Ingénierie des Systèmes d'Information
Vol. 31, No. 2, February, 2026, pp. 471-479
Journal homepage: http://iieta.org/journals/isi

A GeoAI-Based CNN Framework for Urban Flood Potential Mapping via Integration of Rainfall and Waste Data

Leni Novianti[1], Ermatita[2*], Abdiansah[2], Ade Silvia Handayani[3]

[1] Doctoral Program in Engineering, Faculty of Engineering, Universitas Sriwijaya, Palembang 30139, Indonesia
[2] Department of Computer Science, Faculty of Computer Science, Universitas Sriwijaya, Palembang 30139, Indonesia
[3] Management of Informatich Department, Politeknik Negeri Sriwijaya, Palembang 30139, Indonesia

Corresponding Author Email: ermatita@unsri.ac.id

https://doi.org/10.18280/isi.310215

Received: 6 November 2025
Revised: 10 January 2026
Accepted: 21 February 2026
Available online: 28 February 2026

Keywords:
GeoAI, convolutional neural network, flood potential mapping, CHIRPS rainfall data, Sentinel-2 imagery, waste accumulation

ABSTRACT

Urban flooding remains a recurrent challenge in lowland cities, driven by both hydrometeorological conditions and anthropogenic factors. This study proposes a GeoAI-based convolutional neural network (CNN) framework for urban flood potential mapping by integrating multi-source geospatial data, including CHIRPS rainfall information and Sentinel-2-derived waste accumulation indicators. The proposed approach incorporates spatial preprocessing, grid-based sampling, and temporal augmentation to construct a comprehensive spatiotemporal dataset covering the period from 2019 to 2025. A two-dimensional CNN model is developed to capture the spatial interactions between rainfall intensity and waste distribution. Experimental results demonstrate that the proposed framework achieves an overall accuracy of 0.73, with an F1-score of 0.73 for flooded areas and 0.69 for non-flooded areas. The model exhibits stable performance, with a mean F1-score of 0.7080 and an average AUC of 0.7469 under five-fold cross-validation. The generated flood potential maps indicate that areas characterized by high rainfall and concentrated waste accumulation are more susceptible to flooding. The findings highlight the importance of integrating anthropogenic factors into GeoAI-based flood modeling and provide a practical framework for data-driven urban flood monitoring and early warning systems.

1. INTRODUCTION

Flooding is one of the most frequent hydrometeorological disasters in many regions of Indonesia, including Palembang City. This phenomenon is triggered by a combination of natural factors and human activities, such as extreme rainfall, increased river discharge, river narrowing due to sedimentation, drainage damage, and the accumulation of waste along river basins [1]. As a lowland city with an extensive river network particularly the Musi River and its 21 tributaries Palembang has a high level of flood vulnerability [2]. Data from the Meteorology, Climatology, and Geophysics Agency indicate that rainfall in South Sumatra can reach 200–400 mm during the peak of the rainy season, which is above the average for other urban areas [3].

In addition to hydrological factors, one of the significant causes of flooding in Palembang is the accumulation of solid waste [4-10]. According to reports from the Environmental and Forestry Agency and the Sumatra VIII River Basin Authority, Palembang City generates approximately 1,200 tons of waste per day, and nearly 50% of the waste polluting the rivers originates from household sources. This accumulation reduces the water flow capacity, contributes to sedimentation, and clogs drainage channels and rivers, thereby exacerbating flooding events, especially during periods of high rainfall intensity [11-15].

The Palembang City Government, through the Public Works and Spatial Planning Agency, has implemented various flood mitigation measures, including river normalization, construction of drainage outlets, installation of water level monitoring boards (peil scales), establishment of river-care communities, and regulation of illegal buildings along riverbanks. However, these efforts have not yet been able to sustainably address flood problems [16, 17]. One of the main constraints is the limitation of flood monitoring systems that are still manual, non-integrated, non real-time, and have not yet utilized artificial intelligence technologies for predictive analysis [18].

Along with the advancement of digital technology, developments in geospatial technology and artificial intelligence have opened new opportunities for faster and more accurate flood mapping, monitoring, and prediction. The Geospatial Artificial Intelligence (GeoAI) approach, which integrates geospatial data with machine learning and deep learning algorithms, enables efficient processing of large-scale spatial data [19]. Through platforms such as Google Earth Engine (GEE), various data sources including daily rainfall data (CHIRPS), Sentinel-2 satellite imagery, multispectral indices (NDVI, MNDWI, NDBI), and geomorphological information can be comprehensively analyzed [20].

In the context of flood modeling, Convolutional Neural Networks (CNNs) have been widely applied due to their ability to extract spatial patterns from satellite imagery, such as inundated areas, land cover changes, and rainfall distribution [21]. Numerous previous studies have demonstrated that CNNs are effective for flood mapping and detection using remote sensing and hydrometeorological data. However, most of these studies primarily focus on hydrological factors and physical land surface characteristics, while anthropogenic factors particularly waste accumulation have not been explicitly integrated as input variables in flood prediction models [22].

On the other hand, several studies have addressed waste detection and mapping using satellite imagery and deep learning approaches. Nevertheless, these studies are generally conducted independently and are not directly linked to flood risk analysis. Consequently, waste data are still treated as separate information rather than as a critical component in GeoAI-based flood prediction systems, especially in urban areas with drainage systems that are highly susceptible to blockage, such as Palembang City.

Based on this research gap, this study proposes a GeoAI-based approach using a Convolutional Neural Network (CNN) by integrating rainfall data and waste accumulation data as multi-channel inputs. This integration aims to capture the spatial relationships between hydrological and anthropogenic factors in influencing urban flood potential. The utilization of Google Earth Engine enables efficient and consistent processing of multi-source data across broad spatial and temporal scales.

Therefore, this study not only aims to improve the accuracy of real-time and integrated flood monitoring and prediction but also contributes novelty by incorporating waste accumulation data as a key variable in GeoAI-based flood potential modeling. This approach is expected to support faster and more informed decision-making by local governments and to enhance the effectiveness of flood mitigation and early warning systems in Palembang City and other urban areas in Indonesia.

2. METHODOLOGY

The research methodology is structured to illustrate the workflow starting from data collection, the design of the 2DCNN model architecture, up to the evaluation stage and the generated outputs. The methodological diagram in Figure 1 presents the relationships among the main components of the literature study, data collection, model architecture design, evaluation measurement, and output. Each stage is interconnected systematically to produce an accurate flood potential prediction model that can be utilized for risk mapping in Palembang City.

Figure 1 illustrates the sequence of the research process, which begins with the collection of theoretical foundations and data as the primary inputs. The processed data are then utilized in the design of a wo-dimensional Convolutional Neural Network (2D CNN) model consisting of several convolutional layers to generate the required feature representations. The developed model is subsequently evaluated using various performance metrics to ensure its accuracy. The final output of all these stages is a spatial mapping of regional conditions, including the classification of Flooded Area Map (Class 1) and Non Flooded Area Map (Class 0).

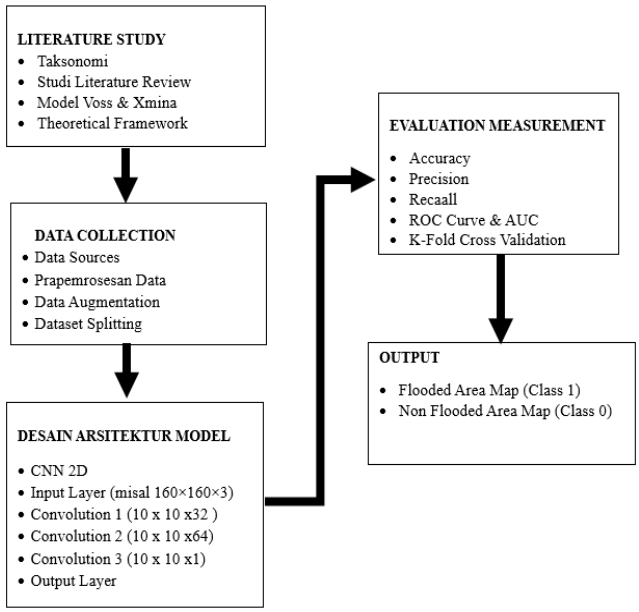

Figure 1. Research method

3. LITERATURE STUDY REVIEW

The literature study review in this research includes the mapping of the research taxonomy, an examination of previous studies, relevant analytical models, and the theoretical framework underlying the development of the 2D CNN–based flood prediction model. The research taxonomy positions this study within the domain of geospatial deep learning, which integrates CHIRPS rainfall data, Sentinel-2 imagery, and waste accumulation indicators as key variables in flood risk analysis. The literature review shows that the use of CNN and GeoAI has been widely applied for flood detection and satellite image classification; however, the integration of waste accumulation as a crucial factor in urban contexts remains very limited, representing a significant research gap. The Voss and Xmina models are used as theoretical references to explain the spatial relationships among variables and the mechanisms of multidimensional feature extraction, serving as the foundation for designing the CNN input–output architecture. Overall, the theoretical framework of this study is constructed based on the concepts of flooding and its triggering factors, GeoAI theory, grid-based spatial representation, and the operational principles of 2D CNNs, thereby providing a strong methodological foundation for building a flood potential classification model in urban environments.

4. DATA COLLECTION

The data collection stage was carried out to obtain all the data required for the 2D CNN modeling process used to monitor flood potential in Palembang City. This process consists of four main steps: data sources, data preprocessing, data augmentation, and dataset splitting. The workflow has been aligned with the research diagram employed in this study.

4.1 Data sources

The data collection stage was conducted to obtain all the

information required for the 2D CNN modeling process used to monitor flood potential in Palembang City. This stage consists of four main components: data acquisition, data preprocessing, data augmentation, and dataset splitting. Each component is systematically aligned with the research workflow outlined in the methodological framework of this study.

4.1.1 Rainfall (CHIRPS – GEE)

Rainfall data were obtained through the GEE platform using the CHIRPS dataset (Climate Hazards Group InfraRed Precipitation with Station Data).

The characteristics of the CHIRPS data are as follows:

- Spatial resolution: approximately 5 km
- Temporal resolution: daily
- Data unit: millimeters (mm)
- Time coverage: 6 years
- Extracted specifically for the Palembang City area

This rainfall data is used as one of the key variables in determining flood occurrences, as illustrated in the rainfall data figure.

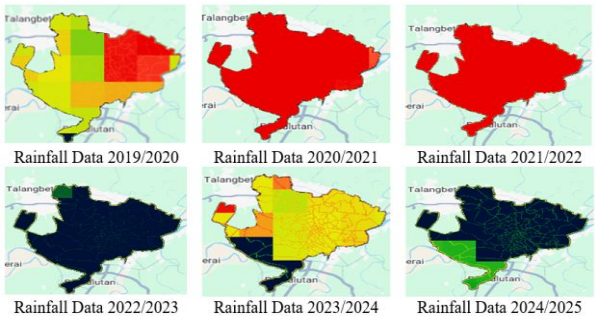

Figure 2. Rainfall for the period 2019 to 2025

Figure 2 presents a visualization of the CHIRPS–GEE rainfall data for the period 2020 to 2025 within the study area. Each map illustrates rainfall intensity using a color gradient, where red indicates high rainfall, green represents moderate rainfall, yellow signifies low rainfall, and blue denotes dry or rainless conditions. The visualization shows clear variations in rainfall intensity across the years, with certain years experiencing significantly higher rainfall while others are dominated by drier conditions. This information is essential for analyzing climatological trends, understanding the dynamics of the rainy and dry seasons, and examining how year-to-year changes in rainfall intensity may influence flood occurrences in the region.

4.1.2 Waste accumulation data (Sentinel-2 MSI)

Waste accumulation data were extracted from Sentinel-2 satellite imagery using a combination of RGB bands and infrared bands. The identification of waste was conducted through:

- Band 4, 3, 2 (True Color)
- Band NIR dan SWIR
- Supporting spectral indices such as NDVI, NDWI, and MNDWI
- Visual detection of flooded or polluted areas that indicate waste accumulation

Sentinel-2 imagery was selected because it provides a 10-meter spatial resolution, allowing detailed visualization of surface conditions in urban areas. Figure 3 presents the waste accumulation map.

Figure 3. Waste accumulation for the period 2019 to 2025

Figure 3 presents the spatial conditions of the study area, emphasizing waste accumulation patterns from 2019 to 2025. Red-colored features represent rivers and watershed networks that function as the primary water flow pathways. Black areas denote the spatial distribution of identified waste deposits, including those located within inundated zones. The visualization clearly demonstrates a strong spatial association between river networks and the concentration of waste accumulation in adjacent areas. Meanwhile, Figure 4 below presents the spatial distribution of urban waste accumulation across the study area over six consecutive observation periods, ranging from 2019/2020 to 2024/2025.

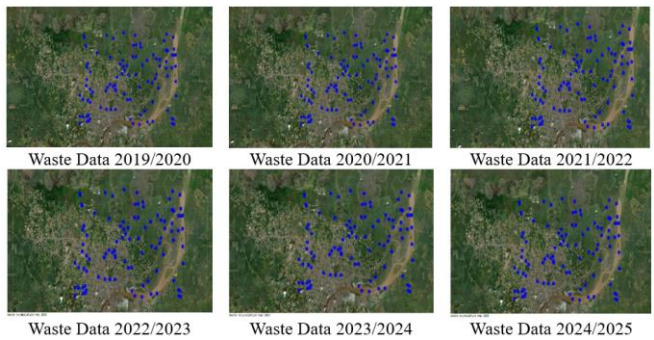

Figure 4. Weste data period 2019/2020 to 2024/2025

Figure 4 presents the spatial distribution of accumulated urban waste within the study area across six consecutive observation periods, spanning from 2019/2020 to 2024/2025. The blue points on each map represent locations of waste accumulation identified through geospatial data processing and satellite imagery analysis, thereby illustrating the temporal and spatial dynamics of waste distribution over time.The observed distribution patterns indicate a clear tendency for waste accumulation to concentrate in residential areas and along river corridors, which potentially reduces flow capacity and impairs urban drainage systems. This visualization provides an essential basis for understanding the role of anthropogenic factors in increasing urban flood vulnerability and supports further analysis of the relationships among waste accumulation, environmental conditions, and flood potential in surrounding areas.

The quantitative amount of waste derived from Figure 4 is summarized in Table 1.

Tabel 1. Waste accumulation for the period 2019 to 2025

No.	Year	Waste Amount (Tons)
1	2019	359,970,216
2	2020	402,944,043
3	2021	294,076,651
4	2022	223,393,678
5	2023	245,877,945
6	2024	177,739,349
7	2025	245,493,169

Table 1 presents the annual waste accumulation from 2019 to 2025, which shows significant fluctuations over the years. The volume of waste increased from 359,970,216 tons in 2019 to its highest peak of 402,944,043 tons in 2020, followed by a sharp decline in 2021 and 2022, reaching 223,393,678 tons. Subsequently, the amount rose again in 2023, then decreased to its lowest point of 177,739,349 tons in 2024, before increasing once more to 245,493,169 tons in 2025. These fluctuations reflect the dynamics of waste production, influenced by various factors such as changes in community activities, environmental conditions, and the presence of surface water pooling and river flow patterns that affect the movement and accumulation of waste in certain areas.

4.2 Data augmentation

Data augmentation in this study was conducted to address the limited number of samples derived from annual datasets and to enhance the representation of spatial and temporal variability within the study area. The augmentation was not performed through synthetic image manipulation, but rather through increasing temporal resolution and applying grid-based spatial sampling while preserving the physical characteristics of the original data.

Temporal resolution enhancement was achieved by generating monthly composites from the multi-source datasets used in this study. Daily CHIRPS rainfall data were aggregated into monthly mean values for each spatial grid. Meanwhile, Sentinel-2 imagery was processed into monthly composites using median values to minimize the effects of clouds, shadows, and atmospheric disturbances. Indicators of waste accumulation were extracted on a monthly basis based on changes in spectral characteristics within river corridors and surrounding areas, enabling a more detailed observation of waste accumulation dynamics over time. Using this approach, data from the period 2019–2025 yielded a total of 72 monthly observation periods.

In addition to temporal enhancement, data augmentation was also performed through grid-based spatial sampling. The administrative area of Palembang City was divided into 10 × 10 pixel grids, with each grid treated as an individual spatial sampling unit for each monthly period. Consequently, each month generated 100 spatial samples representing local rainfall conditions and waste accumulation. Through the combination of monthly compositing and spatial grid partitioning, the total number of samples used in this study reached 7.200 providing sufficient data variability for the 2D CNN model to learn the spatiotemporal relationships among rainfall, waste accumulation, and flood potential.

4.3 Data splitting

The augmented dataset was subsequently divided into training and testing datasets using a train–test split scheme with a proportion of 80% for training data and 20% for testing data. This partitioning was performed after all data had undergone preprocessing stages, including clipping based on the administrative boundaries of Palembang City, construction of 10 × 10 spatial grids, extraction of rainfall and waste accumulation features, and value normalization using the MinMaxScaler method.

The data splitting was conducted randomly while preserving the representation of spatial and temporal variability. As a result, both the training and testing datasets encompass conditions of high, moderate, and low rainfall, as well as varying levels of waste accumulation across the study area. With this approach, the testing dataset serves as a representation of real-world conditions that were not previously observed by the model during the training process.

This data partitioning scheme enables an objective evaluation of the 2D CNN model's performance in terms of its generalization capability for detecting flood potential based on multi-temporal geospatial data. The evaluation results subsequently provide the basis for assessing the effectiveness of integrating CHIRPS rainfall data and Sentinel-2-derived waste accumulation information within a GeoAI-based flood monitoring system.

5. 2D CONVOLUTIONAL NEURAL NETWORK (2D CNN) ARCHITECTURE

The 2D Convolutional Neural Network (2D-CNN) architecture used in this study was designed to learn spatial patterns from 10 × 10 grid data with two input channels, namely rainfall and waste accumulation. The model consists of several main layers, including convolutional layers, activation layers (ReLU), max-pooling layers, dense layers, and dropout, which collectively function to extract features, reduce dimensionality, and perform the final classification of flood potential. Figure 4 presents the overall structure of the model.

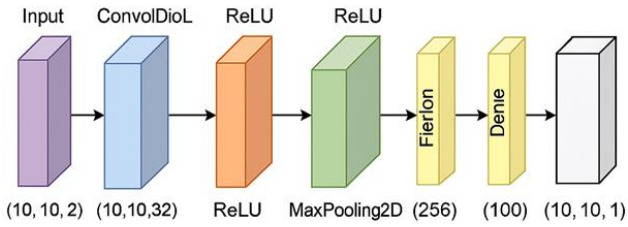

Figure 5. Architecture 2D CNN
Note: CNN = Convolutional Neural Network.

Figure 5 illustrates the 2D Convolutional Neural Network (2D-CNN) architecture used in this study. The model receives an input tensor of size 10 × 10 × 2, which is then processed through a convolutional layer consisting of 32 filters with a size of 3 × 3, producing an output of 10 × 10 × 32. After the ReLU activation and max-pooling operations, the feature map is reduced to 5 × 5 × 32 before being flattened into an 800-element vector. This vector is then passed through fully connected layers with 256 and 100 neurons, respectively, and subsequently reshaped back into a grid representation through a reshaping step and a final dense layer that generates an output of 10 × 10 × 1. The final dimensional transformation follows the sequence: (10 × 10 input → convolution → pooling → flatten (800) → dense → reshape → output 10 ×

10 × 1), where each cell value in the output grid represents the classified flood potential.

6. DETERMINING THE ACCURACY AND LOSS GRAPHS

At this stage, the performance of the 2D CNN model was evaluated by observing the progression of training accuracy and validation accuracy over 50 epochs. The accuracy graph is used to assess the stability of the learning process, the model's generalization ability to unseen data, and to detect indications of overfitting or underfitting. By monitoring the accuracy trends of both curves, it is possible to determine whether the model successfully learns relevant patterns from the normalized rainfall and waste accumulation data.

The accuracy graph in Figure 6 shows a consistent improvement in the model's performance throughout the training process. At the beginning of the epochs, the training accuracy starts at approximately 0.51 and gradually increases, reaching 0.67 by the 50th epoch. This upward trend indicates that the model is able to effectively extract important features from the two input data channels, namely the rainfall channel (CHIRPS) and the waste accumulation channel (Sentinel-2).

Meanwhile, the validation accuracy, which initially started at around 0.56, also shows a continuous improvement, reaching approximately 0.675. It can be observed that the validation accuracy curve tends to be slightly higher than the training accuracy curve across most epochs. This phenomenon indicates that the model does not experience overfitting, as its performance on unseen data remains equal to or better than its performance on the training data. This condition may be attributed to the relatively small dataset size, where the samples in the validation set exhibit simpler or more easily learnable distribution patterns for the model.

Overall, the graph demonstrates that the 2D CNN model successfully learns spatial representations from the two types of geospatial data in a stable manner, with a consistent upward trend in accuracy and without extreme fluctuations. Thus, the model is considered to generalize well to the validation data, although its performance can still be further improved through architectural optimization, additional data expansion, or adjustments to training parameters.

In Figure 6, the graph shows a progressive increase in both training and validation accuracy over 50 epochs. The training accuracy rises from approximately 0.51 to 0.67, while the validation accuracy increases from 0.56 to 0.675. The validation curve, which consistently remains higher than the training curve, indicates that the model is able to generalize well without any significant signs of overfitting. In contrast, the model loss graph provides insight into how the error values decrease throughout the training process.

In Figure 7, this loss graph illustrates the decrease in training and validation loss values over 50 epochs. At the beginning of training, the loss values were around 0.69, indicating that the model's initial performance was still close to random prediction. As the epochs progressed, both loss curves consistently and steadily declined without sharp fluctuations or irregular trends. The parallel downward pattern between the training loss and validation loss indicates that the learning process was effective and that no signs of overfitting were present, as the validation loss remained slightly lower than the training loss across most epochs. By the 50th epoch, the training loss decreased to approximately 0.60, while the validation loss reached around 0.595, demonstrating that the model is capable of generalizing well to unseen data. Overall, the graph confirms that the 2D CNN model learned stably and efficiently from the rainfall and waste accumulation data used in this study. The results of the confusion matrix generated by the 2D CNN model can be seen in Figure 8.

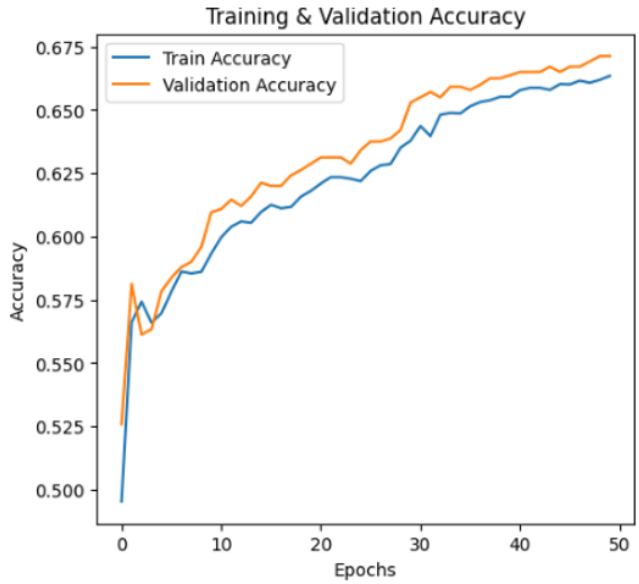

Figure 6. Accuracy Graph of the 2D CNN
Note: CNN = Convolutional Neural Network.

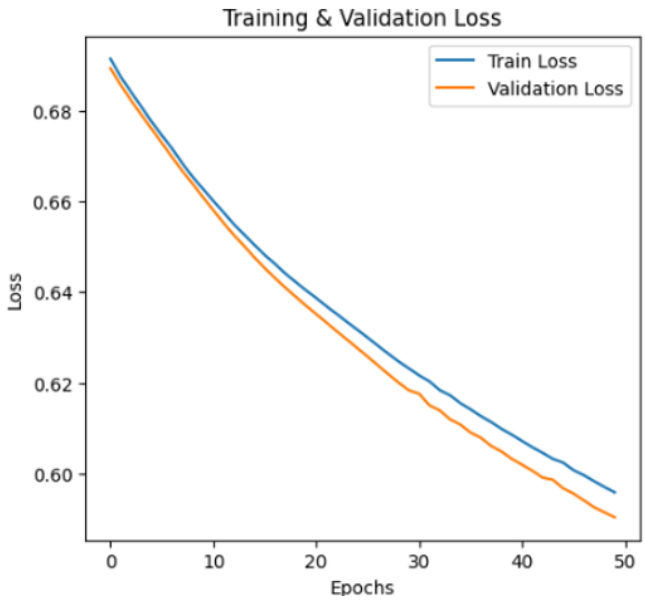

Figure 7. Loss Graph of the 2D CNN
Note: CNN = Convolutional Neural Network.

Based on Figure 8, the 2DCNN confusion matrix demonstrates a fairly good classification performance in distinguishing between flooded and non-flooded areas. The True Positive value (995) and True Negative value (659) indicate that the majority of areas were correctly predicted by the model. However, there are still misclassifications, including False Positives (491), which are non-flooded areas incorrectly predicted as flooded, and False Negatives (255), which are flooded areas that were not detected by the model. This condition suggests that the model is more sensitive in detecting flood occurrences, resulting in a relatively high

number of false alarms, yet it remains capable of identifying flood events with a dominant success rate. Overall, the confusion matrix shows that the model performs well, although improvements are still needed to enhance the accuracy of non-flood predictions.

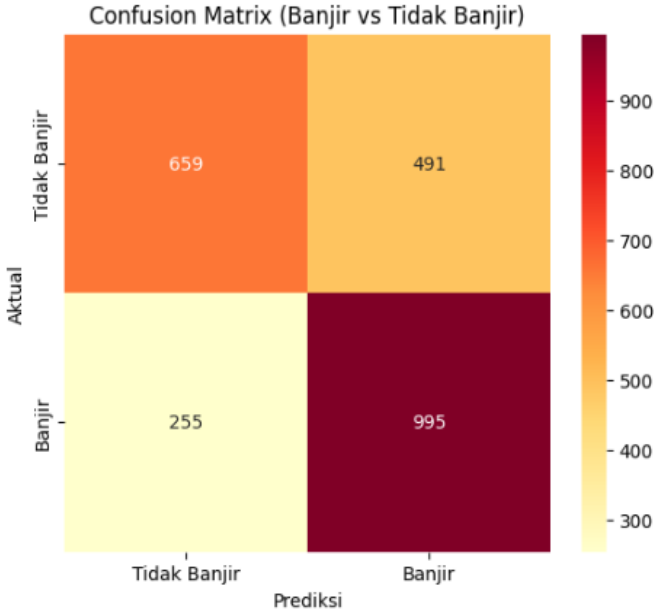

Figure 8. Confusion Martrix 2DCNN
Note: CNN = Convolutional Neural Network.

7. EVALUATION

Evaluation and performance measurement are essential stages in the development of a flood detection model, as they aim to assess the model's ability to accurately distinguish between flooded and non-flooded areas. In this study, the model's performance was evaluated using precision, recall, and F1-score metrics, which provide a comprehensive overview of prediction accuracy, the model's sensitivity in detecting flood events, and the balance between the two. The evaluation values presented in the table reflect the model's performance in classifying the two main classes Not Flooded and Flooded where the testing results yield the overall accuracy of the model.

Table 2. Evaluation 2D CNN

Class	Precision	Recall	F1-score
Not Flooded	0.72	0.69	0.69
Flooded	0.69	0.80	0.73
Accuracy			0.73
Macro Avg	0.76	0.71	0.70
Weighted Avg	0.73	0.71	0.73

Table 2, which presents the evaluation of the 2DCNN model, shows the model's performance in recognizing two classes: Not Flooded and Flooded. For the Not Flooded class, the model achieved a precision of 0.72 and a recall of 0.69, indicating that the model performs fairly well in predicting areas that are truly not flooded, although some misclassifications still occur in capturing all non-flooded samples. Meanwhile, the Flooded class obtained a precision of 0.69 and a higher recall of 0.80, meaning that the model is more sensitive in detecting flood events and is able to correctly identify most of the flooded samples, despite still producing some incorrect flood predictions. The F1-score values for both classes, 0.69 and 0.73 respectively, demonstrate a reasonable balance between precision and the model's ability to detect actual samples. Overall, the model achieved an accuracy of 0.73, indicating a reasonably good classification performance on the test dataset. The consistent macro average and weighted average values (with F1-scores of 0.70 and 0.73) further indicate that the model performs stably on an imbalanced dataset, which aligns with the confusion matrix results showing that the model tends to be more accurate in detecting flooded areas than non-flooded areas.

7.1 ROC curve and AUC analysis

To complement the evaluation based on discrete metrics, this study employs Receiver Operating Characteristic (ROC) analysis and Area Under the Curve (AUC) to assess the discriminative capability of the CNN model across different decision threshold values.

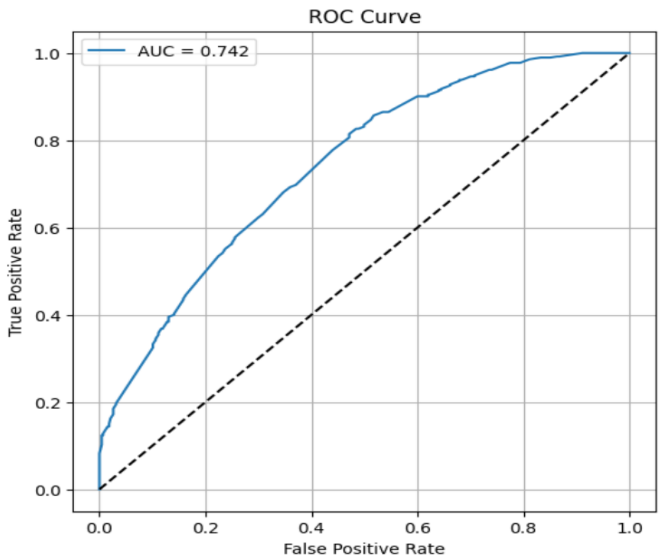

Figure 9 Receiver Operating Characteristic (ROC)

Based on the ROC curve shown in Figure 9, the ROC curve consistently lies above the diagonal line representing random classification. This indicates that the model has a strong ability to distinguish between flooded and non-flooded classes based on the predicted output probabilities. An AUC value of 0.742 indicates that the model exhibits moderate to good discriminative performance. This value implies a 74.2% probability that the model assigns a higher prediction score to flooded areas than to non-flooded areas under random selection. Therefore the ROC-AUC analysis reinforces the results obtained from accuracy- and F1-score-based evaluations and demonstrates that the 2D CNN model performs stably across different decision threshold scenarios.

7.2 ROC curve and AUC analysis

To evaluate the stability and generalization capability of the model under different data partitioning scenarios, this study applies a 5-fold K-Fold Cross-Validation approach. In each fold, the 2D CNN model is retrained using different data subsets and evaluated using the F1-score and AUC metrics, as presented in Table 3.

Table 3. Evaluation results based on 5-fold cross-validation

Metode Evaluasi	Metric	Mean Value
5-Fold Cross Validation	F1-score	0.7080
5-Fold Cross Validation	AUC	0.7469

Note: AUC = Area Under the Curve.

The cross-validation results indicate that the average F1-score is 0.7080, while the average AUC reaches 0.7469. These values are relatively consistent with the evaluation results obtained from the single test dataset, indicating that the model does not suffer from overfitting and exhibits good generalization capability across spatial and temporal data variations. Therefore, the application of K-Fold Cross-Validation provides additional validation that the developed 2D CNN model demonstrates stable performance and is suitable for use as part of a GeoAI-based flood potential monitoring system.

8. VISUALIZATION OF FLOODED AND NON-FLOODED POINTS

This visualization presents the distribution of flooded and non-flooded points based on the processed data of rainfall intensity, waste accumulation, and affected area within the study region. The visualization provides a comprehensive understanding of the spatial distribution of flood events and the underlying environmental conditions. By displaying flooded and non-flooded points separately, the figure helps identify distribution patterns, concentrations of vulnerable areas, and possible relationships between high rainfall, waste accumulation, and changes in inundation extent. This visual narrative plays an important role in supporting further analysis related to the characteristics of the area and the contributing factors that influence the occurrence of flooding.

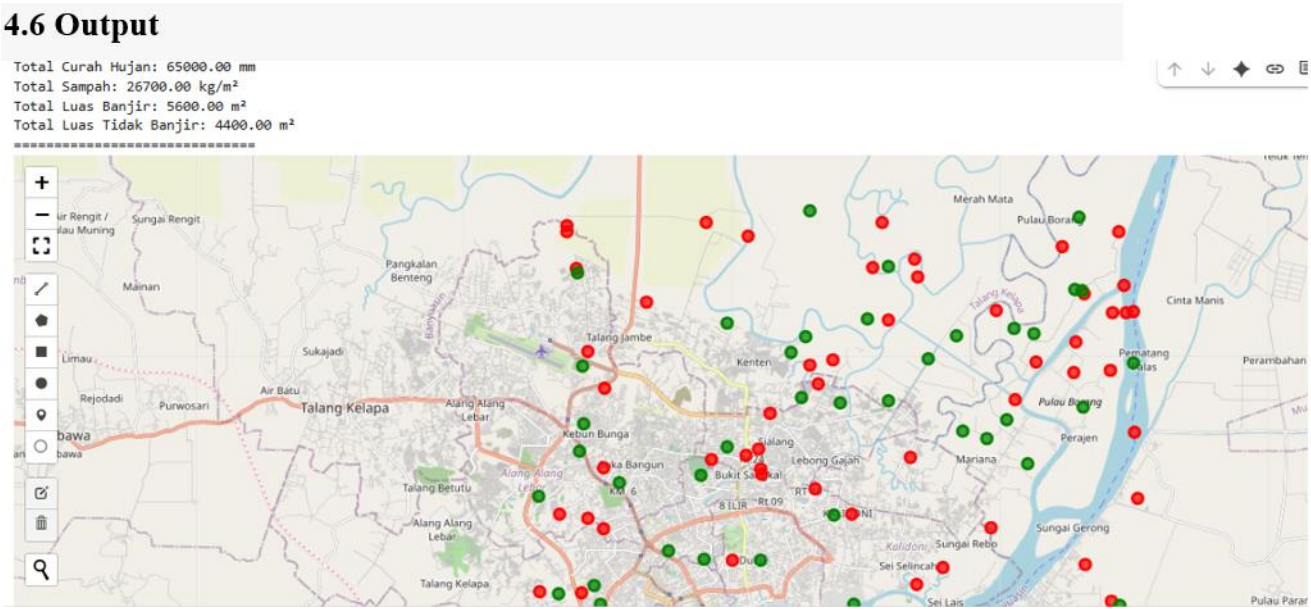

Figure 10. Visualization of flooded and non-flooded locations

Figure 10 presents a visualization of flooded and non-flooded point locations based on environmental data, which include a total rainfall of 65,000 mm, total waste accumulation of 26,700 kg/m², a total flooded area of 5,600 m², and a non-flooded area of 4,400 m². In the map, red symbols represent points that experienced flooding (class 1), while green symbols indicate points that were not flooded (class 0). This visualization is used to observe the spatial distribution of flood events and the patterns of their spread in relation to the environmental parameters being analyzed. It provides an initial understanding of the regional conditions and the level of flood vulnerability based on the available data.

9. CONCLUSIONS

This study proposed and implemented a GeoAI-based flood potential monitoring framework for Palembang City by integrating CHIRPS rainfall data and Sentinel-2–derived waste accumulation information using a 2D CNN. The integration of hydrological and anthropogenic factors represents a novel contribution, as waste accumulation is explicitly incorporated as a key variable influencing urban flood potential, particularly in lowland cities with dense river and drainage networks.

The data processing workflow, which combined grid-based spatial sampling and temporal augmentation through monthly compositing, successfully increased the representativeness of spatial and temporal variability from 2019 to 2025. This approach generated 7,200 spatiotemporal samples, enabling the CNN model to effectively learn complex spatial–temporal relationships between rainfall intensity, waste accumulation, and flood occurrence while preserving the physical characteristics of the original data.

Model evaluation results demonstrate that the proposed 2D CNN achieved a satisfactory performance, with an overall accuracy of 0.73. The flooded class obtained a higher recall (0.80), indicating strong sensitivity in detecting flood events, while the F1-score values for both flooded (0.73) and non-flooded (0.69) classes reflect a balanced classification capability. The ROC–AUC analysis yielded an AUC value of 0.742, confirming moderate to good discriminative performance across varying decision thresholds. Furthermore, the 5-fold cross-validation results, with a mean F1-score of 0.7080 and mean AUC of 0.7469, indicate that the model is stable, does not suffer from overfitting, and generalizes well across different spatial and temporal data partitions.

Spatial flood potential mapping reveals that areas experiencing high rainfall combined with concentrated waste accumulation along river corridors and residential zones exhibit higher flood susceptibility. These findings emphasize the significant role of waste accumulation as an anthropogenic factor that exacerbates urban flooding by reducing drainage and river flow capacity.

In conclusion, the integration of rainfall and waste accumulation data within a GeoAI-based CNN framework provides an effective and scalable approach for urban flood potential monitoring. The proposed model offers practical support for early warning systems, urban drainage management, and data-driven flood mitigation strategies. This framework can be adapted and extended to other flood-prone urban areas with similar environmental and socio-spatial characteristics, contributing to more resilient and intelligent urban flood management systems

ACKNOWLEDGMENT

The authors would like to express their deepest gratitude to Politeknik Negeri Sriwijaya and Universitas Sriwijaya for the support, facilities, and encouragement provided throughout the preparation and completion of this research. Appreciation is also extended to all parties who contributed to the provision of rainfall data, Sentinel-2 imagery, and geospatial processing using Google Earth Engine, enabling this study to be successfully carried out.

The authors also highly appreciate the constructive feedback and valuable suggestions provided by colleagues and peers, which have contributed to improving the clarity of the analysis and the overall quality of this manuscript. This research would not have been completed without the collective contributions of the various institutions and individuals involved.

REFERENCES

[1] Kruse, C., Boyda, E., Chen, S., Karra, K., et al. (2022). Satellite Monitoring of Terrestrial Plastic Waste. arXiv preprint arXiv:2204.01485. https://doi.org/10.48550/arXiv.2204.01485

[2] Tharani, M., Amin, A.W., Maaz, M., Taj, M. (2020). Attention neural network for trash detection on water channels. arXiv preprint arXiv:2007.04639. https://doi.org/10.48550/arXiv.2007.04639

[3] Kim, J., Kim, H., Kim, D.J., Song, J., Li, C. (2022). Deep learning-based flood area extraction for fully automated and persistent flood monitoring using cloud computing. Remote Sensing, 14(24): 6373. https://doi.org/10.3390/rs14246373

[4] Nogueira, K., Fadel, S.G., Dourado, Í.C., Werneck, R.D.O., et al. (2017). Exploiting ConvNet Diversity for Flooding Identification. arXiv preprint arXiv:1711.03564. https://doi.org/10.48550/arXiv.1711.03564

[5] Dong, S., Yu, T., Farahmand, H., Mostafavi, A. (2022). Predictive multi-watershed flood monitoring using deep learning on integrated physical and social sensors data. Environment and Planning B: Urban Analytics and City Science, 49(7): 1838-1856. https://doi.org/10.1177/23998083211069140

[6] Alizadeh, B., Behzadan, A.H. (2023). Scalable flood inundation mapping using deep convolutional networks and traffic signage. Computational Urban Science, 3(1): 17. https://doi.org/10.1007/s43762-023-00090-1

[7] Anshori, S.A., Hadiana, A.I., Kasyidi, F. (2025). An integrated convolutional neural networks and light gradient boosting approach for flood classification using Sentinel-1 SAR satellite imagery. Innovation in Research of Informatics (Innovatics), 7(1): 68-76.

[8] Sarker, C., Mejias, L., Maire, F., Woodley, A. (2019). Flood mapping with convolutional neural networks using spatio-contextual pixel information. Remote Sensing, 11(19): 2331. https://doi.org/10.3390/rs11192331

[9] Youme, O., Bayet, T., Dembele, J.M., Cambier, C. (2021). Deep learning and remote sensing: Detection of dumping waste using UAV. Procedia Computer Science, 185: 361-369. https://doi.org/10.1016/j.procs.2021.05.037

[10] Sallang, N.C.A., Islam, M.T., Islam, M.S., Arshad, H. (2021). A CNN-based smart waste management system using TensorFlow lite and LoRa-GPS shield in Internet of Things environment. IEEE Access, 9: 153560-153574. https://doi.org/10.1109/ACCESS.2021.3128314

[11] Mehmood, H., Conway, C., Perera, D. (2021). Mapping of flood areas using landsat with google earth engine cloud platform. Atmosphere, 12(7): 866. https://doi.org/10.3390/atmos12070866

[12] 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

[13] Ruuhulhaq, M.S., Rohman, A., Shalih, O. (2025). Exploring google earth engine for flood detection (a case study in Bandung City). International Journal for Disaster and Development Interface, 5(1): 24-46. https://doi.org/10.53824/ijddi.v5i1.93

[14] Irbah, N.S., Jaelani, L.M. (2024). Spatial analysis of flood inundation from Sentinel-1 imagery using Google Earth Engine (case study: Bengawan Jero Lamongan Regency). Geoid, 19(2): 202-211. https://doi.org/10.12962/geoid.v19i2.1207

[15] Johary, R., Révillion, C., Catry, T., Alexandre, C., et al. (2023). Detection of large-scale floods using Google earth engine and Google colab. Remote Sensing, 15(22): 5368. https://doi.org/10.3390/rs15225368

[16] Wangchuk, S., Bolch, T., Robson, B.A. (2022). Monitoring glacial lake outburst flood susceptibility using Sentinel-1 SAR data, Google Earth Engine, and persistent scatterer interferometry. Remote Sensing of Environment, 271: 112910. https://doi.org/10.1016/j.rse.2021.112910

[17] Wang, Y., Li, Z., Zeng, C., Xia, G.S., Shen, H. (2019). An urban water extraction method combining deep learning and Google earth engine. arXiv preprint arXiv:1912.10726. https://doi.org/10.48550/arXiv.1912.10726

[18] Giezendanner, J., Mukherjee, R., Purri, M., Thomas, M., Mauerman, M., Islam, A.K.M., Tellman, B. (2023). Inferring the past: A combined CNN-LSTM deep learning framework to fuse satellites for historical inundation mapping. arXiv preprint arXiv:2305.00640. arXiv. https://doi.org/10.48550/arXiv.2305.00640

[19] Sellami, E.M., Rhinane, H. (2024). Google earth engine

and machine learning for flash flood exposure mapping—Case study: Tetouan, Morocco. Geosciences, 14(6): 152. https://doi.org/10.3390/geosciences14060152

[20] Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, 14(5): 778-782. https://doi.org/10.1109/LGRS.2017.2681128

[21] Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., et al. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Scientific Data, 2(1): 150066. https://doi.org/10.1038/sdata.2015.66

[22] Abdi, A.M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GIScience & Remote Sensing, 57(1): 1-20. https://doi.org/10.1080/15481603.2019.1650447