Bhattarai, A., Shiwakoti, R.K. and Uduwage, D.N.L.S., 2026. Deep learning-based flood mapping in monsoon climates using sentinel-1 sar. In: Waidyasekara, K.G.A.S., Jayasena, H.S., Chandanie, H. and Tennakoon, G.A. (eds). *Proceedings of the 14th World Construction Symposium*, 10-11 July 2026, Sri Lanka. pp. 527-538. DOI: https://doi.org/10.31705/WCS.2026.39. Available from: https://ciobwcs.com/papers/

DEEP LEARNING-BASED FLOOD MAPPING IN MONSOON CLIMATES USING SENTINEL-1 SAR

Amit Bhattarai[1], Ranju Kumari Siwakoti[2] and D.N.L.S. Uduwage[3]

ABSTRACT

Monsoon floods pose significant disaster management challenges due to persistent cloud cover, which renders optical remote sensing ineffective for rapid response. To overcome this limitation, this study proposes a computationally efficient, all-weather flood mapping framework that integrates Sentinel-1 Synthetic Aperture Radar (SAR) data with Height Above Nearest Drainage (HAND) topographic constraints within a deep learning pipeline. Three model architectures were compared systematically: a baseline U-Net, a U-Net with spatial augmentation (UNet_Aug), and a U-Net++ with Focal-Tversky loss (UNetPP_TMA). Bi-temporal Sentinel-1 scenes (VV and VH polarisations) from the 2024 Nawalpur flood event in Nepal were combined with a hydrologically corrected 30 m SRTM DEM, utilizing Sentinel-2 imagery and field data as high-fidelity ground truth. Models were trained using patch-based sampling to ensure computational efficiency. Quantitative results demonstrated that while UNet_Aug achieved the highest overall balance of metrics (IoU of 0.6101, F1-Score of 0.7578, Precision of 0.7496, and AUC-ROC of 0.9901), UNetPP_TMA achieved the highest recall of 0.7704. Overall, UNet_Aug is recommended as the most suitable operational model because it provided the best balance across IoU, F1-score, Kappa, accuracy, and AUC, while UNetPP_TMA remains useful as a high-recall option for safety-critical screening. This study contributes a reliable flood delineation approach for cloud-obscured monsoon conditions and provides an open-source pipeline and dataset that can serve as a reproducible benchmark for future disaster management in data-scarce regions.

Keywords: *Change Detection, Deep Learning, Flooding Mapping; Sentinel-1, Synthetic Aperture Radar (SAR).*

1. INTRODUCTION

Flooding is one of the most destructive natural hazards globally, causing severe financial losses and damage to critical infrastructure such as roads, houses, and schools (Amitrano et al., 2024). Nepal ranks tenth on the Global Climate Risk Index due to unpredictable weather, complex topography, and limited institutional resilience (World Bank, 2024). Hydro-meteorological events account for more than 80% of disasters in Nepal, and climate change has increased both their frequency and severity (Bhandari, 2025). Extreme

[1] Student, Department of Electronics and Computer Engineering, Institute of Engineering – Pulchowk Campus. Tribhuvan University, Nepal, 080msdsa003.amit@pcampus.edu.np
[2] Assistant Professor, Department of Electronics and Computer Engineering, Institute of Engineering – Pulchowk Campus. Tribhuvan University, Nepal, ranju.shiwakoti@ioe.edu.np
[3] Senior Lecturer, Department of Building Economics, University of Moratuwa, Sri Lanka, nuwanthas@uom.lk

rainfall from September 26-28, 2024, triggered intense floods and landslides, marking the highest level since 1970. This intense rainfall resulted in the deaths of 250 individuals, 18 missing, and 178 injured. The destruction amounted to NPR 46 billion, affecting 302 telecommunication projects, 26 hydropower plants, 1,639 water supply systems, and 146 bridges (National Disaster Risk Reduction and Management Authority [NDRRMA], 2024). Rainfall exceeded 300 mm in different locations within Kathmandu Valley within three days, producing unprecedented levels of water. To reduce loss of life, infrastructure damage and enhance community resilience against increasingly extreme monsoon climates, accurate weather forecasting and robust early warning systems becomes beneficial.

Historically, flood monitoring primarily relied on in-situ methods, manual rain gauges and community-based watch and warn system (Macherera & Chimbari, 2016). In Nepal, local volunteers were utilised to monitor water level manually which only limited to lead of two to three hours (Budimir et al., 2020). These foundational systems have since evolved through advancements in satellite-based Earth Observation and automated telemetry, enabling real-time, broad-scale monitoring (Peng et al., 2025). While modern systems combine complex probabilistic forecasting, previous manual approaches established the essential framework for today's sophisticated disaster risk management (Budimir et al., 2020; Smith et al., 2017).

Flood monitoring has improved from early manual community gauge monitoring and simple hydrological models. Now it uses sophisticated satellite-based early warning systems (Smith et al., 2017; Revilla-Romero et al., 2014). HAND metrics are topography-based metrics highly effective for rapid prediction of flood inundation. HAND normalises the height of the terrain to the nearest drainage network using only digital elevation model, which allows the topography to be mapped efficiently (Nobre et al., 2011). Despite these technological innovations enhancing flood forecasting capabilities, Nepal's early warning systems continue to face major limitations, similar to those worldwide. These stem from weak government-wide strategies, lack of coordination, unclear institutional roles, and poor dissemination of alerts to vulnerable communities (Bhandari, 2025). Due to these persistent gaps and the urgent need for reliable, all-weather mapping during the cloud-prone monsoon season, this study compares baseline U-Net, UNet_Aug, and UNetPP_TMA using multi-temporal Sentinel-1 SAR imagery and HAND topographic guidance for Nepal’s conditions.

2. LITERATURE REVIEW

Flood detection in Nepal is currently a critical priority because of the increasing frequency of monsoon flooding, which continues to threaten agricultural stability, infrastructure, and livelihoods across the Terai lowlands and high-relief river basins (Bhandari, 2025; Kumar et al., 2026). Historically, flood detection was done manually or using simple automated thresholding, e.g. the Otsu method on SAR images. Nevertheless, these conventional methods were compromised by the phenomenon of surface confusion, when the shadow and wet soil resemble the water signature and the natural speckle noise of the SAR data (Amitrano et al., 2024; Pech-May et al., 2023). Subsequently, SAR has transitioned toward achieving rapid, all-weather situational awareness, overcoming the limitations of optical sensors (e.g., Sentinel-2 and Landsat), which are frequently rendered ineffective by persistent cloud cover during the peak monsoon (Kulk et al., 2023; Tarpanelli et al., 2022). Consequently, operational systems e.g. UNOSAT and

Sentinel Asia have increasingly prioritised SAR-based monitoring. These platforms provide essential rapid mappings; they exhibit systemic limitations in Nepal's environments. Specifically, they suffer from inaccuracies, due to complex terrain, vegetation, and a failure to incorporate hydrological constraints in the West Rapti basin (Poudel et al., 2025; United Nations Satellite Centre [UNOSAT], 2024).

Deep Learning (DL) has significantly advanced automated delineation, with architectures such as U-Net consistently outperforming thresholding methods by learning complex spatial hierarchies (Bentivoglio et al., 2022; Vongkusolkit et al., 2023). Regardless of these advantages, standard U-Net architectures often lack the multi-scale contextual understanding required to reliably distinguish transient floodwater from permanent water bodies in speckle-affected mountainous topography, leading to high false-positive rates (Huang et al., 2024). Recent studies have independently tested the HAND index and bi-temporal change detection to mitigate this limitation. However, a key gap remains: the comparative integration of bi-temporal SAR signals with HAND-based hydrological guidance in a single deep learning framework remains underexplored in the flood-mapping literature (Huang et al., 2024; Tupas et al., 2023).

Standard U-Net uses direct skip connections, but these can leave a semantic gap between encoder and decoder features and limit fine-scale spatial learning in flood segmentation (Zhou et al., 2018). U-Net++ reduces this gap through nested skip pathways, which improve feature refinement, boundary delineation, and discrimination between floodwater and spectrally similar dark surfaces such as terrain shadows and dense vegetation in noisy SAR imagery (Zhou et al., 2018; Melgar-García et al., 2023). For the class-imbalance problem common in flood mapping, Focal-Tversky loss gives more weight to hard-to-classify pixels and helps the model detect small, fragmented flood patches more effectively (Abraham & Khan, 2019). While recent cloud-based studies have started combining Sentinel-1 SAR and HAND topographic data for flood mapping, the comparative use of bi-temporal SAR signals with HAND-based hydrological guidance in complex monsoon settings remains underexplored (Tian et al., 2026; Tupas et al., 2023).

3. METHODOLOGY

This study compares baseline U-Net, UNet_Aug, and UNetPP_TMA using multi-temporal Sentinel-1 SAR imagery and the terrain-derived HAND index. By integrating bi-temporal change detection with hydrological conditioning, the flood mapping process have become faster and highly reliable.

This approach helps address common issues with SAR, such as speckle noise and spectral confusion. As a result, flood predictions stay consistent with the actual topography of the area (Huang et al., 2024; Tupas et al., 2023). This means flood mapping can be both computationally efficient and geophysical reliable. Figure 1 shows the entire workflow, covering complete workflow from data collection and preprocessing to model inference and performance evaluation.

To accurately capture the flooding and bi-temporal change detection, Sentinel-1 Synthetic Aperture Radar (SAR) images for the study area was used. The pre-flood image was captured on August 3, 2024, while the post-flood image was captured on October 2, 2024. These images helped to map surface water and identify temporary flood features.

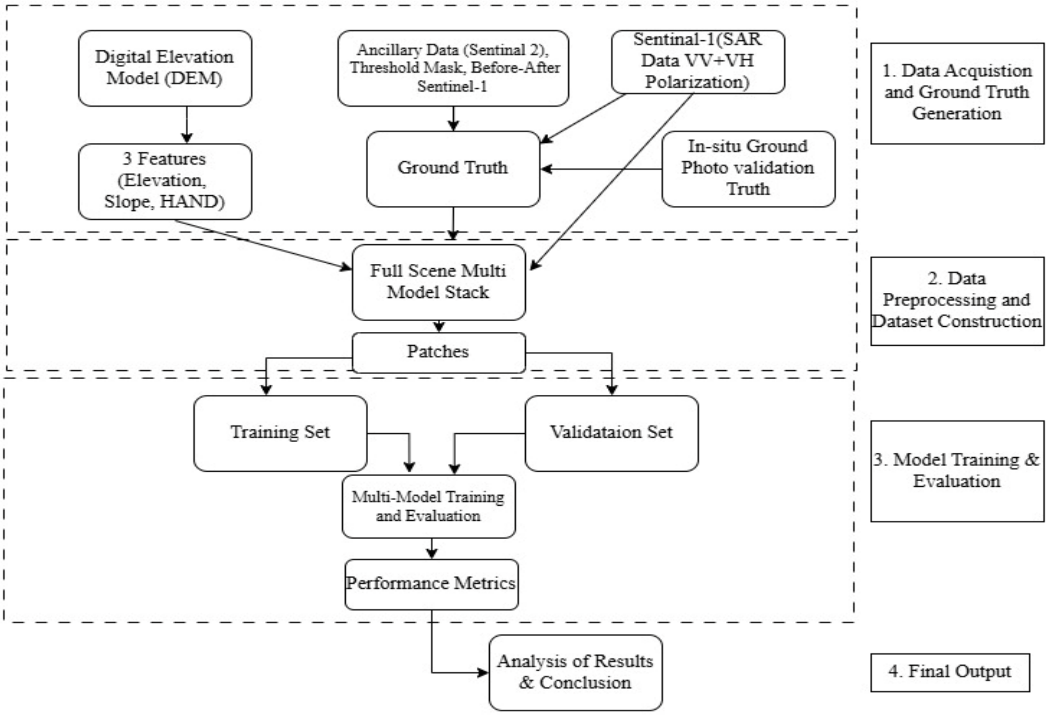

Figure 1: Methodology of the proposed deep learning framework.

The methodology to develop proposed model is illustrated in Figure 1, which outlines the comprehensive pipeline for disaster mapping using multi-source geospatial data. The methodology is structured into four distinct functional phases.

3.1 PHASE 1: DATA ACQUISITION AND GROUND TRUTH GENERATION

The flooded data collection area covers Nawalpur district of Nepal, with a focus on flood event. Multi-source data were integrated for model training and evaluation. Sentinel-1 SAR imagery (VV,VH polarisations) was acquired for pre- and post-flood dates. Sentinel-2 optical data, threshold masks, and bi-temporal SAR comparisons were fused to generate a ground mask. These masks were further refined using in-situ ground photographs to ensure high fidelity and representativeness of real-world flood conditions. Topographic data, including elevation, slope, and HAND indices, were derived from the NASADEM Digital Elevation Model using Google Earth Engine.

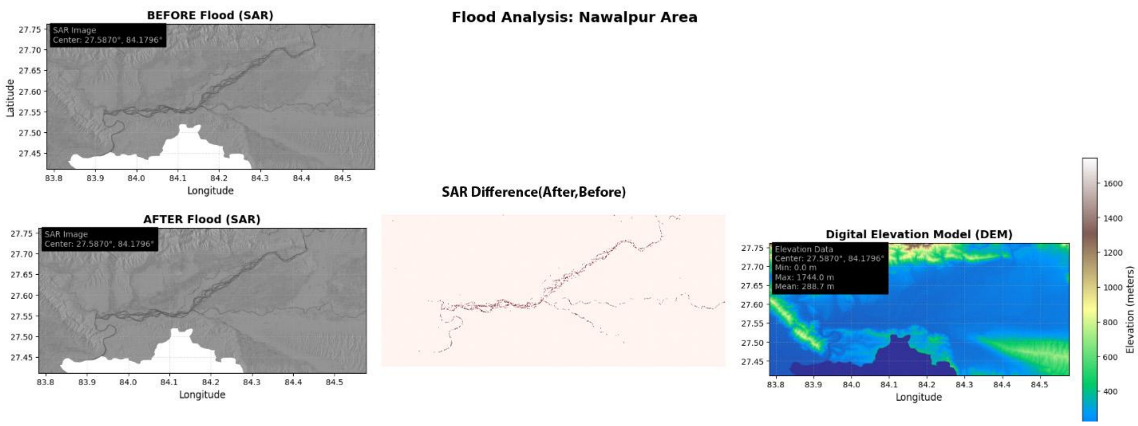

Figure 2: Before flood, after flood, SAR difference and digital elevation model

Figure 2 shows the study area before flood on dry season observed on July 16, 2024, of Nawalpur district, Nepal, from Sentinel-1. The post-flood image, acquired on 2nd October 2024, and pre-flood image were used to analyse surface changes on SAR Difference image. SAR imagery in VV and VH polarisations was utilised to detect changes in water extent, comparing conditions "before" and "after" the event. Digital Elevation Model (DEM) of the collected data shows elevation, slope and HAND with the goal of providing an environmental context of the Nawalpur region. The NASADEM elevation data was processed in Python with RichDEM and NumPy to compute the HAND index. Flow direction and accumulation rasters were first generated from the DEM, after which the nearest drainage network was extracted, and HAND was calculated as the vertical height of each cell above its downslope drainage path.

3.2 PHASE 2: DATA PREPROCESSING AND DATASET CONSTRUCTION

Raw Sentinel-1 data underwent standard preprocessing, including orbit file application, thermal noise removal, radiometric calibration to σ^0, terrain correction using the SRTM 30 m DEM, 5×5 refined Lee speckle filtering, and conversion to a decibel scale. Pre- and post-event SAR channels (VV, VH) were then combined with the normalised Height Above Nearest Drainage (HAND) index. Five channel input tensor was independently scaled between 0 and 1. A HAND threshold of 15 m was selected as a conservative, study-specific operational buffer caused by radar shadow and DEM uncertainty. Because HAND-derived flood extent is sensitive to local topography, DEM preprocessing, and drainage-network definition, the threshold was treated as a calibrated parameter rather than a universal cutoff.

Pre-processed full scene stack (3,910 × 8,851 pixels) and corresponding ground truth were divided into 128×128 pixel non overlapping 1500 patches of spatial tiles. Strict geographic split was used rather than randomly shuffling the patches to prevent spatial data leakage. Which is a critical issue in remote sensing where overlapping or adjacent patches artificially inflate validation metrics. The broader scene was spatially partitioned to ensure that training and validation patches were drawn completely distinct geographic blocks, preventing any boundary pixels from leaking across sets. The dataset was divided into 80/20 ratio, resulting in 1,200 independent training samples and 300 validation samples.

3.3 PHASE 3: MODEL TRAINING AND EVALUATION

Model development and training were done in Python using the TensorFlow framework to maintain computational efficiency. Deep learning model was trained on the patched dataset using the combined SAR and topographic features, where the attention mechanism focused on relevant spatial features while suppressing noise inherent in the raw Sentinel-1 data. For comparative evaluation, baseline U-Net, UNet_Aug, and UNetPP_TMA were benchmarked under identical training conditions. Following an 80/20 training and validation split, all three deep learning architectures were trained from scratch for exactly 50 epochs utilising a batch size of 16. Optimisation was performed using the Adam optimiser with an initial learning rate of 0.0001, which dynamically updated the network weights by minimising a combined binary cross-entropy and Dice loss function.

As illustrated in the third phase of Figure 1, three CNN models—baseline U-Net, UNet_Aug, and UNetPP_TMA—were implemented and compared:

1. Baseline U-Net which takes the five-channels as the input (before flood VV/VH, after flood VV/VH, and HAND).

2. U-Net with augmentation (UNet_Aug), which uses the same architecture as the baseline but trained using data augmentation, by random flips and 90-degree rotations.

3. U-Net++ with Focal-Tversky Loss (UNetPP_TMA), uses more complex U-Net++ architecture with dense skip connections to aggregate features. UNetPP_TMA uses focal-Tversky loss, which is determined by the formula: TFL= (1 - TI)^γ, where TI is Tversky Index), with α=0.7 (higher penalty on false positives to avoid over-prediction in imbalanced data), β=0.3 (lower on false negatives, as per Abraham & Khan, 2019, for small-target segmentation like floods), and γ=0.75 (equivalent to 3/4, focusing on hard examples by down weighting easy ones, as this non-linearity balances convergence without over-suppression, empirically optimal in original studies and variants for rare-class detection).

3.4 PHASE 4: ANALYSIS OF RESULTS AND CONCLUSION

The trained model generates flood inundation maps. Results are analysed through quantitative metrics. The quantitative assessment of the methodology, based on the final stage of Figure 1, is presented through precision, recall, and F1-score. Precision measures the reliability of positive predictions by measuring the proportion of correctly classified flooded pixels out of all pixels classified as flooded and how many are actually flooded out of all the pixels that are classified as flooded, as Precision = TP / (TP + FP), where TP and FP are the number of pixels correctly identified as flooded and the number of pixels incorrectly identified as flooded respectively.

Recall measures the completeness of the model in identifying actual flooded areas, which is Recall = TP / (TP + FN), where FN denotes false negatives and TP denotes true positive. Recall reduces errors of omission and ensures that the regions at risk of being flooded do not escape disaster management and monitoring. F1 score, a harmonic mean of the precision and recall measures, is a single-value measure of the classification performance which is more useful with class imbalance caused by the large number of non-flooded pixels in flood datasets. It is calculated with F1 = 2 (Precision * Recall)/ (Precision+ Recall). This measure is commonly cited in the deep learning flood segmentation literature as it combines both error types.

Additionally, Cohen's Kappa coefficient (κ) accounts for agreement beyond chance, formulated as $\kappa = (p_o - p_e) / (1 - p_e)$, where p_o is the observed agreement and p_e is the expected agreement by chance. The Area Under the Receiver Operating Characteristic Curve (ROC-AUC) evaluates threshold-independent performance by integrating the true positive rate against the false positive rate across classification thresholds, expressed as Area Under Curve.

4. RESULTS AND PERFORMANCE EVALUATION

Quantitative metrics, including IoU, F1-score, Precision, Recall, Cohen's Kappa, and AUC, were employed to comparatively evaluate three deep learning architectures: the baseline U-Net, U-Net with augmentation (UNet_Aug), and U-Net++ with Focal-Tversky loss (UNetPP_TMA).

Table 1: Quantitative comparison of deep learning models for flood segmentation.

Model	IoU	F1-Score	Precision	Recall	Kappa	Overall Accuracy	AUC
Baseline U-Net	0.5803	0.7344	0.7201	0.7493	0.7205	0.9736	0.9846
UNet_Aug	**0.6101**	**0.7578**	**0.7496**	0.7663	**0.7453**	**0.9761**	**0.9901**
UNetPP_TMA	0.5686	0.725	0.6846	**0.7704**	0.71	0.9715	0.9855

From Table 1, the UNet_Aug model achieved a IoU (0.6101), indicating better spatial overlap between the predicted and ground truth flood segments. In the context of semantic segmentation, IoU values within the 0.50–0.70 range are generally recognised as indicating moderate and practically useful performance. The model yielded a 3.2% increase in F1-Score over the baseline. Consistent with the data scarcity challenges outlined by Bentivoglio et al. (2022), these performance gains demonstrate that spatial augmentation effectively enhances model generalisation and robustness in complex SAR-based flood mapping. The UNet_Aug model effectively balanced Precision (0.7496) and Recall, resulting in the highest overall F1-Score and an AUC-ROC of 0.9901. This indicates excellent discriminative ability (>0.90 benchmark; Mandrekar, 2010).

Conversely, UNetPP_TMA achieved the highest Recall (0.7704), minimising omission errors, yet at a lower Precision (0.6846), increasing commission errors and reducing overall IoU and F1-Score. The Focal-Tversky Loss, with parameters α=0.7 (penalising false negatives), β=0.3 (penalising false positives), and γ=0.75, effectively targeted the imbalanced flood class, however, compromised prediction purity (Abraham & Khan, 2019). UNet_Aug is the recommended overall model because it achieved the best balance of IoU, F1-score, precision, Kappa, overall accuracy, and AUC.

Figure 3: Intersection over Union (IoU) of U-Net, Unet_Aug, and UnetPP_TMA

Figure 3 illustrates the convergence of the Intersection over Union (IoU) across training epochs, representing the spatial overlap between predictions and ground truth. While the baseline U-Net displays a smooth learning curve, the UNet_Aug model demonstrates the most consistent and superior convergence, ultimately achieving the highest overall segmentation accuracy. UNetPP_TMA exhibits fluctuations in validation performance; however, it reaches competitive IoU levels in later epochs, indicating that despite initial training instability, it performs robustly by the end of the training cycle.

Figure 4 shows how precision changes over the training epochs for three models. For each model training and validation curve shows how often models are correct when it predicts a flood pixel. Higher curves show few false positive flood detections.

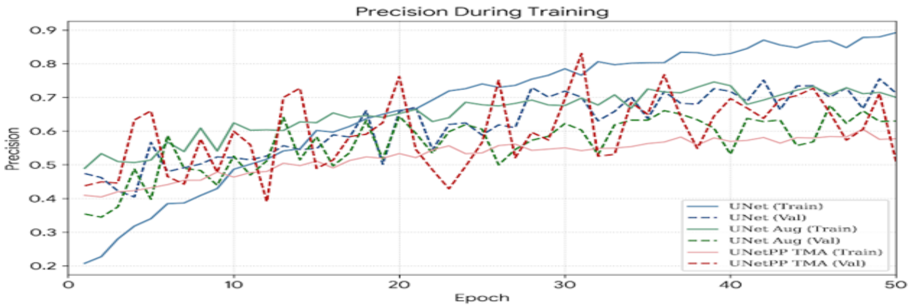

Figure 4: Precision of U-Net, Unet_Aug, and UnetPP_TMA

Solid lines are used for training precision, while dashed lines represent validation precision. Blue curves are for U-Net, green curves for UNet_Aug and then red curves are for UNetPP_TMA. The U-Net training curve rises steadily from a low starting point and reaches the highest final precision, which indicates the model is improving over time. UNet_Aug exhibits more consistent relationship between training and validation precision, which indicates improved generalisation and greater training stability. UNetPP_TMA displays stronger fluctuations in both curves, which shows less stable convergence. Overall, UNet_Aug shows balanced learning among the evaluated models.

All three models were able to distinguish temporary floodwater from permanent water bodies via bi-temporal change detection. Qualitative comparisons of the outputs with and without HAND integration (Figure 5) demonstrated that the topographical prior successfully eliminated implausible predictions. This effectively restricted flood extents to low-lying, drainage-adjacent areas, significantly improving physical realism for emergency response applications.

A B C

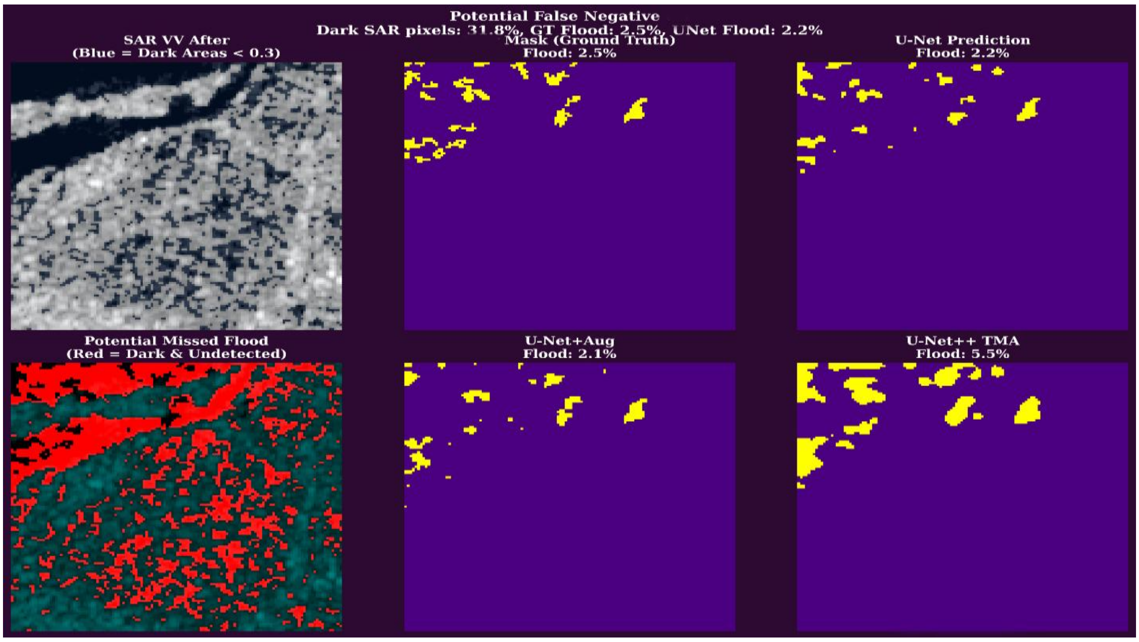

Figure 5: Qualitative comparison of flood predictions with and without HAND-based hydrological mask.

Figure 5 shows flood predictions from the evaluated models over a challenging patch that is prone to false negatives where HAND is used as topographic guidance. Figure 5(A) shows the post-event Sentinel-1 VV SAR image, where dark, low-backscatter areas is less than 0.3 and cover approximately 31.8% of the pixels. The ground truth (Figure 5B) contains about 2.5% flooded pixel, the baseline U-Net (Figure 5C) contains about 2.2%, the augmented U-Net about 2.1% flooded pixel, and the UNetPP_TMA about 5.5%. Figure 5A contains extensive dark low-backscatter regions, which may represent floodwater, smooth surfaces, radar shadow, or speckle rather than flood. When compared with the ground truth, the baseline U-Net and UNet_Aug produce similar flood proportions, whereas UNetPP_TMA predicts a larger and more continuous flooded extent along drainage-connected areas. The comparison shows that HAND improves the physical realism of the predictions by constraining flood extent to low-lying terrain and reducing implausible detections outside hydrologically plausible zones.

This visual comparison confirms deep learning model was able to completely capture flood extents completely and realistically than the simple U-Net baselines. This result maps are physically consistent and well suited for emergency response.

A B C D E

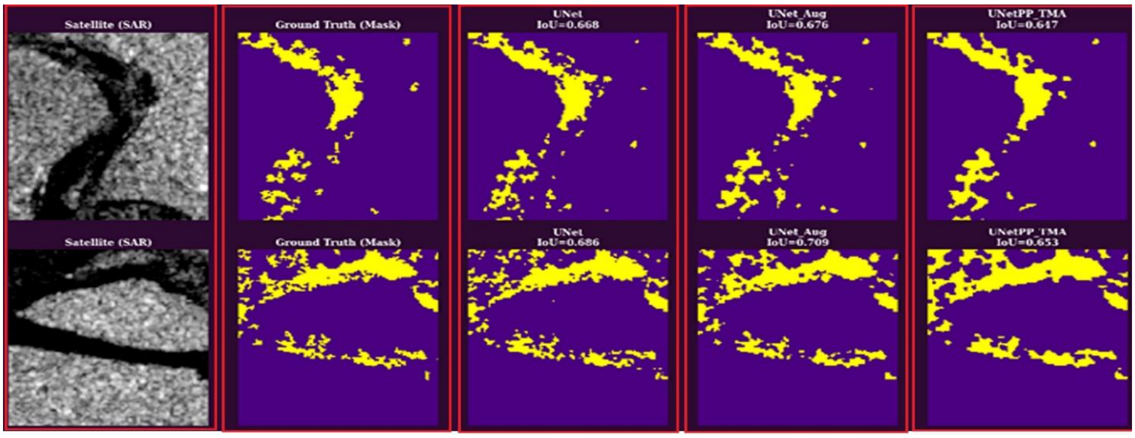

Figure 6: Prediction of flood from 3 different model and satellite SAR image

Figure 6 provides a visual comparison of the flood-extent predictions produced from three deep-learning models. The input SAR VV polarisation image and ground truth binary mask are shown in figures 6A and 6B, respectively. The base line model in Figure 6C tends to produce conservative flood estimation with fragmented flood patches. This might be due to limited sensitivity to subtle flood signal under SAR speckle noise. figure comprises multiple panels: input SAR VV-polarization imagery is shown in Figure 6(A), and the ground truth binary mask is shown in Figure 6B. These visualisations illustrate distinct qualitative differences in flood delineation. The simple baseline model (refer Figure 6C) often yields conservative estimates, characterised by smaller, fragmented yellow regions that may under-represent subtle inundated areas obscured by SAR speckle noise. The augmented model (refer Figure 6D) demonstrates moderate gains in generalisation and spatial extent which captures the strongest overall flood extent representation. This comparative evaluation highlights the proposed framework's ability to generate accurate, practical flood maps suitable for emergency response applications.

The bi-temporal Sentinel-1 SAR image combined with HAND in U-Net has been successfully applied to develop deep learning techniques for flood mapping in a monsoon climate. The Overall Accuracy of the UNet_Aug model was 0.9761, IoU was 0.6101 and F1-Score was 0.7578. In general, standard U-Net on complex SAR datasets without topographic constraints have IoU ranging from 0.35 to 0.48. The proposed framework was able to tackle the complexity of the monsoon floods, by including spatial augmentation and hydrological conditioning, avoiding the large performance drop and high false-positive rate often seen in urban and topographically challenging environments (Bentivoglio et al., 2022). The model's competitive IoU is largely attributable to the successful removal of permanent water false positives via bi-temporal change detection. While these results are highly robust for complex terrains, future work could further improve boundary delineation by integrating finer-resolution Digital Elevation Models (DEMs) to reduce topographical backscatter errors, aiming for the high-resolution improvements demonstrated in advanced urban SAR fusion studies.

UNet_Aug emerges as the optimal model architecture for operational flood detection because it achieved the strongest overall performance. This framework, combining SAR robustness, hydrological physics, and efficient computation, advances reliable, all-weather flood mapping in data-scarce regions while maintaining computational practicality for real-time deployment.

5. CONCLUSIONS AND RECOMMENDATIONS

This study presents a robust, deep learning framework for operational flood detection in topographically complex monsoon regions, successfully validated using the 2024 Nawalpur flood event in Nepal. UNet_Aug achieved the highest overall balance of metrics (IoU = 0.6101, F1-score = 0.7578, precision = 0.7496, Kappa = 0.7453, overall accuracy = 0.9761, and AUC = 0.9901), while UNetPP_TMA achieved the highest recall (0.7704). In the context of emergency response, missing an active flood zone is unacceptable due to life-safety. Therefore, UNet_Aug emerges as the optimal operational model because it provides the strongest overall balance of flood-mapping performance.

This approach addresses critical gaps in current flood monitoring systems in disaster-prone regions. By prioritising sensitivity to inundated areas, the framework supports evidence-based emergency resource allocation, thereby enhancing regional resilience to increasing climatic hazards. This study recommends the operational deployment of UNet_Aug within Nepal’s flood early warning systems, while UNetPP_TMA may be retained as a high-recall alternative for emergency screening.

Future research should extend this framework by including DEMs from high resolution LiDAR and output from real-time hydrologic models to further enhance the topographic guidance over a variety of terrain types. Furthermore, combining multiple deep-learning architectures with uncertainty quantification techniques such as Bayesian inference would greatly improve the predictive accuracy of near-real-time operations. Finally, the scale expansion of the framework to other natural hazards, such as landslides and wildfires, and for future global SAR missions (e.g., NASA-ISRO Synthetic Aperture Radar) will help to foster broader adoption of this important tool in data-sparse developing contexts.

6. REFERENCES

Abraham, N., & Khan, N. M. (2019). A novel focal Tversky loss function with improved attention U-Net for lesion segmentation. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (pp. 683–687). IEEE. https://doi.org/10.1109/ISBI.2019.8759329

Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., & Ruello, G. (2024). Flood detection with SAR: A review of techniques and datasets. *Remote Sensing*, *16*(4), 656. https://doi.org/10.3390/rs16040656

Bentivoglio, R., Isufi, E., Jonkman, S. N., & Taormina, R. (2022). Deep learning methods for flood mapping: A review of existing applications and future research directions. *Hydrology and Earth System Sciences*, *26*(16), 4345–4378. https://doi.org/10.5194/hess-26-4345-2022

Bhandari, D. (2025). Challenges and opportunities in Nepal's early warning communication. *International Journal of Disaster Risk Reduction*, *119*, 105318. https://doi.org/10.1016/j.ijdrr.2025.105318

Budimir, M., Donovan, A., Brown, S., Shakya, P., Gautam, D., Uprety, M., Cranston, M., Sneddon, A., Smith, P., & Dugar, S. (2020). Communicating complex forecasts: An analysis of the approach in Nepal's flood early warning system. *Geoscience Communication*, *3*(1), 49–70. https://doi.org/10.5194/gc-3-49-2020

Dantas, A. A. R., & Paz, A. R. (2021). Use of HAND terrain descriptor for estimating flood-prone areas in river basins. *Revista Brasileira de Ciências Ambientais*, *56(3)*, 501–516. https://doi.org/10.5327/z21769478892

Huang, B., Li, P., Lu, H., Yin, J., Li, Z., & Wang, H. (2024). WaterDetectionNet: A new deep learning method for flood mapping with SAR image convolutional neural network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 14471–14485. https://doi.org/10.1109/JSTARS.2024.3440995

Kumar, A., Pandey, G., & Kale, R. V. (2026). Ensemble machine learning and deep learning framework for flood susceptibility mapping in the transboundary Rapti River Basin. *Environmental Earth Sciences*, *85*, 168. https://doi.org/10.1007/s12665-026-12912-6

Macherera, M., & Chimbari, M. J. (2016). A review of studies on community based early warning systems. *Jàmbá: Journal of Disaster Risk Studies*, *8*(1), a206. https://doi.org/10.4102/jamba.v8i1.206

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, *5*(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

Melgar-García, L., Martínez-Álvarez, F., Tien Bui, D., & Troncoso, A. (2023). A novel semantic segmentation approach based on U-Net, WU-Net, and U-Net++ deep learning for predicting areas sensitive to pluvial flood at tropical area. *International Journal of Digital Earth*, *16*(1), 3661–3679. https://doi.org/10.1080/17538947.2023.2252401

National Disaster Risk Reduction and Management Authority. (2024). *Nepal: 2024 September floods and landslides situation report #4 as of 16 October 2024*. ReliefWeb. https://reliefweb.int/report/nepal/nepal-2024-september-floods-and-landslides-situation-report-4-16-october-2024

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., & Saleska, S. (2011). Height above the nearest drainage – a hydrologically relevant new terrain model. *Journal of Hydrology*, *404*(1-2), 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051

Pech-May, F., Aquino-Santos, R., & Delgadillo-Partida, J. (2023). Sentinel-1 SAR images and deep learning for water body mapping. *Remote Sensing*, *15*(12), 3009. https://doi.org/10.3390/rs15123009

Peng, X., Chen, S., Miao, Z., Xu, Y., Ye, M., & Lu, P. (2025). Automatic flood monitoring method with SAR and optical data using Google Earth Engine. *Water*, *17*(2), 177. https://doi.org/10.3390/w17020177

Poudel, P., Pokhrel, P., & Thapa, S. (2025). Hazard modeling using Sentinel data for risk assessment and management of Tsho Rolpa Glacier, Nepal. In *Proceedings of the International Conference on Remote Sensing* (pp. 1–8). SCITEPRESS. https://www.scitepress.org/Papers/2025/132887/132887.pdf

Revilla-Romero, B., Thielen, J., Salamon, P., De Groeve, T., & Brakenridge, G. R. (2014). Evaluation of the satellite-based global flood detection system for measuring river discharge: Influence of local factors. *Hydrology and Earth System Sciences*, *18*(11), 4467–4484. https://doi.org/10.5194/hess-18-4467-2014

Smith, P. J., Brown, S., & Dugar, S. (2017). Community-based early warning systems for flood risk mitigation in Nepal. *Natural Hazards and Earth System Sciences*, *17*(3), 423–437. https://doi.org/10.5194/nhess-17-423-2017

Tarpanelli, A., Mondini, A. C., & Camici, S. (2022). Effectiveness of Sentinel-1 and Sentinel-2 for flood detection assessment in Europe. *Natural Hazards and Earth System Sciences*, *22*(8), 2473–2489. https://doi.org/10.5194/nhess-22-2473-2022

Thomas, J., Mohan, S., Mallik, S., Oommen, T., Xue, P., Meadows, G., Thalakkottukara, N. T., & Williams, R. (2025). Development of a HAND-based flood risk assessment tool in Google Earth Engine for a data-scarce region in the US. *Journal of Great Lakes Research*, 102510. https://doi.org/10.1016/j.jglr.2025.102510

Tian, D., Wang, L., Liu, H., Cohen, S., & Shu, S. (2026). DeepSAR Flood Mapper: global flood mapping on Google Earth Engine cloud platform using MLP deep learning model with Sentinel-1 SAR imagery and HAND topographic data. *GIScience & Remote Sensing*, *63*(1), Article 2612306. https://doi.org/10.1080/15481603.2025.2612306

Tupas, M. E., Roth, F., Bauer-Marschallinger, B., & Wagner, W. (2023). Improving Sentinel-1 flood maps using a topographic index as prior in Bayesian inference. *Water*, *15*(23), 4034. https://doi.org/10.3390/w15234034

United Nations Satellite Centre (UNOSAT). (2024, October 2). *Satellite detected water extent Koshi and Madhesh Provinces, Nepal as of 01 October 2024*. United Nations. https://unosat.org/products/3990

Vongkusolkit, J., Peng, B., Wu, M., Huang, Q., & Andresen, C. G. (2023). Near real-time flood mapping with weakly supervised machine learning. *Remote Sensing*, *15*(13), 3263. https://doi.org/10.3390/rs15133263

World Bank. (2024). *Nepal disaster resilience second DPC with Cat DDO (NDRC): Annex - Climate change technical note*. World Bank Group. https://documents.worldbank.org/pt/publication/documents-reports/documentdetail/099511210252413915

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. In D. Stoyanov, Z. Taylor, B. Kainz, G. Zheng, R. R. Shamir, V. Diaz-Pinto, P. A. Cook, Y. Hu, S. Khan, T. Arbel, M. Blaschko, G. Carneiro, & D. Samaras (Eds.), *Deep learning in medical image analysis and multimodal learning for clinical decision support* (Lecture Notes in Computer Science, Vol. 11045, pp. 3–11). Springer. https://doi.org/10.1007/978-3-030-00889-5_1