# Flood Hazard Modelling Assessment Using Deep Learning and Earth Observation

**Binita Shahi[1], Arjun Poudel[2], Ashmeera Dahal[3] & Sandesh Upadhyaya[4]**
shahiii.binita@gmail.com, arjpou268@gmail.com, ashmeeradahal04@gmail.com, sandeshupadhya1@gmail.com
[1]Land Management Training Center, [2]Survey Office Palpa, [3]Ministry of Land Management and Cooperatives and Poverty Alleviation, [4]Survey Department

## KEYWORDS

*Flood Modelling, U-Net, Deep Learning, CNN and Softmax*

## ABSTRACT

*This study explores the integration of satellite imagery with deep learning models, specifically the U-Net architecture, for flood hazard modelling and detection in flood-prone regions of Uttar Pradesh, India. The research focuses on using Sentinel-1 satellite data, combined with multispectral imagery, to develop a robust flood detection system. The methodology involves preprocessing satellite images, creating flood polygons through change detection, and training a U-Net model with a ResNet34 backbone on a dataset of 28,896 image patches. The model achieved an accuracy of 92.82% and an Area Under the Curve (AUC) score of 90%, demonstrating strong performance in identifying flooded areas. The study highlights the effectiveness of combining remote sensing and deep learning techniques for flood mapping, providing high-resolution flood maps validated against ground truth data. The results underscore the potential of AI-driven approaches in enhancing flood risk assessment, disaster response, and management strategies. In order to address class imbalance, optimizing threshold selection is necessary and incorporating additional data sources, such as LiDAR and hydrological models, can improve model performance and applicability. This research contributes to advancing flood monitoring and mitigation efforts, offering a scalable and efficient solution for flood-prone regions.*

## 1 INTRODUCTION

Flooding in the state of Uttar Pradesh, India affected over 245,000 people in August 2022 (ReliefWeb, 2022). The Ganges was above the danger mark in at least 5 locations, including Varanasi (IMD, 2022). According to the National Emergency Response Centre (NERC) in India, flooding has affected over 1,000 villages across 22 districts in the state since 26 August 2022. Authorities had opened 386 relief camps, which, as of 29 August, were housing 16,562 evacuees. As of 31st August 2022, NERC reported 245,585 people were affected and at least 4 people died during the flood.

Satellite imageries combined with U-Net-based deep learning methods provide a powerful approach for detecting and analyzing floods (Pech-May et al., 2023). Satellites equipped with remote sensing technology capture high-resolution images of the Earth's surface. These images can show land use, water bodies and changes over time. By taking images before, during, and after flooding events, satellites provide a temporal perspective, which is crucial for assessing the extent and impact of floods (NASA, 2020). As satellites can capture data in various spectral bands (e.g., visible, infrared), this allows for identifying water bodies and differentiating flooded areas from dry land based on their spectral signatures.

Machine learning (ML) is a subset of artificial intelligence (AI) that involves the development of algorithms and statistical models that enable computers to perform tasks without explicit programming (Ian et al., 2016). ML allows systems to learn from data, identify patterns, and make decisions with minimal human intervention (Mitchell, 1999). Unlike traditional computing methods that follow predefined rules, machine learning systems improve their performance over time as they are exposed to more data, adapting to new information and conditions.

Remote sensing involves the collection of data from a distance, typically using satellites, aircraft, or UAVs (unmanned aerial vehicles) equipped with various sensors to capture imagery of the Earth's surface (Panda et al., 2015). The data collected by these sensors can include optical, infrared, radar, and LiDAR (Light Detection and Ranging) measurements, offering a rich and complex representation of the environment (Panda et al., 2015). However, analyzing these large datasets, especially in real-time, poses a significant challenge due to the vast amount of information and the high-dimensional nature of the data. This is where machine learning comes in. ML algorithms excel at handling large volumes of data and can automatically identify patterns, classify objects, and extract meaningful information from remote sensing images (Lary et al., 2016). These capabilities are invaluable for applications such as environmental monitoring, disaster management, urban planning, and agricultural monitoring.

Machine learning techniques can process remote sensing data more efficiently than traditional methods. For example, in the context of satellite imagery, ML algorithms can be trained to detect land use changes, classify vegetation types, or identify flood-prone areas. These automated systems can analyze vast regions and provide insights at speeds far greater than manual analysis. Moreover, machine learning can improve over time as more data is available, allowing for increasingly accurate predictions and analyses.

U-Net is a convolutional neural network designed for biomedical image segmentation, but it also excels at segmenting satellite images (Ronneberger et al., 2015). It consists of an encoder-decoder structure with skip connections, allowing for precise localization of features (Ronneberger et al., 2015). A U-Net model can be trained on labeled datasets where different classes (e.g., water, land, urban areas) are annotated. This training enables the model to learn the characteristics of flooded and non-flooded regions. Once trained, the U-Net can process new satellite images to segment and identify flooded areas by classifying every pixel as either "flood" or "non-flood". This segmentation is critical in providing detailed maps of affected regions. U-Net can help estimate the extent of flooding soon after an event using real-time satellite images, allowing for timely responses (Pech-May et al., 2023). By comparing images over time, U-Net can effectively highlight changes in land cover due to flooding, helping in damage assessment. Flood maps generated using this method can be integral for disaster response planning, risk assessment, and developing management strategies. Combining U-Net with other machine learning techniques or algorithms (like ensemble methods) can improve accuracy in detecting and classifying flooded areas (Pech-May et al., 2023). Thus, the integration of satellite imagery with U-Net deep learning models provides a robust mechanism for effectively monitoring and managing flood events, facilitating better preparedness and response strategies.

# 2 MATERIALS AND METHODS

## 2.1 Study Area

For this study, about 1635 sq. km area was chosen that lies within the Ghazipur, Ballia, Siwan, Saran and Bhojpur districts of the Uttar Pradesh State of India, which is located in the bank of the Ganges River focusing on the high-flood areas during heavy rainfall in August and September 2022 as shown below in Figure 1.

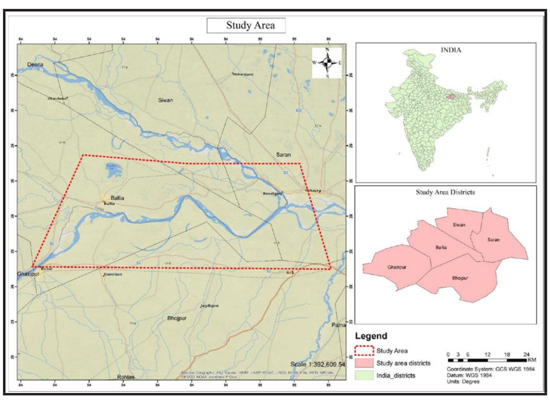


*Figure 1: Study Area Map*

## 2.2 Research Method

The methodology adopted is described as a process for flood mapping using Sentinel-1 data and Multispectral data, combining remote sensing techniques with deep learning to achieve precise flood extent detection. The process starts with the acquisition of pre- and post-flood satellite images, followed by preprocessing in the Sentinel Application Platform (SNAP). This stage involves performing subsetting, radiometric and geometric corrections, collocation of data and applying band math to differentiate water surfaces. The flood extent is then extracted through methods like thresholding, change detection and GIS-based digitization. To enhance the classification accuracy, Multispectral optical data is incorporated. The resulting flood extent is compared to ground truth data using Area Under the Receiver Operating Characteristic Curve (AUC-ROC) analysis. If the accuracy does not meet the required threshold, a Convolutional Neural Network (CNN)-U-Net deep learning model is trained with labeled data to further refine the flood extent using AI-powered segmentation. After validation, the final flood map is generated, offering improved precision for flood assessment.

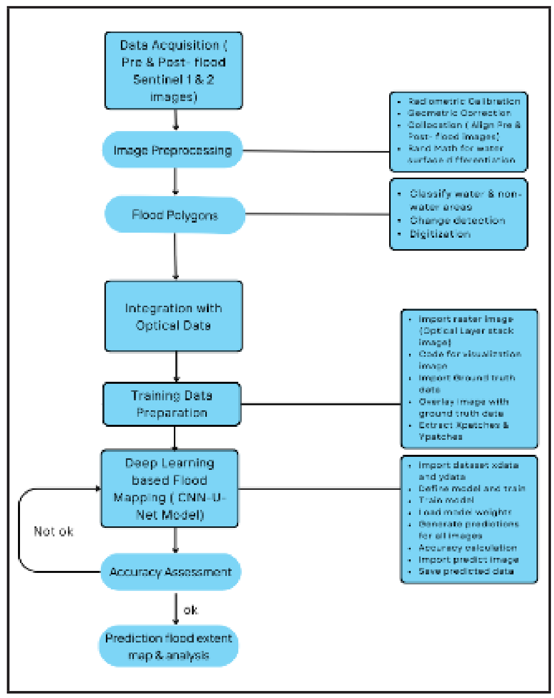


*Figure 2: Methodological workflow for Flood Modelling*

### *2.2.1 Training Data Preparation*

#### *2.2.1.1 Data Acquisition*

In order to develop the ground truth data, Sentinel 1A images with a resolution of 10 meters were used. Similarly, for creating training datasets and model development, High-resolution multispectral imagery with a resolution of 3 meters was used.

#### *2.2.1.2 Image Preprocessing*

The SNAP tool was used for preprocessing images, including subsetting, radiometric and geometric correction, multilooking, collocation as well as band math operations like Theta/Theta for master and slave images of collocated datasets.

#### 2.2.1.3 Flood Polygons

The change detection process was carried out for the pre-flood and post-flood images after being preprocessed, depicting water and non-water areas. After that, flood areas were digitized in Quantum Geographic Information System (QGIS) environments, resulting in a vector polygon dataset called flood polygons. Each polygon was carefully delineated based on visual interpretation and the flood extent derived from Synthetic Aperture Radar (SAR) ratio calculations.

### 2.2.2 Integration with Optical Data (Training Sample in CNN)

Training samples are the datasets used to train the model and let it learn the features and patterns of the data (Lary et al., 2016). Without high-quality training data, even the most efficient machine learning algorithms will fail to perform (Lary et al., 2016). The need for quality, accurate, complete, and relevant data is required in the initial training process (Mohammed et al., 2022). Only if the algorithm is fed with good training data, it can easily pick up the features and find relationships that it needs to predict (Sanyal et al., 2021). The training process involves presenting the model with input data (features) along with the corresponding correct output (labels or targets), allowing the model to learn the patterns and relationships within the data (Zhao et al., 2020). The quality and quantity of training samples directly impact the performance of the trained model (Halevy et al., 2009).

Optical images from the planet website before and after the flood were imported and visualization was carried out in the Collab environment, where vector data of flood polygons were rasterized and all layers were overlaid to define x patches and y patches of the training sample. The training samples were assigned with corresponding class labels, forming the basis for supervised classification and model training. This dataset served as a crucial input for flood mapping and accuracy assessment, enhancing the reliability of the final classification results.

### 2.2.3 Deep Learning based Flood Mapping (CNN-U-Net Model)

The U-Net model, a type of convolutional neural network (CNN), is widely used for flood data modeling due to its ability to perform pixel-wise segmentation (Ronneberger et al., 2015). It follows an encoder-decoder architecture, where the down-sampling (encoder) path extracts spatial features, and the up-sampling (decoder) path reconstructs the segmented image while preserving spatial details (Ronneberger et al., 2015).

For flood detection, SAR (Sentinel-1) and multispectral (Sentinel-2) images were used as input. The training dataset, created from digitized flood extent polygons, helped the U-Net learn the distinction between flooded and non-flooded areas. The activation function (e.g., ReLU) enhanced feature extraction, while the sigmoid or SoftMax function at the output layer classified flood-affected regions.

To optimize performance, hyperparameter tuning (learning rate, batch size and number of filters) was applied. Loss functions such as binary cross-entropy or Dice loss improved segmentation accuracy. The trained U-Net model effectively predicted flood-affected areas, providing high-resolution flood maps for disaster response and management.

### 2.2.4 Accuracy Assessment

After training, the model was tested on an unseen validation image, a georeferenced GeoTIFF file, provided by project partners. The model predicted flood-prone areas, which were binarized using a 0.50 threshold.

The predicted results were compared against digitized ground truth datasets derived from SNAP-processed imagery and validated using QGIS. Visual comparisons showed a strong spatial correlation between predicted flood areas and ground truth samples.

## 3 RESULTS

### 3.1 Flood Mapping Dataset Preparation

Initially, the layer stack of pre and post-flood multispectral images were used as an input and then, ground truth data were overlaid on layer stacked image. Then, the patch dimensions of XData obtained were (28896, 32, 32, 6) and that of YData obtained were (28896, 32, 32, 1) for model training. Then, a randomly selected patch (index 555) was visualized to examine its spectral properties and corresponding flood classification. The final dataset was stored in NumPy (.npy) format for further deep-learning applications.

#### *3.1.1 Ground Truth Data*

The ground truth flood and non-flood polygons were extracted from Sentinel 1A imagery of pre and post-flood after applying correction, band math and collocation. Then, the digitization of sample flood polygons was integrated and overlaid onto the raster for validation, which is shown in Figure 3.

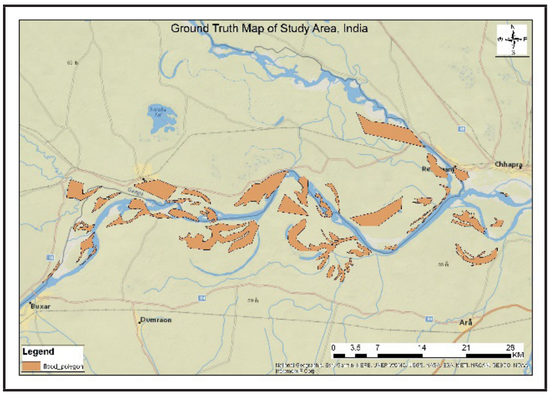


*Figure 3: Ground Truth Data of Study Area*

#### *3.1.2 Input Dataset with Ground Truth Data*

The multispectral images of the study area, captured before and after the flood, were used as the input dataset after layer stacking. This dataset was defined as XData and overlaid with ground truth data, which was defined as YData which is shown in below Figure 4.

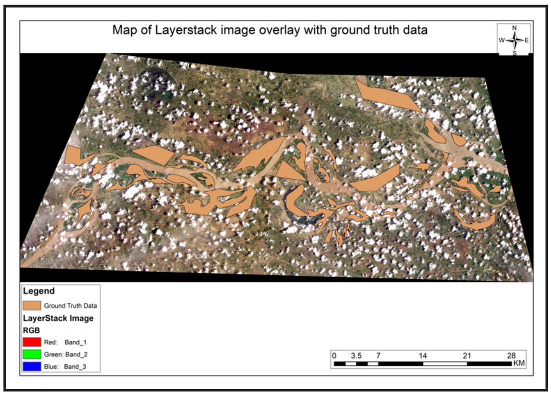


*Figure 4: Layer stack image overlaid with ground truth data.*

### 3.2 Data Modeling

In this phase, XData and YData were the input dataset. The AUC score was calculated using a weightage file where the XData and YData were split into 70% training (20,227 samples) and 30% testing (8,669 samples) and was obtained 90% and that of accuracy score was obtained 92.82%. Then, Predictions were compared with ground truth datasets from SNAP-processed imagery and validated in QGIS.

### 3.3 Flood Area Prediction

Now, after the model is trained, the multispectral image of another site was used for the prediction of flooded and non-flooded areas.

#### *3.3.1 Flood Extent Map of Prediction Area*

The flood extend map of the new multispectral image of Bangladesh was obtained from the above-trained model which is shown in below Figure 6, which shows non-flooded areas as “0” with the color “Cyan blue” and flooded areas as 0.998 with the color “Dark blue”.

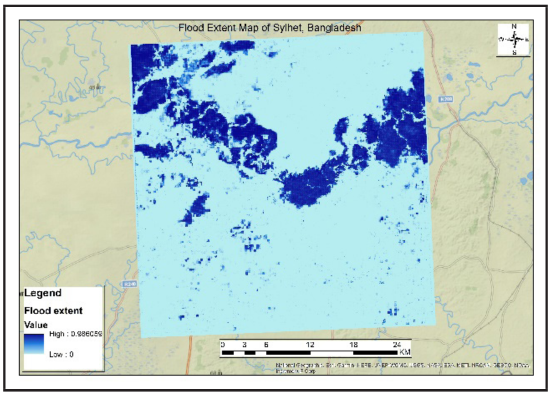


*Figure 5: Flood extent map*

### *3.3.2 Output Overlay with Prediction Image*

The flooded areas as predicted from the model trained were overlaid with the prediction image which is shown in below Figure 6.

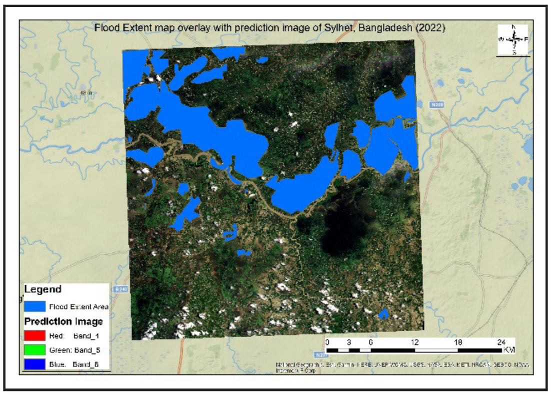


*Figure 6: Predicted Flooded areas of Sylhet, Bangladesh*

## 4 CONCLUSION

The devastating floods in Uttar Pradesh in August 2022 underscored the urgent need for advanced, rapid-response flood monitoring systems. This study addressed this challenge by developing a novel framework that uses satellite remote sensing with deep learning, especially Sentinel-1 SAR and Sentinel-2 multispectral data processed through a U-Net model. The U-Net model, trained on a dataset of 28,896 image patches, achieved an accuracy of 92.82% and an AUC score of 90%, indicating strong performance in identifying flooded areas. This approach not only achieved high accuracy in flood detection but also demonstrated scalability by successfully mapping floods in a new region (Bangladesh), proving its potential for global applicability.

The results showed that the model could effectively segment and classify flooded areas, providing high-resolution flood maps that were validated against ground truth data. The integration of remote sensing and deep learning techniques proved to be a powerful tool for flood monitoring, offering timely and accurate information for disaster response and management. The study highlights the potential of AI-driven approaches in improving flood risk assessment and mitigation strategies.

## 5 RECOMMENDATION

The study's findings highlight the potential of integrating satellite imagery with deep learning models for effective flood detection and mapping. However, several areas for improvement and future work are recommended to enhance the model's performance and applicability. Firstly, if we can perform a comparative analysis in order to validate the predicted areas with that of actual flooded areas such as field data and other reliable secondary sources, the prediction through this study would be more reliable. Additionally, optimizing the binarization threshold used for flood detection could improve precision and recall, potentially leading to better overall performance. Incorporating additional data sources, such as LiDAR, weather data, or hydrological models, could further enhance the model's accuracy and provide a more comprehensive understanding of flood dynamics.

## REFERENCES

Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. *IEEE Intelligent Systems*, *24*(2), 8–12. https://doi.org/10.1109/MIS.2009.36

Ian, G., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org

IMD. (n.d.). *Flood Situation in Uttar Pradesh*. 2022. Retrieved March 4, 2025, from https://mausam.imd.gov.in/

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, *7*(1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003

Mitchell, T. M. (1999). Machine learning. In *Software Testing, Verification and Reliability* (Vol. 9, Issue 3). McGraw-Hill Science/Engineering/Math. https://doi.org/10.1002/(SICI)1099-1689(199909)9:3<191::AID-STVR184>3.0.CO;2-E

Mohammed, S., Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., Naumann, F., & Harmouch, H. (2022). *The Effects of Data Quality on Machine Learning Performance*. *1*(1). http://arxiv.org/abs/2207.14529

NASA. (n.d.). *Satellite observations for disaster response and resilience. NASA Earth Science Division*. 2020. https://appliedsciences.nasa.gov/

Panda, S., Rao, M., Thenkabail, P., & Fitzerald. (2015). Remote Sensing Systems. In *Platforms and Sensors: Aerial, Satellites, UAVs, Optical, Radar, and LiDAR*.

Pech-May, F., Sanchez-Hernández, J. V., López-Gómez, L. A., Magaña-Govea, J., & Mil-Chontal, E. M. (2023). Flooded Areas Detection through SAR Images and U-NET Deep Learning Model. *Computación y Sistemas*, *27*(2). https://doi.org/10.13053/cys-27-2-4624

ReliefWeb. (2022). *Floods in Uttar Pradesh displace thousands*. https://reliefweb.int/

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Sanyal, A., Chatterji, V., Vyas, N., Epstein, B., Demir, N., & Corletti, A. (2021). *Fix your Models by Fixing your Datasets*. *NeurIPS*. http://arxiv.org/abs/2112.07844

Zhao, Y., Chen, J., & Oymak, S. (2020). *On the Role of Dataset Quality and Heterogeneity in Model Confidence*. 1–25. http://arxiv.org/abs/2002.09831

## Author's Information


| | | |
|---|---|---|
| Name | : | Binita Shahi |
| Academic Qualification | : | Bachelor's in Geomatics Engineering |
| Organization | : | Land Management Training Center |
| Current Designation | : | Instructor |
| Work Experience | : | 5 years |
| Published paper/article | : | 2 |