Deep learning-based segmentation of multi-temporal satellite imagery for flood detection

Vita Kashtan[*,†] and Volodymyr Hnatushenko[†]

Dnipro University of Technology, Dmytra Yavornytskoho Ave 19, Dnipro, 49005, Ukraine

Abstract

A deep learning-based pixel-based flood zone segmentation approach is proposed using multi-temporal satellite images and topographic and hydrological information. It is proposed to combine heterogeneous data (satellite images before and after the flood, digital elevation model, and hydrographic characteristics) into a single input tensor, allowing the neural network to consider the area's spatial and temporal dynamics and morphometric features. The architecture of the model ensures the preservation of the spatial detail of the flooded area through skip-connection mechanisms, which contributes to the correct identification of flood boundaries. Comparative analysis with FCNN, DeepLabv3, and BASNet confirmed the superiority of the proposed approach (F1-score 82%, Dice 82% for the category 'flooded areas'), which indicates its effectiveness for accurately detecting flooded areas.

Keywords

deep learning, pixel segmentation, flooding, multi-temporal satellite imagery, neural network, classification

1. Introduction

Floods are among the most devastating natural disasters, causing severe damage to infrastructure, human casualties, and significant economic losses. Climate change, the intensification of extreme weather events, and the expansion of urban areas further increase the vulnerability of regions to such hazards. Timely and accurate flood detection, along with continuous monitoring of their progression, are crucial components of effective emergency response, evacuation planning, resource allocation, and risk reduction for affected populations.

Traditional flood detection methods are typically based on ground observations or hydrological models, which come with several limitations, including a high dependency on the quality of input data, labor-intensive processes, delays in obtaining results, and challenges in scaling to large areas. This issue is particularly critical in urban areas, where complex terrain morphology, shallow and temporary flooding, and other water bodies significantly complicate flood detection [1]. While high-precision hydrological models can be effective in limited areas, their application at the community scale is constrained by the need for substantial computational resources.

Satellite imagery has become a vital resource for flood monitoring due to its ability to capture data over large areas with detailed spatial representation rapidly. However, flood zone segmentation remains challenging, particularly in urban environments [2].

Firstly, urban areas are characterized by high structural complexity and numerous water channels and drainage systems, often very narrow and sometimes less than one meter wide.

ISW-2025: Intelligent Systems Workshop at 9th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2025), May 15–16, 2025, Kharkiv, Ukraine

[*] Corresponding author.

[†] These authors contributed equally.

vitalionkaa@gmail.com (V. Kashtan); vvgnat@ukr.net (V. Hnatushenko)

0000-0002-0395-5895 (V. Kashtan); 0000-0003-3140-3788 (V. Hnatushenko)

Secondly, urban flooding is typically shallow and short-lived, making it difficult to detect and monitor using satellite-based methods. Thirdly, permanent water bodies, such as ponds or reservoirs, create difficulties in distinguishing temporary flood zones, especially under conditions of dynamic change. Given the complexity of urban terrain, the transient nature of flooding, and the presence of permanent water features, flood detection requires highly effective automated satellite image analysis methods capable of processing large volumes of data and accounting for the spatiotemporal variability of flood events.

2. Related works

Detecting flooded areas through precisely delineating water bodies using segmentation methods is a key component of satellite-based flood monitoring systems. This approach enables rapid emergency response and effective risk management by providing timely information, thereby reducing threats to the population. Additionally, the resulting data are critically important for spatial planning, particularly regarding land use [3] and the development of resilient infrastructure, considering the need to protect critical facilities such as hydroelectric power stations from potential flooding impacts. In turn, it contributes to minimizing socio-economic losses [4, 5].

Deep learning methods utilize multilayer neural networks to identify patterns in data, making them particularly promising for analyzing complex satellite imagery. Recurrent neural networks (RNN) are widely used to analyze water bodies and land cover using Sentinel imagery [6, 7]. In particular, studies [8, 9] have proposed approaches incorporating recursive and convolutional operations for effective spatiotemporal data processing. The authors of [10] presented a method based on convolutional neural networks (CNN) for rapid flood mapping using Sentinel-1 SAR imagery. This approach reduced map production time by 80% and enabled accurate monitoring under various conditions. Fully convolutional networks (FCN) have outperformed traditional superpixel-based segmentation methods, and their performance has been further enhanced by the use of Conditional Random Fields [11].

In the study [12], convolutional and recurrent neural networks were used to predict the likelihood of flash floods in Golestan Province, Iran. CNN models achieved higher accuracy due to using geospatial databases and the SWARA weighting method. The authors in [13] proposed a modified U-Net architecture called UFLOOD, which enables the prediction of two-dimensional water depth maps during urban flooding. The model utilizes hyetographic and topographic data to generate fast and accurate forecasts. Another approach, described in [14], involves using CNN models such as YOLOv3 and Fast R-CNN to detect flood indicators through integrated computer vision systems. The method incorporates edge detection and the analysis of objects' geometric parameters for real-time flood monitoring. The study [15] focuses on accurately identifying flooded areas using a fully convolutional network based on dual patches, which leverages deep learning-based feature fusion. FCNs are independently trained on synthetic aperture radar and multispectral images, enabling them to capture distinctive features combined to enhance flood detection capabilities.

Despite significant advances in applying deep learning for flood detection, current image segmentation methods still face several limitations. While techniques such as convolutional neural networks and fully convolutional networks achieve high accuracy, they demand substantial computational resources. This can slow down processing of large data volumes in real-time, which limits their effectiveness in emergency response scenarios where timely information is crucial. Furthermore, these models often require large annotated datasets for effective training. The limited availability of multi-temporal satellite imagery with corresponding expert annotations also reduces the models' generalization ability and performance across different geographic regions and flood types.

This study aims to develop a deep learning-based approach for the segmentation of multi-temporal satellite imagery to improve the accuracy of flood zone detection, enhance early warning systems, optimize risk management, and support more rapid emergency response.

3. Deep learning-based image segmentation

The proposed approach for detecting flooded areas is based on integrating multi-temporal satellite imagery and additional geospatial data using deep neural network architecture. The overall structure of the proposed approach is shown in Figure 1.

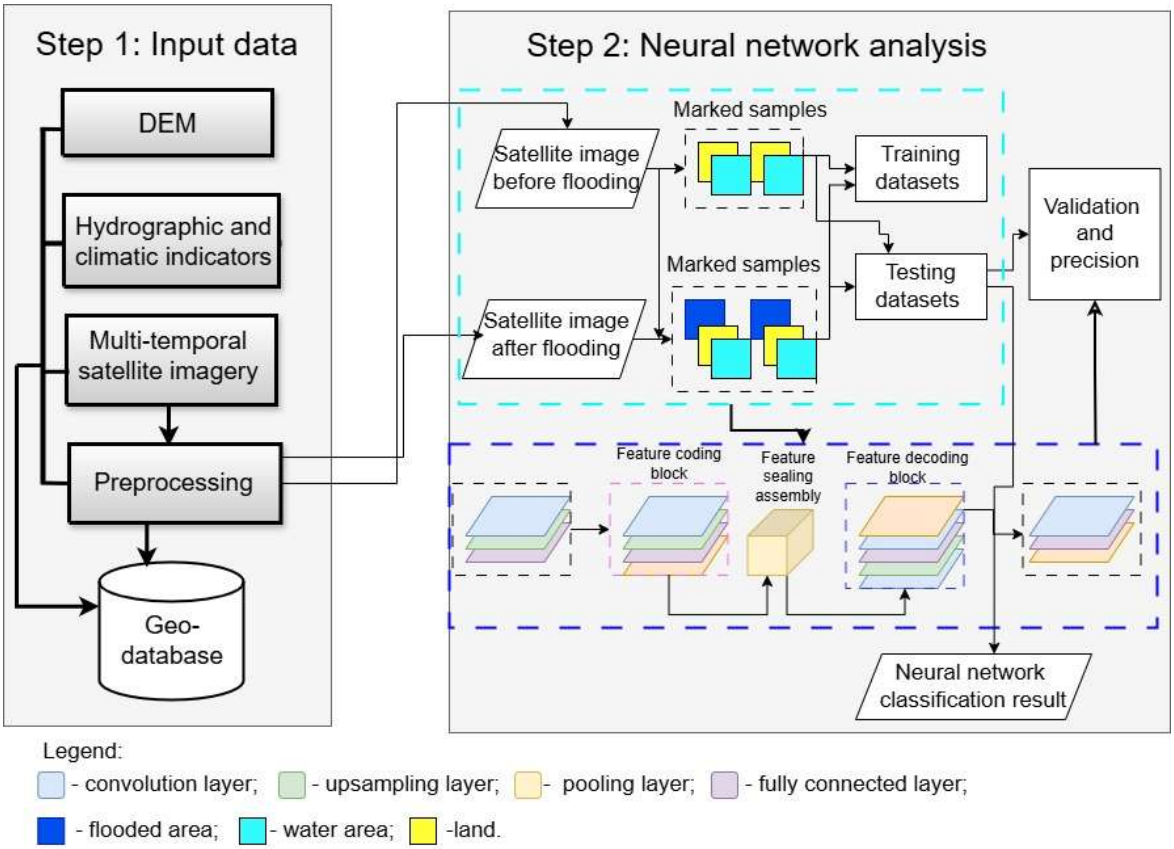


Figure 1: Diagram of the proposed approach.

In the first step (Input data), the collection and preprocessing of all necessary geospatial data occur, which will serve as input layers for the neural network. The data are stored and managed in a Geo-database, ensuring their integration and accessibility. The Digital Elevation Model (DEM) is the initial step for flood zone analysis. The DEM represents a three-dimensional digital model of the Earth's surface, including elevations over a specific area. This model is used to calculate water flows and identify potential flood zones. The DEM is obtained from relevant sources such as the USGS or platforms providing LIDAR data. DEM processing includes data smoothing, error correction, and gap filling to represent the terrain accurately. Terrain analysis calculates slopes, flow directions, and other characteristics for modeling water flow and flood zones. All elevation values in this model are stored in meters relative to the WGS84 EGM96 geoid. This geoid is based on the WGS84 ellipsoid, with coefficients computed from a global database of 30-minute mean free-air gravity anomalies and data obtained from satellites and direct altimetry measurements.

Loading of climatic indicators includes data on precipitation, temperature, humidity, wind speed, and other climatic factors that may influence the likelihood and extent of floods. These data are usually obtained from meteorological stations or climate models. For the study in the Dubai region, data from four meteorological stations were selected, significantly contributing to the flooding analysis in this area. These stations include Dubai, Sharjah, Al Ain, and Jebel Ali, as precipitation data from these stations are critical for regional flood modeling. Hydrological indicators include information about rivers, lakes, reservoirs, and other water bodies. They contain data on water levels, flow velocity, water volume, and other characteristics of flood dynamics.

Optical images from Sentinel-2 or Landsat-8,9 satellites before and after the flood are downloaded from the Copernicus Open Access Hub platform [16]. After downloading, the images undergo preprocessing, including radiometric and atmospheric corrections. These corrections eliminate distortions such as atmospheric effects, solar illumination, and other factors during image capture and transmission from space. Radiometric correction ensures that the data correspond to real physical quantities by converting pixels' digital numbers (DN) into reflectance values. This process includes sensor effect correction and instrument calibration [17]:

$$Reflective\ capacity = \frac{L_{TOA}}{E_{sun} \cdot \cos(\theta) \cdot d^2}, \quad (1)$$

where L_{TOA} is the top-of-atmosphere radiance; E_{sun} is the spectral solar constant; θ is the solar zenith angle; d is the Earth-Sun distance in astronomical units.

Atmospheric correction removes the effects of light scattering and absorption in the atmosphere, significantly improving image quality for further analysis.

The study proposes using the Dark Object Subtraction (DOS) method, which assumes that the darkest objects in the image (usually water or shadows) have zero reflectance, and any non-zero value results from atmospheric scattering [17]:

$$Corrected\ value = DN - DN_{min}, \quad (2)$$

where DN_{min} is the minimum DN value in the image.

The next stage involves the creation of a structured geodatabase (GDB) within the QGIS environment, which ensures data integrity and efficient information storage. This process includes defining tables and fields and importing source raster data, climatic and hydrographic data, and digital elevation model data [18]. To maintain data integrity, logical relationships between tables are established. Within this stage, the data are prepared for further analysis of the relationships between descriptive features and target labels. After the geodatabase is created, the data are optimized and indexed to enhance performance and access speed, essential for accurately identifying flooded areas.

The second step focuses on the automated pixel-level segmentation of flooded areas using a deep neural network architecture. The process begins with forming labeled samples from satellite images captured before and after the flood. These samples consist of image patches sized 321x321 pixels. Each pixel within these patches is assigned to one of three classes according to expert labeling: water area, which includes permanent water bodies like rivers, lakes, and reservoirs that existed before the flood and are not its consequence; flooded area, denoting temporarily inundated regions which appeared due to the flooding event; and land, indicating areas not covered by water both before and after the flooding event. These labeled samples are strategically divided into training and testing datasets in a 70% to 30% ratio. The neural network is trained using the training dataset. Throughout the training process, the model updates its internal parameters (weights and biases) through backpropagation, employing the Stochastic Gradient Descent optimizer to reduce the loss function, which measures the discrepancy between the model's predictions and the reference annotations.

The U-Net architecture was selected to address the segmentation task due to its capability for high-precision pixel-level segmentation [19, 20], which is crucial for accurately identifying flooded areas. The input to the U-Net is formed from multichannel data, including satellite imagery captured before and after the flood and DEM and hydrographic indicators, as described in step 1. This integrated input tensor has a size of 321x321 pixels. The U-Net architecture comprises a feature encoding block, a feature sealing assembly, and a feature decoding block. The feature encoding block includes several levels, each employing successive two-layer convolutional operations. It enables the network to extract hierarchical features from the image, ranging from low-level features (such as edges and textures) to high-level features (such as spatial-spectral characteristics of flooded areas).

The input layer accepts images of size:

$$H \times W \times C, \tag{3}$$

where H is height; W is width; C is the number of bands.

The convolution layer applies filters to extract features:

$$Y_{i,j,k} = \sum_{m,n} X_{i+m,j+n} \cdot W_{m,n,k} + b_k, \tag{4}$$

where X is the input data; W is the filter weights; b is the offset; Y is the output features.

After each series of convolutional layers, a max-pooling operation is applied, which reduces the spatial resolution of the feature maps while preserving the most significant features. This process enables the model to effectively capture contextual details indicating flooded areas while reducing computational load.

Max-pooling reduces the dimensionality of the feature maps:

$$Y_{i,j,k} = maxX_{s(i,m),s(j,n),k}, \tag{5}$$

where s is the size of the summation window.

The smoothing layer converts multidimensional data into a one-dimensional vector.

The dense layer uses an activation function for training:

$$Y_i = f\left(\sum_j W_{ij} \cdot X_j + b_i\right), \tag{6}$$

where f is the activation function.

An intermediate layer is positioned between the encoding and decoding modules. This layer delivers a highly condensed yet informative encoding of the extracted features. It comprises two convolutional operations with five filters, enabling the model to emphasize the most significant and abstract features related to flooded areas before restoring spatial information. The Feature Decoding Block is responsible for gradually recovering the image's spatial resolution to its initial dimensions. It comprises a series of upsampling layers employing bilinear interpolation, followed by convolutional layers that expand the spatial size of the feature maps. The extracted features are merged at every decoding stage with matching high-resolution features transferred directly from the corresponding encoding stage through skip connections. This feedback mechanism allows the model to preserve the overall structure and fine details of flooded areas, which is crucial for accurate boundary detection [21].

In the final layer of the decoding block, a convolution with a single filter and the Softmax activation function is applied. This function transforms the output values into probabilities representing the likelihood that each pixel belongs to one of the three defined classes: "water area," "flooded area," or "land." As a result, the output of the U-Net generates a segmentation mask in which each image pixel is classified according to its most probable state. The neural network comprehensively analyzes spatial and temporal changes between pre- and post-flood satellite images while integrating topographic data (DEM) and hydrological/climatic factors. It learns to identify spectral and textural features that indicate the presence of temporary water (flooded area), distinguishing it from water features and land. The architecture with encoding/decoding blocks and skip connections enables the model to detect high-level contextual features (e.g., large flooded areas, their shape, and relation to topography) and fine details (e.g., narrow flooded streets or small inundated patches).

4. Experiment

The United Arab Emirates is a country located on the Arabian Peninsula, known for its arid desert climate. In April 2024, the country was affected by a robust system of slow-moving storms that led

to significant rainfall, exceeding the annual average within just a few days. It resulted in flash floods in the eastern regions, causing road inundation and disruptions to the transportation infrastructure. Landsat-8 satellite images taken before (Fig. 2a) and after (Fig. 2b) the flood were used to analyze the impact of the flood on the region. The pre-flood images, taken in early April 2024, were used to determine the baseline state of the area. Following the storms, the images acquired on April 19, 2024, enabled the assessment of the extent of flooding and its impact on infrastructure.

a) b)

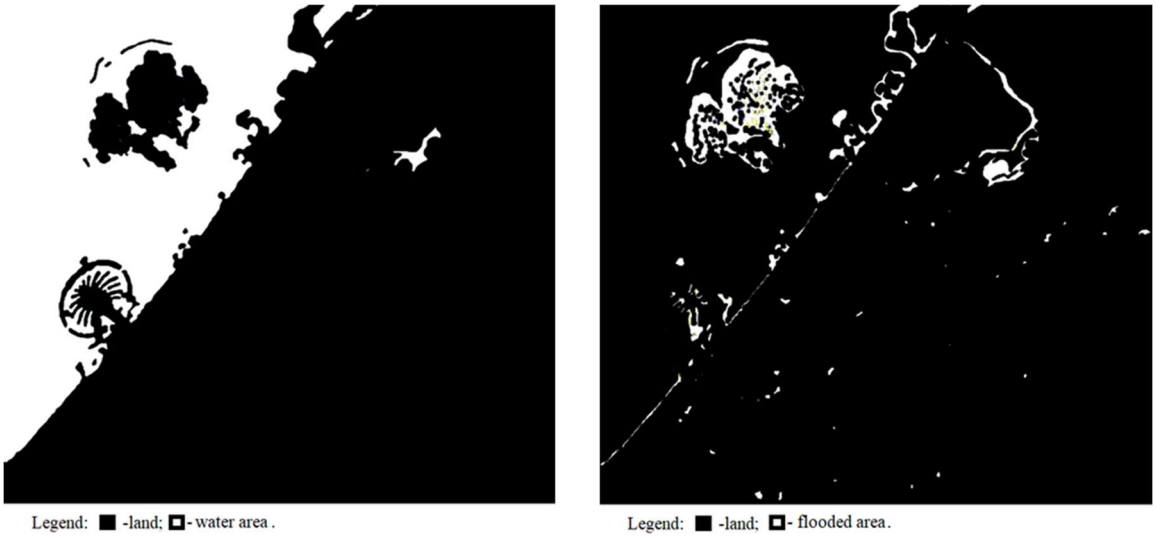
Figure 2: Satellite images from the Landsat 8-9 spacecraft synthesized into R-G-B channels: a) 18 March 2024; b) 19 April 2024.

This study conducted a quantitative assessment to evaluate the effectiveness of the proposed deep learning-based approach for pixel-wise segmentation of flooded areas. To train the model, specialized training samples of flooded zones were created using multi-temporal satellite images from Sentinel-2 and Landsat-8/9, following a preprocessing stage. The dataset comprises 1,043 images. Of these, 730 patches were allocated for training, while 313 patches were used for model testing. The data preparation process involved exporting the constructed training samples (pairs of "image + segmentation mask") in a format compatible with the selected neural network architecture. Key export parameters included the raster image from which the samples were derived; the size of each training patch, set to 256×256 pixels; the stride distance for the subsequent image, fixed at 321 pixels, which controls the degree of overlap when generating chips from large input images; the metadata format indicating the category of classified tiles; and the image format, chosen as TIFF to preserve high quality and multi-band information. Several hyperparameters were adjusted during the training of the U-Net model. The maximum number of training epochs ranged from 25, 50, and 75 to 100, allowing an investigation into the effect of training duration on model convergence and performance. The batch size, defining the number of samples processed simultaneously per iteration, was fixed at eight based on the available hardware resources. The input patch size fed into the neural network was 321×321 pixels. The training was conducted using the QGIS geospatial platform integrated with the PyTorch library. Model parameters were initialized with random values drawn from a standard normal distribution. The number of training samples used for each epoch configuration is provided in Table 1.

The output data from the training process included detailed information about the resulting model, such as the learning rate (which controls the magnitude of updates to the weight coefficients), the loss function values during both training and validation phases (indicating the model's fit to the data), and accuracy — the average proportion of correct predictions on the validation samples. Table 2 presents the evaluation metrics for the segmentation model's performance on the 'Flood area' class under various training configurations, including Precision, Recall, F1-score, mIoU, and Pixel Accuracy.

To visually assess the effectiveness of the proposed approach to satellite image segmentation, binary masks (Fig. 3) were obtained and formed by a neural network, which reflects spatial changes in the state of the territory before and after the flood event. In these masks, pixels corresponding to

water bodies (including permanent water areas) and flood zones are 1, while pixels of dry land are marked with 0.



a) b)

Figure 3: Binary mask by a neural network: a) before flooding; b) after flooding.

Figure 3a shows the initial state of the territory, where permanent water areas are highlighted in white and land areas are shown in black. This mask represents the initial state of the territory, where permanent water areas (such as rivers, lakes, etc.) exist before the flood event are highlighted in white. The black regions correspond to unflooded land surfaces. This image serves as a baseline for comparing and identifying newly flooded areas. Figure 3b displays two categories: unflooded areas (marked in black) and flooded areas (marked in white). Comparing this with Figure 3a allows visual identification of new areas covered by water due to the flood. The white regions in this mask indicate territories that were land before the event but became flooded afterward. Correspondingly, the black areas represent land that remained unflooded. It is important to note that this mask specifically emphasizes temporary flooded zones, distinguishing them from permanent water bodies. This distinction is achieved through the neural network's ability to analyze multi-temporal data and detect changes between the "before" and "after" states.

The results of the neural network classification of Landsat-9 satellite image pixels obtained using U-Net are presented in Figure 4 for a more detailed and multi-class analysis. This step demonstrates the outcome of pixel segmentation, where the model successfully distinguishes three categories: land, permanent water bodies, and temporarily flooded areas. Comparative analysis of Fig. 4a and Fig.4b highlights the effectiveness of the developed neural network model in detecting flood zones. The model accurately differentiates permanent water bodies from temporary flooding, critical for precise flood mapping. Using multi-temporal satellite images (before and after the event) combined with the U-Net architecture enables the identification of dynamic surface changes indicative of floodwater presence. The obtained results confirm the capability of the proposed approach to generate clear and informative masks for flood zone detection and that they can be applied for operational monitoring and impact assessment.

To quantitatively evaluate and compare the segmentation results of multi-temporal satellite images obtained using the proposed architecture, along with alternative architectures (FCNN, U-Net, DeepLabv3, and BASNet), the following metrics were employed [19, 22]: Pixel Accuracy (PA), Precision, F1-score, Recall, and mean Intersection over Union (mIoU). For a more detailed evaluation of the recognition accuracy for specific land cover classes (water bodies, flooded areas, and land), the Dice Similarity Coefficient was applied. This metric is commonly used in semantic segmentation tasks to quantify the overlap between the predicted mask and the ground truth label. The Dice coefficient ranges from 0 to 1, where 1 signifies complete mask overlap, and 0 indicates

no overlap. The metrics above were computed based on segmentation outcomes produced by the model on the test dataset. Generally, higher values of these metrics correspond to improved segmentation performance [22]:

$$Dice\ Coefficient = \frac{2 \times |A \cap B|}{|A| + |B|}, \tag{7}$$

where A is the set of pixels predicted by the model; B is the set of pixels in the reference mask; $|A \cap B|$ denotes the number of elements in both sets.

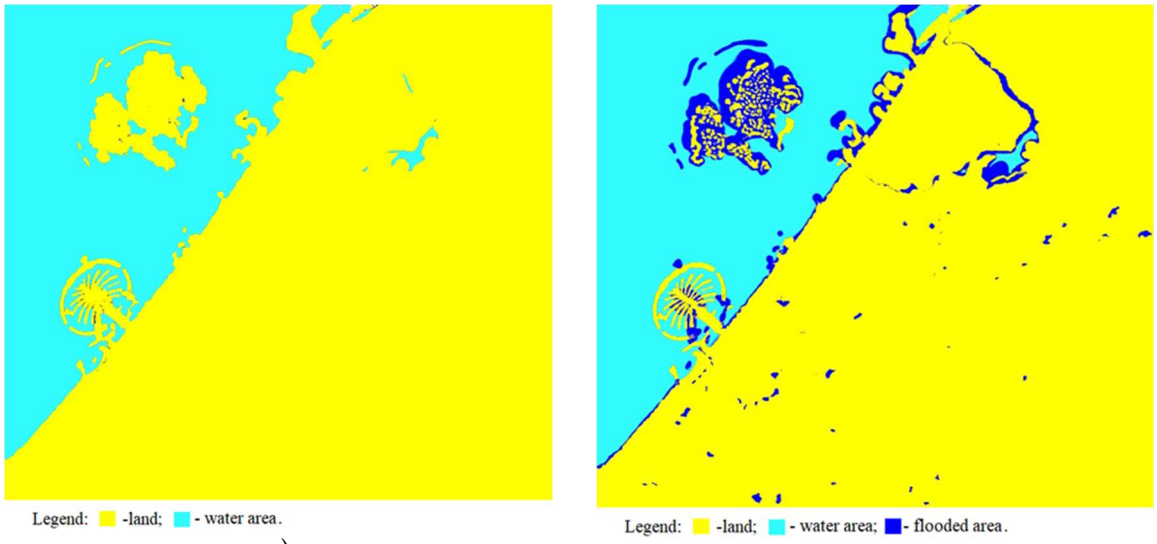


Figure 4: Neural network classification of Landsat 9 satellite imagery: a) before the flood on 18 March 2024; b) after the flood on 19 April 2024.

An essential step in ensuring reliable validation of segmentation results obtained using neural network models was the creation of ground truth masks. In this study, the authors manually generated these masks based on visual interpretation of satellite images captured before and after flooding. The process involved expert knowledge and specialized geographic information system tools. The resulting ground truth masks served as an objective "truth" for comparison with the model's predicted segmentation masks, allowing validation of its performance and assessment of the consistency between predictions and the actual state of the Earth's surface.

5. Results and Discussion

Table 1 shows the distribution of the number of samples used at each training stage. As the number of epochs increases, the training sample volume also grows, directly affecting the classification quality. When the number of epochs rises from 25 to 100, there is a significant increase in the number of samples used, resulting in a direct improvement in model accuracy. Specifically, at 25 epochs, the accuracy is only 82%, and the *F1 score* is 53%, indicating a low balance between classification correctness and the completeness of flood zone detection. The most balanced results are achieved at 50 and 75 epochs, with *F1 scores* of 78% and 73%, respectively. However, the highest values for all three metrics: accuracy (94%), recall (92%), and *F1 score* (93%) were recorded at 100 training epochs based on 1036 samples. These results highlight the importance of sufficient training data to achieve high classification quality and demonstrate a direct relationship between the number of training epochs, sample size, and model performance metrics.

Table 1
Evaluation of the performance of the proposed approach for flood zone segmentation by different learning parameters

Number of training epochs	Number of samples	Accuracy	Recall	F1 score
25	256	82	40	53
50	566	81	78	78
75	716	74	72	73
100	1036	94	92	93

The segmentation models' performance for flood zone detection was evaluated using a Landsat 9 satellite image. Five architectures were tested for comparative analysis: FCNN, U-Net, DeepLabv3, BASNet, and the proposed model. The summarized results are presented in Table 2. The U-Net model showed the lowest performance, with an *F1 score* of 62% and *mIoU* of only 45%, indicating a limited ability to delineate flood zone boundaries accurately. The FCNN model demonstrated better performance, but its *F1 score* remained at 66%, with a *mIoU* of 49.3%. DeepLabv3 achieved higher metrics, with an *F1 score* of 69.2% and a mIoU of 53%, reflecting improved capability for detailed segmentation. Even better results were obtained using the BASNet architecture, which achieved an F1 score of 78% and a mIoU of 64%, significantly outperforming the previous approaches. The proposed model reached the highest values across all metrics: precision of 89.5%, recall of 85.0%, *F1 score* of 82.0%, *mean IoU* of 69.5%, and *PA* of 92.8%.

Table 2
Metrics for evaluating the Landsat 9 satellite image segmentation model

Architectures	Precision (%)	Recall (%)	F1-score (%)	mIoU (%)	PA(%)
FCNN	72.5	68.1	66.0	49.3	85.2
U-Net	70.8	65.5	62.0	45.0	83.5
DeepLabv3	75.1	72.8	69.2	53.0	87.1
BASNet	80.5	76.8	78.0	64.0	90.3
Proposed	**89.5**	**85.0**	**82.0**	**69.5**	**92.8**

Table 3 shows the Dice coefficients used to evaluate the accuracy of image segmentation in three categories: water area, flooded area, and land. All models were assessed under the same experimental conditions, and the average Dice score served as a generalized metric of the effectiveness of each architecture. All architectures demonstrated the highest Dice coefficient values during the segmentation of the 'land' category due to the dominance of this class in the images and its stable spectral-spatial characteristics. Segmentation of flooded areas remains the most resource-intensive task, which is explained by their temporary nature, high spatial-temporal variability, and spectral similarity to permanent water bodies, complicating their correct identification. The U-Net model showed the lowest results for the flooded area (62.0%), while DeepLabv3 and BASNet demonstrated higher accuracy — 69.2% and 78.0%, respectively. The proposed model showed the best results for the 'Flooded Area' class at 82.0%, indicating its high ability to detect temporary water objects.

Table 3
Dice coefficients (%) for segmentation of different land cover classes

Architectures	Water area	Flooded area	Land	Average Dice
FCNN	88.2	66.0	92.5	82.2
U-Net	87.5	62.0	91.8	80.4

DeepLabv3	89.1	69.2	93.1	83.8
BASNet	90.5	78.0	94.2	87.6
Proposed	**91.8**	**82.0**	**95.5**	**89.8**

Figure 5 shows the flooded areas from April 15 to April 19, 2024. The graph indicates that the flooded area gradually increased from 10.2 hectares on April 15 to 42.61 hectares on April 19. This growth in the flooded area resulted from intensified meteorological phenomena, particularly significant precipitation, contributing to the worsening flood.

Figure 5: Change in Flooded in UAE, April 2024.

6. Conclusions

The study developed and experimentally evaluated an approach for pixel-wise segmenting flooded areas based on deep learning using multi-temporal satellite imagery. A unified model framework was proposed that integrates heterogeneous input data — pre- and post-flood satellite images, a digital elevation model (DEM), and hydrographic features. This integration enhanced the model's ability to accurately classify temporary water bodies, especially for the challenging "flooded area" class, which is difficult to detect due to its temporal variability and visual similarity to permanent water bodies. A comparative analysis of the performance of various deep learning architectures, including FCNN, U-Net, DeepLabv3, BASNet, and the proposed model, was conducted. Quantitative assessment of segmentation effectiveness on satellite images, performed using metrics (Precision, Recall, F1-score, mIoU, PA, and Dice), demonstrated the advantage of the proposed approach in semantic segmentation tasks, particularly for the class "flooded area," which is the most challenging to recognize due to its temporal variability and visual similarity to other water bodies (Precision 89.5%, Recall 85.0%, F1-score 82.0%, mIoU 69.5%, and PA 92.8%). The proposed model achieved the highest precision, recall, and consistency values in delineating flood boundaries, as confirmed by the Dice coefficient of 82% for the "flooded area" class and an average of 89.8%. Creating ground truth masks based on visual interpretation of multi-temporal satellite images and expert annotation enabled objective validation of model performance. Visual analysis of the segmentation results demonstrated a high level of spatial agreement between predicted masks and reference annotations. The model effectively identifies different land cover categories, including land, permanent water bodies, and newly formed flooded areas, which is crucial for operational flood event mapping. The proposed flood zone detection approach can be used for further analysis, flood risk management strategy development, public information dissemination, and infrastructure planning.

Acknowledgements

The study was conducted as part of the international educational project “Safe Artificial Intelligence: The European Legal Dimension” [101176092, a joint project of Dnipro University of Technology, Erasmus+ Jean Monnet Foundation, and the European Education and Culture Executive Agency (EACEA)]. Support from the European Commission for the publication of this work does not imply endorsement of its content, which solely reflects the views and opinions of the authors, and the Commission cannot be held responsible for any use that may be made of the information contained therein. 

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] Shih-Chun, Hsiao, Wen-Son, Chiang, Jiun-Huei, Jang, Han-Lun, Wu, Wei-Shiun, Lu, Wei-Bo, Chen, Yun-Ta, Wu. "Flood risk influenced by the compound effect of storm surge and rainfall under climate change for low-lying coastal areas." Sci. Total Environ, 2021, volume 764. 144439. doi:10.1016/j.scitotenv.2020.144439.

[2] Hartanti, Bella, Krisnadhi, Adila, Rahadianti, Laksmita, Susanti, Wiwiek, Shomim, Achmad. "Flood-MATE: A Flood Segmentation Model in Urban Regions through Adaptation of Mean Teacher and Ensemble Approach." IET Image Processing, 2025. doi:19. 10.1049/ipr2.70023.

[3] V. Kashtan, V. Hnatushenko, S. Zhir, Information Technology Analysis of Satellite Data for Land Irrigation Monitoring: Invited Paper. 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), 2021, pp. 1-4, doi: 10.1109/UkrMiCo52950.2021.9716592.

[4] J. Jang, M. Park, N. Lee, J. Lee, D. Yang, Real-time prediction of urban inundation based on SWMM 1D-1D model. Journal of the Korean Society of Hazard Mitigation. 2020, 401–411. doi: 10.9798/KOSHAM.2020.20.1.401

[5] V. Kashtan, D. Ivanov, V. Hnatushenko. Geoinformation Technology for Modeling and Mapping Flooding Territory in the Event of the Dnipro Hydroelectric Power Station Dam Failure. In: Babichev, S., Lytvynenko, V. (eds) Lecture Notes in Data Engineering, Computational Intelligence, and Decision-Making, Volume 1. Lecture Notes on Data Engineering and Communications Technologies, 27 December 2024, vol 219, pp 93–115. Springer, Cham. Print ISBN 978-3-031-70958-6, Online ISBN 978-3-031-70959-3. https://doi.org/10.1007/978-3-031-70959-3_5

[6] D. Ienco, R. Gaetano, R. Interdonato, K. Ose, D. Ho-Tong-Minh, "Combining Sentinel-1 and Sentinel-2 time series via RNN for object-based land cover classification." IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 2019, pp. 4881–4884. doi: 10.1109/IGAR SS.2019.8898458.

[7] V. Kashtan, V. Hnatushenko, I. Laktionov, H. Diachenko, Intelligent sentinel satellite image processing technology for land cover mapping. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, 2024, № 5, pp. 143–150. doi:10.33271/nvngu/2024-5/143.

[8] R. Marc, K. Marco, "Multi-temporal land cover classification with sequential recurrent encoders". ISPRS International Journal of Geo-Information, 2018, Vol. 7, No. 4, pp. 1–18. doi: 10.3390/ijgi7040129.

[9] M. Volpi, D. Tuia, "Dense semantic labeling of subdecimeter resolution images with convolutional neural networks." IEEE Transactions on Geoscience and Remote Sensing, 2017, Vol. 55, No. 2, pp. 881–893. doi: 10.1109/TGRS.2016.2616585.

[10] E. Nemni, J. Bullock, S. Belabbes, and L. Bromley, Fully Convolutional Neural Network for Rapid Flood Segmentation n Synthetic Aperture Radar Imagery, Remote Sensing, 2020, vol. 12, no. 16, p. 2532. doi:10.3390/rs12162532.
[11] S. Sarp, M. Kuzlu, Y. Zhao, M. Cetin, O. Guler, A comparison of deep learning algorithms on image data for detecting floodwater on roadways. Computer Science and Information Systems, 2022, 19(1), 397–414. doi: 10.2298/csis210313058s.
[12] M. Panahi et al., Deep learning neural networks for spatially explicit prediction of flash flood probability, Geoscience Frontiers, 2021, vol.12, no.3, p.101076, doi:10.1016/j.gsf.2020.09.007.
[13] R. Löwe, J. Böhm, D. Jensen, G. Leandro, J. Rasmussen, U-FLOOD – Topographic deep learning for predicting urban pluvial flood water depth. Journal of Hydrology, 2021, 603, 126898. doi: 10.1016/j.jhydrol.2021.126898.
[14] R. Pally, S. Samadi, Application of image processing and convolutional neural networks for flood image classification and semantic segmentation. Environmental Modelling & Software, 2022, 148, 105285. doi: 10.1016/j.envsoft.2021.105285.
[15] J.Emily, N. Sudha, DeepFlood: A deep learning based flood detection framework using feature-level fusion of multi-sensor remote sensing images. JUCS -Journal of Universal Computer Science, 2022, 28(3), pp. 329-343, doi: 28. 329-343. 10.3897/jucs.80734.
[16] Copernicus Open Access Hub. URL: https://www.sentinel-hub.com/explore/eobrowser/.
[17] R. Khanna, I. Sa, J. Nieto, R. Siegwart, "On field radiometric calibration for multispectral cameras." 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017, pp. 6503-6509. doi:10.1109/ICRA.2017.7989768.
[18] N. Notarangelo, C. Wirion, F. Winsen, STURM-Flood: a curated dataset for deep learning-based flood extent mapping leveraging Sentinel-1 and Sentinel-2 imagery. Big Earth Data, 2025, 1–27. doi:10.1080/20964471.2025.2458714.
[19] J. Yoo, J. Lee, S. Jeung, S. Jung, M. Kim Development of a Deep Learning-Based Flooding Region Segmentation Model for Recognizing Urban Flooding Situations. Sustainability. 2024, 16(24):11041. doi:10.3390/su162411041.
[20] V. Hnatushenko, O. Honcharov, Land Cover Mapping with Sentinel-2 Imagery Using Deep Learning Semantic Segmentation Models. CEUR Workshop Proceedings. Vol. 3909: Proc. of the XI International Scientific Conference "Information Technology and Implementation", 2024, pp1–18. https://ceur-ws.org/Vol-3909.
[21] Y. Sun, L. Lei, Z. Li, G. Kuang, Similarity and Dissimilarity Relationships Based Graphs for Multimodal Change Detection. ISPRS Journal of Photogrammetry and Remote Sensing, 2024, pp 70–88. doi:10.1016/j.isprsjprs.2024.01.002.
[22] S. Mou, T. Chowdhury, et al., AI-Driven Water Segmentation with deep learning models for Enhanced Flood Monitoring, 2025. doi:10.48550/arXiv.2501.08266.