Y., Rahardjo, H., Nistor, M.M. Satyanaga, A., Leong, E.C., Sham, A.W.L. 2022. Assessment of critical rainfall scenarios for slope stability analyses based on historical rainfall records in Singapore. *Environmental Earth Sciences*, 81, 39.

Du, G, Cao, Cao, X, Liang, J, Chen, X, Zhan, Y, 2020. Medical Image Segmentation based on U-Net: *A Review in Journal of Imaging Science and Technology*, 020508-1 - 020508-12

Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., and Derakhshani, R. 2021. Deep learning-based landslide susceptibility mapping. *Scientific reports*, 11(1), 24112.

Collini, E., Palesi, L. I., Nesi, P., Pantaleo, G., Nocentini, N., and Rosi, A. 2022. Predicting and understanding landslide events with explainable AI. *IEEE Access*, 10, 31175-31189.

Ronneberger, O., Fischer, P., and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer international publishing.

Jeong, S., Hong, M., Song, J. 2022. Soil depth estimation in mountainous areas by using GIS and satellite images. *Landslides*, 19(11), 2711-2726

LANDFLOW Program version 3 (2019), LANDFLOW - A Numerical Modeling of GIS-based Rainfall-induced Landslides (Initiation) and Debris Flows, SW Management number: DG-2019-0088.

Heo, J.H., Lee, J.T., and Kim, G.D. 2000. Rainfall Intensity-Duration-Frequency of Seoul. *Journal of Korea Water Resources Association*, 33(S1), 164-169.

Seoul Metropolitan Government. (2013*). Landslide investigation report for mountainous areas in Seoul*. Seoul Institute of Technology.