International Journal of Applied Earth Observation and Geoinformation 152 (2026) 105482

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

From flood extent mapping to mechanism-aware flood products: integrating flood type classification into satellite-based flood monitoring

Ziming Wang [a], Jeffrey Neal [a], Peter M. Atkinson [b,c,d], Ce Zhang [a,e,*]

[a] School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK
[b] Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YR, UK
[c] Geography and Environmental Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[d] College of Surveying and Geo-Informatics, Tongji University, No.1239, Siping Road, Shanghai 200092, PR China
[e] UK Centre for Ecology & Hydrology, Library Avenue, Lancaster LA1 4AP, UK

ARTICLE INFO

Keywords:
Floodtype
Flood mapping
Deep learning
CNN
Remote sensing
Flood risk management

ABSTRACT

Flood type information is critical for effective flood risk management because dominant flood types are associated with distinct hydrodynamic behaviour, contamination pathways, and recovery trajectories. However, most operational flood mapping products provide only binary inundation extent, offering limited information for interpreting the dominant flood type and its likely impact characteristics. Existing flood type classification approaches rely predominantly on hydrometeorological observations and modelling, which are often unavailable in data-scarce regions and can be unstable in mechanism-complex environments such as estuarine deltas, urban river corridors, and coastal cities. To address these limitations, this study proposes a multi-CNN framework that integrates flood type classification directly into satellite-based flood mapping. The framework first uses a U-Net model for flood extent segmentation and then applies a CNN-based classifier for scene-level flood type identification by combining satellite imagery with auxiliary topographic and hydrological-context features, including DEM information and water connectivity ratios. Several CNN architectures were compared for flood type classification, with Inception-ResNet selected based on its performance–complexity trade-offs. To support model training and evaluation, this study introduces the World Flood Type (WFT) dataset, a new multi-event flood dataset containing 464 flood scenes from 120 flood events across 48 countries. Results show effective performance in both inundation segmentation and flood type classification. The U-Net model achieved 83.8% overall accuracy on an independent test dataset, while the final Inception-ResNet-based classifier achieved 93.0% scene-level overall accuracy and a scene-level macro-F1 score of 0.73 under dominant-mechanism conditions. These findings demonstrate the feasibility of extending conventional flood extent products with scene-level attribution of dominant flood types using remotely sensed imagery and ancillary geospatial data, thereby providing more decision-relevant information for emergency response and post-disaster assessment.

1. Introduction

Flooding is one of the most destructive natural hazards globally, causing substantial economic losses and posing increasing threats to human safety under climate change and rapid urbanisation (Shahapure et al., 2010; Zhang et al., 2020). Effective flood risk management relies not only on timely identification of inundation extent, but also on understanding the underlying flood-generating mechanisms that shape hydrodynamic behaviour, contamination pathways, and recovery trajectories (Peng et al., 2021; Schumann et al., 2022).

Over the past decade, satellite-based flood mapping has advanced significantly and now provides near-real-time delineation of inundated areas from regional to global scales. However, most operational products remain extent-centric, focusing primarily on mapping flood boundaries. Such binary representations provide limited insight into the processes driving flooding and therefore constrain mechanism-level interpretation. Different flood types are associated with distinct hydrological characteristics and impact profiles. For example, pluvial flooding is often linked to urban drainage exceedance and localised contamination risks (Welch et al., 2022), fluvial flooding may involve prolonged inundation and sediment transport (Vázquez-Tarrío et al., 2024), and coastal flooding can include saline intrusion and compound tidal effects

* Corresponding author at: School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK.
E-mail address: ce.zhang@bristol.ac.uk (C. Zhang).

https://doi.org/10.1016/j.jag.2026.105482
Received 18 February 2026; Received in revised form 26 June 2026; Accepted 13 July 2026
Available online 15 July 2026

(Barui et al., 2025). Identifying flood type therefore offers a mechanism-aware perspective that can support emergency response, infrastructure planning, insurance assessment, and post-disaster recovery (Moftakhari et al., 2017). In this study, "mechanism-aware" refers to the identification of dominant flood types, namely fluvial, pluvial, and coastal flooding, rather than pixel-level reconstruction of detailed hydrodynamic processes.

Existing flood type classification approaches rely predominantly on hydrometeorological and physical drivers (Fernandes, 2025; Tai et al., 2025). While physically interpretable, such methods are often constrained by data availability and can be difficult to implement reliably in observation-sparse regions (El Khalki et al., 2020; Brunner et al., 2021). Temporal misalignment among rainfall, river stage, and tidal data further complicates classification during short-duration events (Huang et al., 2018). Moreover, in mechanism-complex settings such as estuarine deltas, urban river corridors, and coastal cities, compound drivers may interact, making single-driver inference prone to misclassification (Green et al., 2025; Jo et al., 2025). These limitations highlight the need for complementary, observation-driven approaches to flood type attribution that can operate independently of dense hydrometeorological infrastructure.

With continuous improvements in spatial and temporal resolutions of satellite sensors, remote sensing imagery offers an increasingly viable pathway for observation-driven flood type inference (Bentivoglio et al., 2022). Different flood-generating mechanisms often produce distinct inundation morphologies. For instance, fluvial flooding is commonly shaped by river channel overflow and therefore often appears as inundation connected to river channels and extending across low-lying floodplain areas (Call et al., 2017). Remote sensing imagery can capture these mechanism-related spatial patterns, while ancillary geospatial variables such as elevation and water connectivity can provide additional contextual information for distinguishing flood types.

Because both flood extent extraction and flood type recognition rely on spatial patterns within imagery, flood type classification can be embedded within existing flood mapping workflows to produce integrated and computationally efficient mechanism-aware products. Distinguishing flood types from imagery requires capturing inundation morphology, boundary structure, terrain context, water connectivity, and broader spatial configuration rather than relying solely on pixel-level spectral differences. Convolutional neural networks (CNNs), which effectively model spatial relationships and texture features, are therefore well suited for this task (Panahi et al., 2021; Bai et al., 2022).

However, existing CNN-based flood mapping research has mainly focused on flood extent extraction, while relatively limited attention has been paid to flood type identification (Bentivoglio et al., 2022). Although many recent studies have incorporated multi-source data, such as remote sensing imagery, DEMs, and water body information, into flood mapping workflows to enhance the representation of flood boundaries, terrain context, and water connectivity, these fusion strategies have generally been designed to improve inundation boundary delineation rather than to support flood type attribution (Muñoz et al., 2021; Chen et al., 2026). In addition, when CNNs are used for flood type classification, further consideration is needed regarding how to robustly evaluate classification accuracy and result reliability, especially when class imbalance and compound flood mechanisms coexist.

Building on these considerations, this study investigates whether dominant flood types can be attributed at the scene level using remote sensing imagery and ancillary geospatial data, and develops an integrated workflow combining flood extent segmentation with dominant flood type attribution. To support model training and validation, we constructed the World Flood Type (WFT) dataset, containing 464 flood scenes corresponding to 120 flood events across 48 countries. A U-Net architecture is employed as the segmentation backbone (Konapala et al., 2021), and multiple CNN architectures are systematically benchmarked for flood type classification to identify an appropriate model for integration.

The main contributions of this study are as follows:

1. We extend conventional satellite-based flood extent mapping by introducing scene-level attribution of dominant flood types, including fluvial, pluvial, and coastal flooding.
2. We construct the World Flood Type (WFT) dataset, comprising 464 flood scenes from 120 flood events across 48 countries, together with flood extent labels and auxiliary topographic and hydrological-context data.
3. We introduce Water Connectivity Ratio features to characterise the spatial relationship between mapped inundation and permanent inland or marine water bodies.
4. We demonstrate how flood extent segmentation and dominant flood type attribution can be combined within a unified workflow to provide flood products enriched with both inundation extent and dominant flood type information.

The remainder of this paper is structured as follows: section 2 reviews the related literature. Section 3 describes the datasets and their geographic coverage. Section 4 presents the proposed methodology and CNN architectures. Section 5 reports the experimental results, followed by discussion in section 6. Section 7 concludes the study.

2. Related research

2.1. Flood types and their characteristics

Flood events are commonly categorised into three hydrological types: fluvial, pluvial and coastal (Rentschler et al., 2023). These types differ fundamentally in their driving mechanisms and consequently in their hydrodynamic behaviour, spatial manifestation, duration, and water chemistry. Coastal flooding is typically driven by storm surge and tidal interaction, leading to seawater inundation, saline intrusion, and corrosion risks to infrastructure (Lane et al., 2013; Le Roy et al., 2015). Fluvial flooding occurs when river discharge exceeds channel capacity, often producing extensive floodplain inundation and transporting sediment and pollutants over large areas (Merz et al., 2010). Pluvial flooding is commonly triggered by intense local rainfall and insufficient drainage capacity, resulting in short-duration but spatially heterogeneous inundation patterns in urban environments (Falconer et al., 2009; Rosenzweig et al., 2018).

Because these flood types differ in hydrodynamic behaviour and contaminant pathways, they imply distinct damage mechanisms, exposure risks, and recovery trajectories. However, such process-level distinctions are rarely reflected in satellite-based flood mapping products, which typically focus on binary inundation extent. This disconnect between process understanding and mapping representation motivates the exploration of approaches capable of inferring flood-generating mechanisms directly from observable spatial patterns.

2.2. State-of-the-art flood mapping products

Over the past decade, several operational platforms have emerged to provide near-real-time flood extent information at regional to global scales. For example, the Copernicus Emergency Management Service (EMS) delivers rapid flood delineation maps derived from satellite imagery to support disaster response. The Dartmouth Flood Observatory (DFO) produces global inundation maps using moderate-resolution optical imagery. NASA's Global Flood Monitoring System (GFMS) integrates satellite-derived precipitation data with hydrological modelling to generate near-real-time flood monitoring products. Google's Flood Hub combines satellite observations, hydrological forecasts, and gauge data to provide riverine flood alerts across multiple countries (Abu and Ibebuchi, 2025).

These systems have substantially improved situational awareness by identifying where flooding occurs. Nevertheless, most operational products remain extent-centric and do not explicitly distinguish among flood-generating mechanisms. As a result, downstream users need to infer flood type indirectly from ancillary information or local knowledge. This extent-only representation limits the ability of emergency responders, insurers, and planners to rapidly interpret the behaviour, duration, and impact characteristics of flood events. Integrating flood type attribution into flood mapping workflows therefore represents a critical step toward more decision-relevant flood intelligence.

2.3. CNNs for flood mapping and flood type attribution

Convolutional neural networks (CNNs) have become a dominant approach in remote sensing-based flood mapping due to their ability to capture spatial context and multi-scale features (Wang et al., 2020; Hou et al., 2021; Farhadi et al., 2024). This ability is particularly relevant for flood-related applications, where inundation patterns are shaped by interactions among terrain, drainage networks, land cover, and flood-generating processes. Recent hydrological deep learning studies have similarly shown that CNN-based architectures can extract process-relevant information from spatially structured environmental data (He et al., 2026). In flood mapping, most existing CNN applications focus on inundation extent extraction as a semantic segmentation task. Architectures such as U-Net have demonstrated strong capability in delineating flood boundaries and preserving fine spatial details, particularly in heterogeneous landscapes and data-constrained settings (Löwe et al., 2021).

Beyond extent mapping, CNNs have been effective in remote sensing classification tasks where class distinctions depend on morphology, texture, and spatial context rather than solely on spectral differences, such as crop type recognition and land use classification (Seydi et al., 2022; Blickensdörfer et al., 2023; Zhao et al., 2023). Flood type classification similarly relies on recognising differences in inundation form, spatial arrangement relative to terrain and water bodies, and broader contextual patterns. This makes CNN-based architectures potentially suitable for learning mechanism-related spatial signatures from flood imagery.

Despite this conceptual alignment, explicit application of CNNs to flood type attribution remains limited. Existing studies focus primarily on inundation extraction, and there is a lack of systematic evaluation of CNN architectures for mechanism-based flood classification. This gap highlights the need for rigorous evaluation of deep learning models to determine suitable architectures for integrating flood type classification into operational flood mapping workflows.

3. Dataset

This study constructed a geospatial dataset for flood type identification, named the World Flood Type dataset (WFT dataset). The dataset was developed from the WorldFloodsv2 dataset (Mateo-Garcia et al., 2021), which provides Sentinel-2 flood scenes and associated flood extent vector labels. By filtering the flood vector attributes, we selected 464 flood scenes that contained mapped flood extent excluding permanent water bodies. These scenes correspond to 120 flood events across 48 countries.

The dataset was partitioned into training, validation, and testing subsets using a 6:2:2 ratio, resulting in 279 training scenes, 92 validation scenes, and 93 testing scenes. The partition was performed at the scene level because each flood scene represents a distinct spatial unit with its

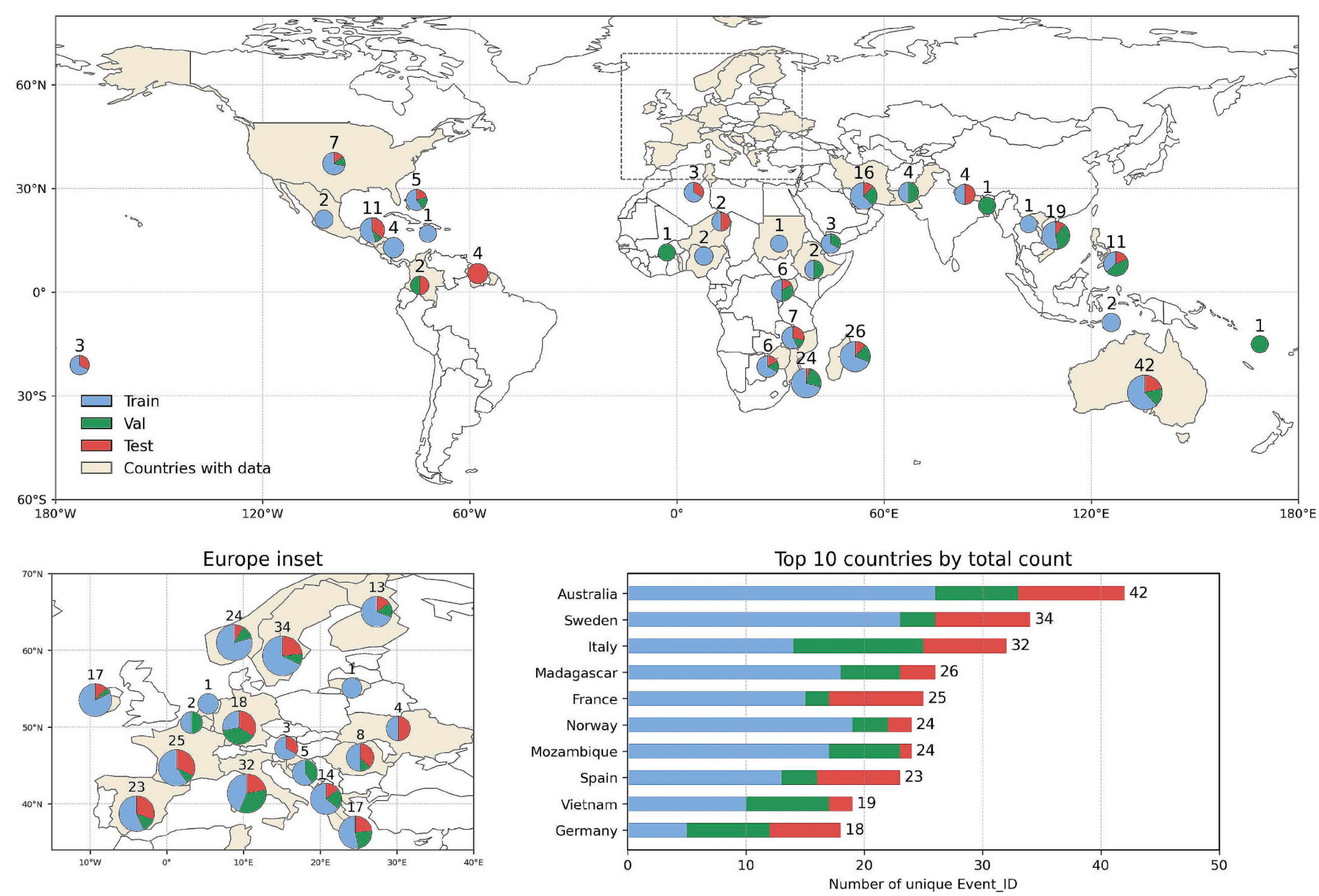

Fig. 1. Global distribution and train, validation, and test split of WFT dataset.

own local terrain, land-cover, water-connectivity, and image conditions. Because flood type labels were assigned according to the dominant mechanism of each flood event, the reported results should be interpreted as scene-level dominant flood type classification rather than strict event-level generalisation. The global distribution of the selected flood scenes and their train–validation–test split is shown in Fig. 1.

The WFT dataset was organised using flood events and their corresponding image scenes in WorldFloodsv2 as the basic units. For each scene, Sentinel-2 RGB bands, including B02, B03, and B04, were retained to represent surface spectral characteristics during flood events. In addition, DEM data and permanent water extent data were added to provide topographic and hydrological-context information. Specifically, DEM data were used to characterise terrain elevation, while permanent water extent data were used to estimate the spatial connectivity between flooded areas and existing water bodies (Fig. 2). Study area extent and cloud cover information were also retained to support data processing and filtering.

The DEM data in WFT dataset were derived from FathomDEM (Uhe et al., 2025). Original DEM tiles were clipped and mosaicked according to the spatial extent of each flood scene. Permanent water information was constructed separately for inland water bodies and ocean water bodies to better represent hydrological context. For inland water bodies, including rivers, lakes, and reservoirs, water extent was generated using the SWOT River Database (SWORD) and the Prior Lake Database (PLD). River vector data were reconstructed using node information and river reach widths from SWORD (Altenau et al., 2021). The reconstructed river vectors were then smoothed and merged with the lake vector data from PLD to generate the inland permanent water mask (Wang et al., 2025). For ocean water bodies, OpenStreetMap-derived water polygons were used to generate the ocean water mask.

To support flood type classification, dominant flood type labels were assigned to all flood scenes. The labelling process was conducted by four researchers. First, three researchers independently assessed the flood type of each event based on event descriptions and supporting documentation. A fourth researcher then reviewed the event descriptions together with the three independent assessments to determine the dominant flood type. This procedure was designed to reduce individual labelling bias and improve label consistency.

Among the 464 scenes, 400 were labelled as fluvial floods, 33 as pluvial floods, and 31 as coastal floods. This distribution reflects the dominance of fluvial events in the selected WorldFloodsv2 scenes, but also introduces substantial class imbalance. Therefore, model evaluation in this study reports not only overall accuracy but also class-sensitive metrics, including macro-F1, to provide a more balanced assessment of classification performance. For scenes with substantial disagreement among annotators, a secondary dominant flood type was also retained to document potential ambiguity associated with compound or mixed flood mechanisms.

4. Methodology

To integrate flood type attribution into satellite-based flood mapping, we designed a two-stage modelling framework that separately addresses inundation detection and mechanism classification. The framework first performs pixel-level flood extent segmentation using a U-Net architecture, and subsequently conducts scene-level flood type classification using convolutional neural networks (CNNs). Six representative CNN architectures were benchmarked to identify a suitable classifier for flood type recognition. The selected classifier was then combined with the segmentation backbone to form a multi-CNN framework capable of jointly producing spatially explicit flood extent maps and corresponding flood type labels. Fig. 3 illustrates the overall system architecture and processing workflow.

4.1. Data preparation and two-stage modelling framework

The proposed framework consists of two sequential stages: flood extent segmentation followed by flood type classification. This design reflects the conceptual distinction between pixel-level inundation detection and scene-level flood mechanism attribution. In the first stage, a U-Net model was trained to generate binary inundation masks. These masks were subsequently used to derive water connectivity features for flood type classification by quantifying the spatial relationship between flooded areas and permanent inland or marine water bodies. A detailed description of the connectivity features is provided in Section 4.3.1. In the second stage, the flood type classification model used RGB imagery, DEM information, and water connectivity ratio (WCR) features as inputs to classify each flood scene into one of three dominant flood types: fluvial, pluvial, or coastal. During model training, WCR features were derived from the ground-truth flood extent masks in the WFT dataset to

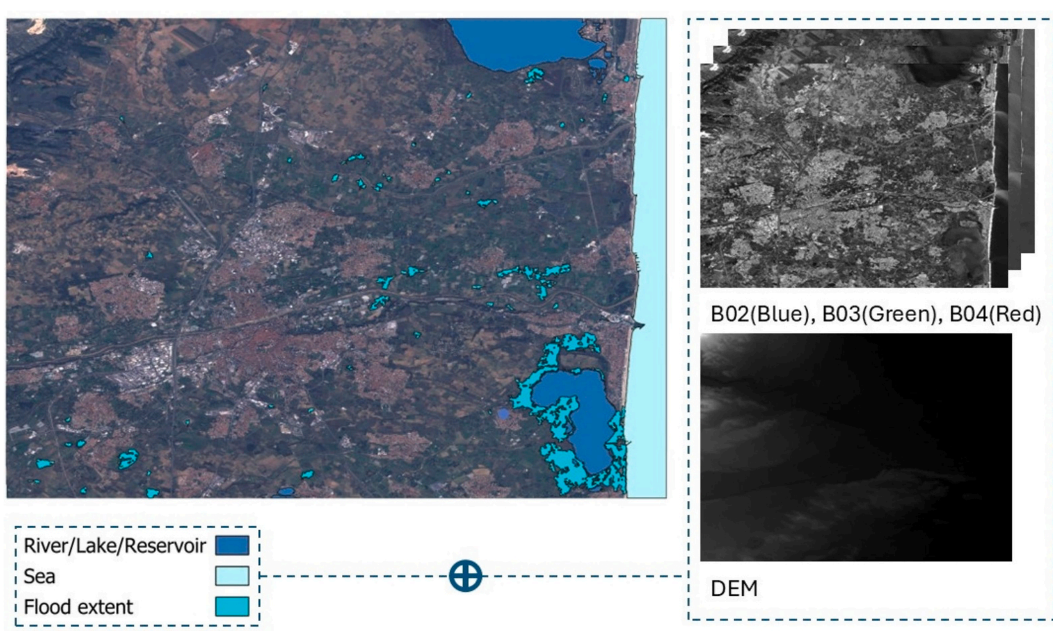

Fig. 2. Overview of remote sensing and auxiliary data included in each flood scene of the WFT dataset.

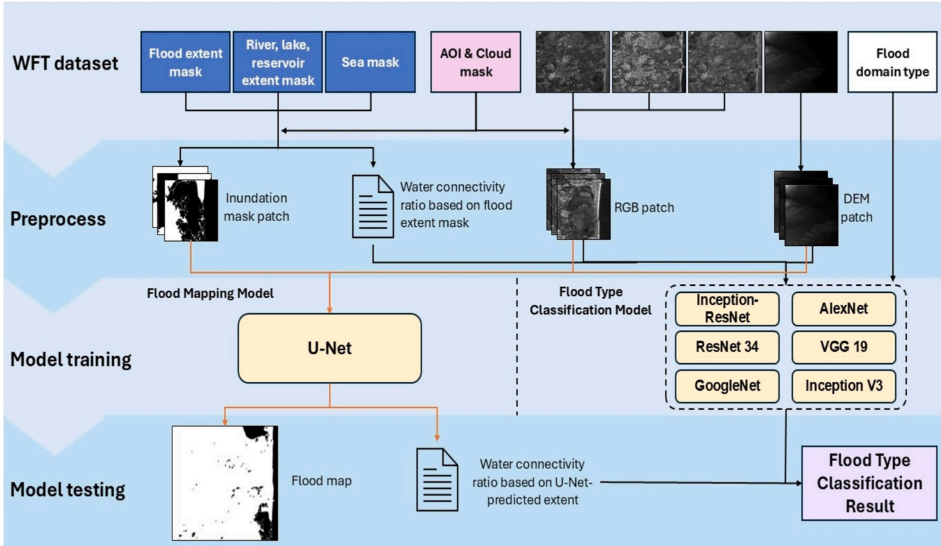

Fig. 3. System flow diagram describing the methodology for flood extent segmentation and flood type classification.

avoid propagating segmentation errors into the classifier learning process. During end-to-end workflow inference, WCR features were derived from the predicted inundation masks generated by the U-Net model, thereby enabling flood extent mapping and flood type attribution to be performed within a unified pipeline.

Before model training, all flood scenes in WFT dataset were preprocessed using a consistent workflow. First, all input data were spatially aligned using the Sentinel-2 B02, B03, and B04 bands as the reference grid. Vector data were then rasterised into masks. Sentinel-2 imagery and DEM data were normalised, and invalid areas were removed using the cloud mask and study area mask. Each scene was organised as multi-channel input data, while the flood extent mask and flood type label were retained for pixel-level flood extent segmentation and scene-level flood type classification, respectively.

To support model training, all scenes were divided into non-overlapping patches of 256 × 256 pixels. This patch-based strategy reduced computational cost while preserving local inundation morphology and surrounding spatial context. Patches with very limited flood coverage were removed because they provide little information for either segmentation or flood type classification. Specifically, patches

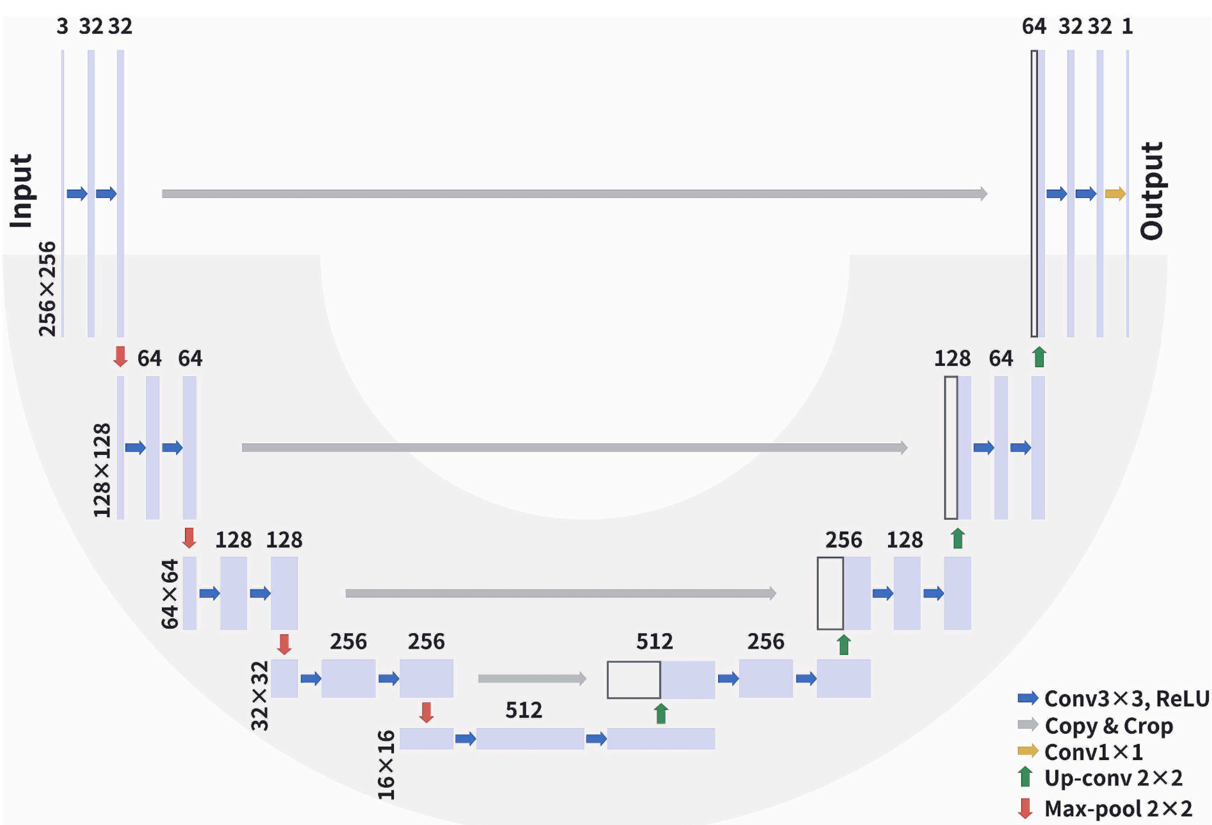

Fig. 4. U-Net model adopted for flood area segmentation.

were retained only when they contained at least 200 flood pixels and a minimum flood ratio of 2%. This filtering step reduced redundant samples while preserving patches with sufficient inundation information for model training.

4.2. Flood extent mapping model

Flood extent segmentation was performed using a U-Net architecture, selected for its ability to capture multi-scale spatial context and preserving boundary details in heterogeneous landscapes (Ronneberger et al., 2015). U-Net has been widely adopted for remote sensing-based water body segmentation because its encoder–decoder structure and skip connections enable effective fusion of global context and local spatial features. This choice is also consistent with the ML4Floods pipeline used for flood segmentation in the original WorldFloods v2 dataset, where U-Net serves as the segmentation backbone (Portalés-Julià et al., 2023).

The implemented architecture is illustrated in Fig. 4. The encoder consists of repeated 3 × 3 convolutional layers followed by ReLU activation and 2 × 2 max pooling with a stride of 2 for spatial downsampling. The number of feature channels is doubled at each downsampling step. The decoder symmetrically upsamples feature maps and concatenates them with corresponding encoder features via skip connections, followed by 3 × 3 convolutions and ReLU activations to refine spatial details.

For the flood extent segmentation stage, we used the filtered split derived from WorldFloodsv2, consisting of 426 training scenes, 20 validation scenes, and 20 independent test scenes. This split was used only to train and evaluate the inundation segmentation backbone. The flood type classification stage used the WFT train–validation–test split described in Section 3.

The U-Net flood extent segmentation model was trained using four-channel inputs consisting of RGB imagery and DEM data. The model was optimised with AdamW using an initial learning rate of 1×10^{-4}, a weight decay of 1×10^{-4}, and a batch size of 48 for up to 150 epochs. The loss function combined masked binary cross-entropy with logits and masked Dice loss, where invalid pixels identified by the valid mask were excluded from loss calculation. During evaluation and prediction, sigmoid probabilities were converted into binary flood masks using a threshold of 0.5. A scheduler was used to reduce the learning rate by a factor of 0.5 when validation performance did not improve for eight epochs, and early stopping was applied with a patience of ten epochs. Training was conducted on an NVIDIA GeForce RTX 5090 GPU with 32 GB memory.

4.3. Flood type classification model

Flood type classification was formulated as a three-class scene-level classification task, including fluvial, pluvial, and coastal flooding. Because each flood scene was divided into multiple patches, CNN predictions were first generated at the patch level and then aggregated to the scene level using majority voting. The final flood type of each scene was assigned as the class predicted for the largest number of valid patches. Flood type classification requires not only spectral information from flood imagery but also contextual information describing the spatial relationships among inundation, terrain, and permanent water bodies. DEM data were therefore incorporated to provide topographic context. In addition, this study introduced water connectivity ratio features to quantify the spatial connectivity between flood extent and different types of permanent water bodies.

4.3.1. Water connectivity ratio

This study proposed the Water Connectivity Ratio (WCR) to quantify the proportion of flooded pixels spatially connected to permanent inland water bodies or the sea. The underlying assumption is that different flood types exhibit different water connectivity patterns. Fluvial floods are more likely to be connected to rivers, lakes, or reservoirs; coastal floods are more likely to be connected to the sea; and pluvial floods may contain a higher proportion of inundated pixels that are not directly connected to permanent water bodies.

To calculate WCR, the flood extent mask was first reconstructed at the scene level and combined with permanent inland water and marine water masks. Connectivity was then evaluated by determining whether connected flood components intersected buffered water source regions. In this study, a 50 m buffer was applied to permanent inland water bodies, and a 300 m buffer was applied to the sea. To reduce unstable connections caused by narrow or noisy links, a minimum connection width of 100 m was used in the connectivity calculation (Fig. 5). These parameters were selected as practical thresholds based on the 10 m analysis grid and the spatial characteristics of the permanent-water masks. The 50 m inland-water buffer was intended to account for small spatial offsets between mapped flood extent and river or lake boundaries, equivalent to approximately five Sentinel-2 pixels. The larger 300 m marine buffer was used to accommodate greater uncertainty in coastline delineation and the potentially broader transition zone between coastal inundation and mapped marine water. The 100 m minimum connection width was applied to remove narrow and potentially spurious links while retaining broader flood connections that were less likely to be artefacts of mask geometry or rasterization. After morphological filtering, flood components touching buffered water source regions were regarded as water-connected flooded pixels. WCR was calculated as the ratio between the number of connected flooded pixels and the total number of flooded pixels.

Based on this design, four scene-level WCR features were extracted: (1) the proportion of flooded pixels connected to permanent inland water bodies; (2) the proportion of non-inland-water-connected flooded pixels connected to the sea; (3) the proportion of all flooded pixels connected to the sea; and (4) the proportion of flooded pixels connected to neither permanent inland water bodies nor the sea. Because these features were calculated at the scene level, the same four WCR values were assigned to all patches from the corresponding scene. In the classification model, the WCR features were passed through a lightweight auxiliary feature encoder and projected into a 16-dimensional representation. This auxiliary representation was then concatenated with CNN-extracted features from RGB and DEM inputs before being passed to the final classification head.

4.3.2. Model selection and training settings

To evaluate the suitability of different feature extraction mechanisms for flood type classification, six representative CNN architectures were benchmarked: AlexNet, VGG19, ResNet34, GoogLeNet, InceptionV3, and Inception-ResNet. These models represent distinct architectural paradigms. AlexNet, and VGG19 correspond to classic convolutional design with progressively increasing depth and representational capacity (Krizhevsky et al., 2012; Simonyan and Zisserman, 2014). ResNet34 introduces residual connections to facilitate deeper feature learning and increase optimisation stability (He et al., 2016). GoogLeNet and InceptionV3 employ multi-branch convolutional modules that capture features at multiple receptive field scales, which is relevant for recognising heterogeneous inundation morphology (Szegedy et al., 2015, Szegedy et al., 2016a). Inception-ResNet combines multi-scale feature extraction with residual learning, potentially enhancing both representation capacity and training stability (Szegedy et al., 2016b).

For each architecture, the final classification layer was adapted to output three classes. The models were trained using four-channel inputs, including RGB imagery and DEM data, together with scene-level WCR features. All models were optimised using AdamW with an initial learning rate of 5×10^{-5} and a weight decay of 5×10^{-4} for up to 120 epochs. To reduce the influence of class imbalance, inverse-frequency class weights were calculated from the number of training samples in each flood type and applied to the cross-entropy loss, giving higher loss weights to under-represented classes. In addition, an inverse-frequency

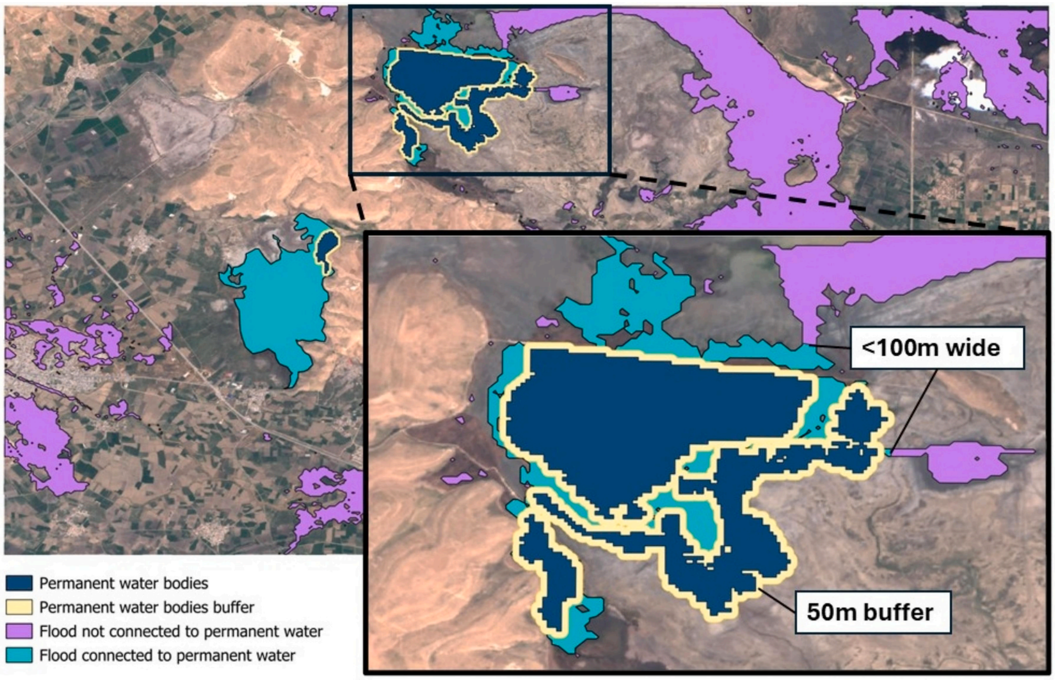

Fig. 5. Schematic diagram of WCR-based flood connectivity delineation, illustrated by identifying flooded pixels connected to permanent inland water bodies. The selected case is the Lake Urmia flood event in Iran.

weighted random sampler was used to increase the probability of sampling patches from minority classes during training. Mixed-precision training was enabled, and early stopping was applied with a patience of 40 epochs.

4.4. Evaluation criteria

Flood extent segmentation was evaluated as a binary classification task, distinguishing flood pixels from non-flood pixels. Performance was assessed using overall accuracy, precision, recall, and F1 score, with overall accuracy and F1 score used as the primary evaluation metric. The standard definitions are:

$$OverallAccuracy = \frac{TN + TP}{TN + TP + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

Flood type classification was evaluated as a three-class classification task. Overall accuracy was calculated at the scene level. Because the dataset is strongly imbalanced across flood types, macro-averaged precision, recall, and F1 score were also reported to assign equal importance to fluvial, pluvial, and coastal classes. The macro-averaged metrics are defined as:

$$MacroPrecision = \frac{1}{C}\sum_{c=1}^{C} Precision_c \tag{5}$$

$$MacroRecall = \frac{1}{C}\sum_{c=1}^{C} Recall_c \tag{6}$$

$$MacroF1score = \frac{1}{C}\sum_{c=1}^{C} F1score_c \tag{7}$$

where C is the number of flood type classes. Overall accuracy and Macro-F1 score were used as the main reported metrics for flood type classification.

5. Results

5.1. Flood extent segmentation

The trained U-Net model was evaluated on the WFT validation and independent test datasets to assess its flood extent segmentation performance. A representative example is shown in Fig. 6. The predicted inundation mask generally follows the spatial pattern of the reference flood extent, capturing the main flooded areas while preserving parts of the flood boundary structure.

Across all validation regions, the model achieved an overall accuracy of 92.2% and an F1 score of 81.1%. On the independent test dataset, overall accuracy decreased to 83.8%, with an F1 score of 64.4%. This performance drop indicates that flood extent segmentation remains challenging when applied to unseen scenes with heterogeneous land cover, cloud contamination, variable flood morphology, and differing image conditions. Nevertheless, the test results remain within a reasonable range for global flood mapping across diverse geographic settings and are broadly consistent with performance levels reported in recent deep learning-based flood mapping studies (Abid et al., 2021; Roth et al., 2023). The segmentation model therefore provides a usable inundation backbone for subsequent flood type attribution, while also highlighting the importance of accounting for segmentation uncertainty in end-to-end flood product generation.

5.2. Flood type classification

Six CNN architectures were trained and evaluated on the WFT dataset under identical settings to ensure a fair comparison. The test

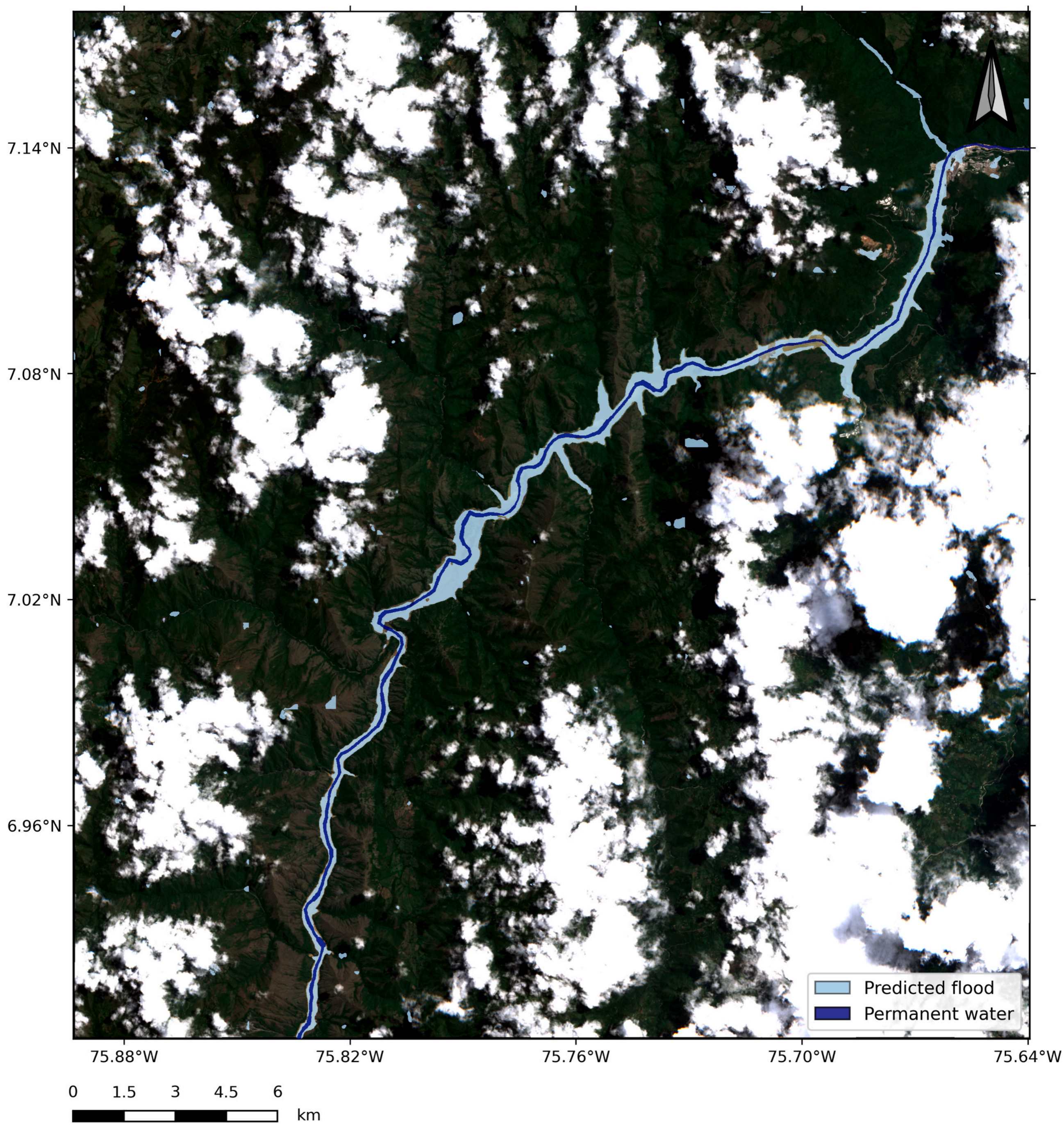

Fig. 6. Example of flood extent mapping results. The displayed sample is from the flood event north of Ituango on 22 May 2018.

dataset comprised 90 flood scenes, from which 3,020 patches were generated. To assess the contribution of different input components, four model configurations were evaluated for each CNN backbone: RGB-only input, RGB with WCR features, RGB with DEM information, and the complete model combining RGB, DEM, and WCR features.

As shown in Fig. 7, most models achieved scene-level overall classification accuracy above 80%. The complete model based on Inception-ResNet achieved the highest overall accuracy of 93.0%. The results indicate that different CNN backbones respond differently to the DEM and WCR modules. Adding DEM generally improved classification accuracy, suggesting that topographic context provides useful information for distinguishing flood types. In contrast, the effect of WCR features varied across architectures. AlexNet, GoogLeNet, and InceptionV3 showed performance degradation after adding WCR features, whereas VGG19 and Inception-ResNet were better able to benefit from the additional water connectivity information. This suggests that the usefulness of scene-level connectivity features depends on the model's capacity to combine auxiliary contextual variables with image-derived features.

As shown in Fig. 8, scene-level macro-F1 scores were consistently lower than corresponding overall accuracy values, indicating that class imbalance affected model evaluation. The complete Inception-ResNet model achieved the highest macro-F1 score of 0.73, consistent with its best performance in overall accuracy. The gap between overall accuracy and macro-F1 highlights the difficulty of achieving balanced recognition across all flood types, particularly given the limited number of pluvial and coastal scenes. AlexNet achieved relatively competitive macro-F1 despite its lower overall accuracy. This may be because its simpler

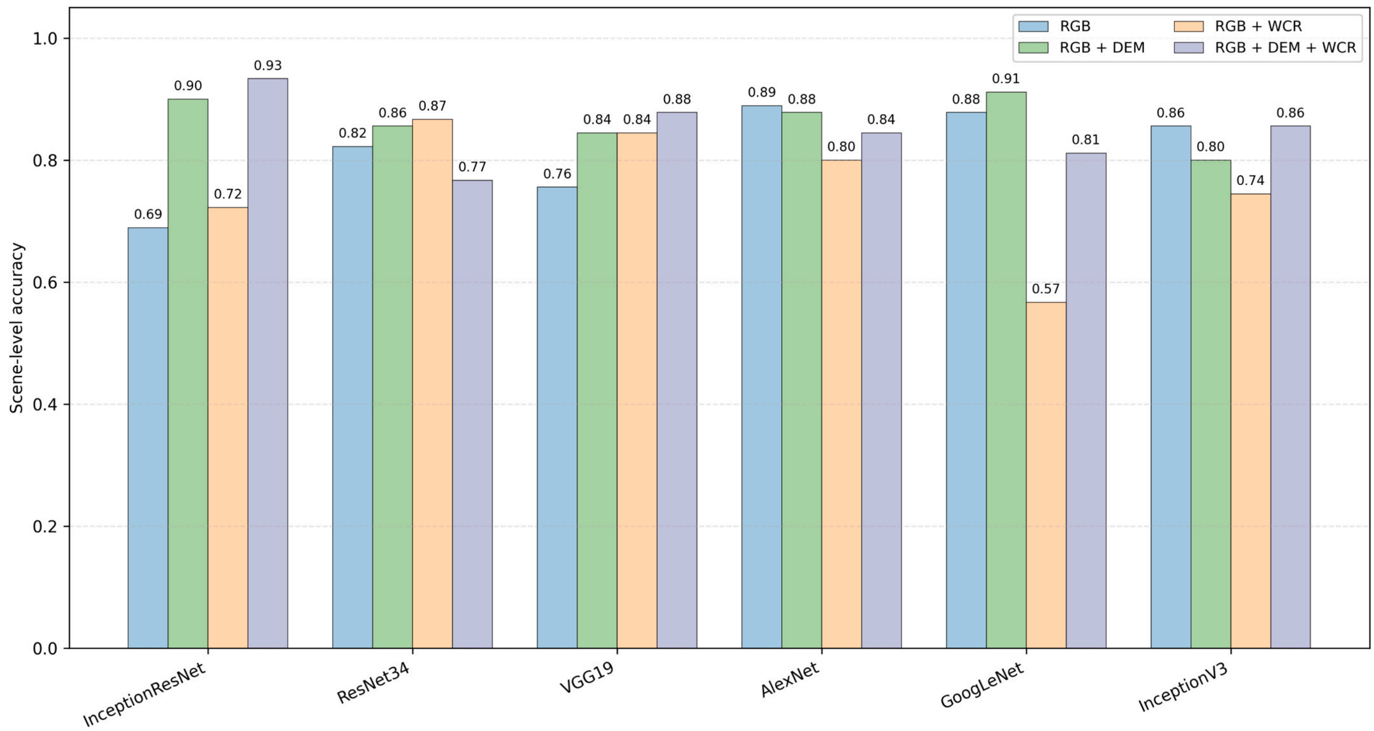

Fig. 7. Ablation analysis of scene-level classification accuracy across different CNN backbones with DEM and WCR modules.

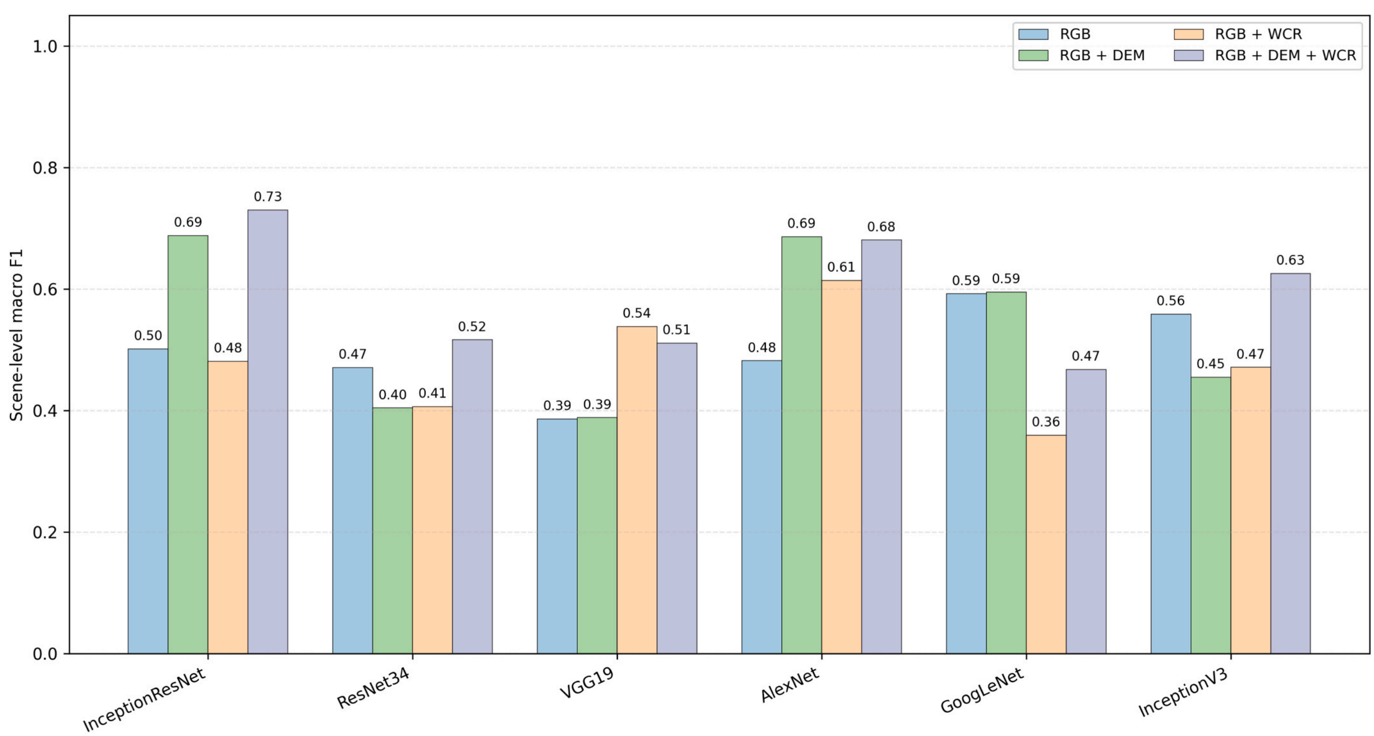

Fig. 8. Ablation analysis of DEM and WCR modules based on scene-level macro-F1 across different CNN backbones.

architecture is less prone to overfitting to the majority fluvial class under class-imbalanced conditions. In contrast, the macro-F1 performance of GoogLeNet and ResNet34 was less stable, suggesting that multi-scale feature aggregation or residual learning alone was insufficient to ensure balanced class-wise recognition without effective integration of contextual features.

Table 1 reports the scene-level class-wise performance of the complete Inception-ResNet configuration on the WFT test set. Overall accuracy was influenced by the predominance of fluvial scenes, for which the model achieved high precision (0.963), recall (0.987), and F1 score (0.975). In contrast, minority-class performance was less balanced. Pluvial flooding achieved perfect precision but low recall (0.333), indicating that the model identified only a small proportion of pluvial scenes while making few false pluvial predictions. Coastal flooding

Table 1
Scene-level class-wise performance of the complete Inception-ResNet configuration (RGB + DEM + WCR) on the WFT test set.

Flood type	Precision	Recall	F1 score	Support
Fluvial	0.963	0.987	0.975	78
Pluvial	1.000	0.333	0.500	6
Coastal	0.625	0.833	0.714	6
Macro average	0.863	0.718	0.730	90

achieved relatively high recall (0.833) but lower precision (0.625), indicating that some non-coastal scenes were misclassified as coastal. These class-specific results explain the lower macro-F1 score of 0.730 and demonstrate that overall accuracy alone does not fully reflect balanced recognition across flood types. The limited support for pluvial and coastal classes, with only six test scenes each, should be considered when interpreting these estimates, as misclassifications have a substantial effect on class-specific recall and F1 scores.

The confusion matrices provide further insight into class-wise performance. As shown in Fig. 9, most errors occurred in the pluvial class. Compared with fluvial and coastal floods, pluvial floods were more difficult to identify reliably. This is likely because pluvial inundation is often fragmented, localised, and strongly influenced by urban drainage conditions, making its spatial signature less visually separable from other flood types in some scenes.

Classification was performed by generating patch-level predictions and aggregating them to the scene level using majority voting. Therefore, the reported accuracy reflects the model's ability to identify the dominant flood-generating mechanism of each scene, rather than pixel-level process differentiation. While the results indicate clear separability among dominant flood types in the WFT dataset, performance may be less stable in compound or ambiguous events where multiple flood mechanisms coexist.

We further examined whether errors in the U-Net-derived flood masks propagated through the WCR features and affected flood type classification performance. Table 2 quantifies the effect of replacing reference-mask-derived WCR with U-Net-predicted-mask-derived WCR. Performance was more sensitive to the WCR source for RGB + WCR models without DEM, with mean OA and macro-F1 decreasing from 0.802 to 0.757 and from 0.511 to 0.478, respectively. When DEM was included, the corresponding decreases were smaller, from 0.857 to 0.848 for mean OA and from 0.594 to 0.588 for mean macro-F1. Notably, the final Inception-ResNet configuration (RGB + DEM + WCR) achieved identical OA (0.933) and macro-F1 (0.730) under both WCR sources.

Beyond classification accuracy and macro-F1, computational efficiency is important for practical deployment. To assess model complexity, we compared the number of parameters and floating-point operations (FLOPs) across the six CNN architectures (Fig. 10).

Inception-ResNet achieved the highest macro-F1 score while maintaining moderate model complexity, with 115.97 million parameters and 60.64 million FLOPs. In contrast, VGG19 required substantially more parameters and computational cost without achieving better classification performance. This favourable performance–complexity trade-off supports the selection of Inception-ResNet as the final flood type classifier. On the WFT training set, each model required approximately 2 h for training on average, and the testing procedure produced classification outputs within 5 min on the specified GPU environment. These results suggest that the tested CNN models have potential for timely flood type classification, although operational deployment would still depend on data availability, preprocessing time, and computing infrastructure.

Based on the benchmarking and ablation results, Inception-ResNet was selected as the backbone of the final flood type classification model. The final classifier incorporated RGB imagery, DEM information,

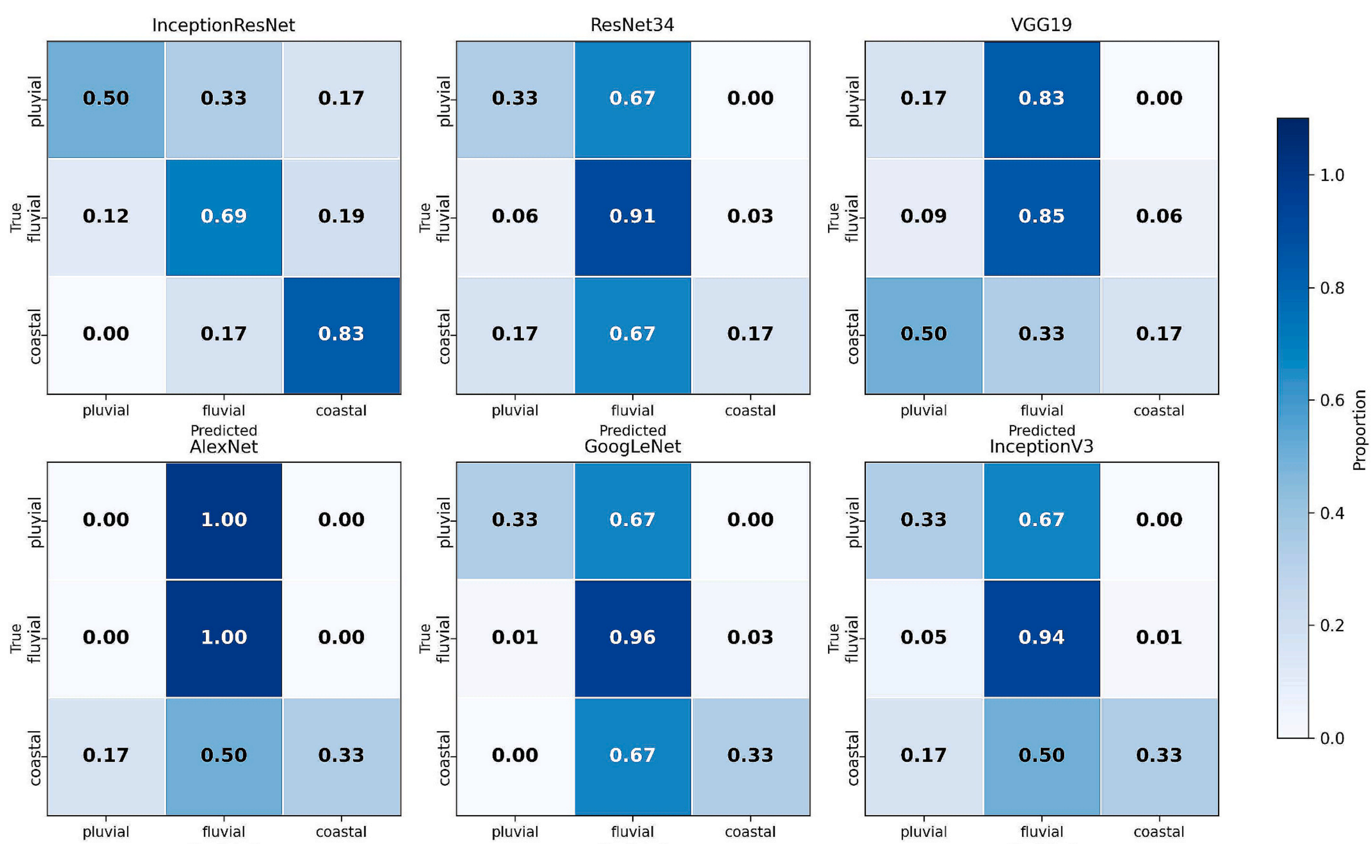

Fig. 9. Confusion matrices of baseline CNN models for the flood type classification stage.

Table 2
Scene-level flood type classification performance using WCR derived from reference flood masks and U-Net-predicted flood masks.

Model	Input	OA, reference	OA, U-Net	ΔOA	Macro-F1, reference	Macro-F1, U-Net	ΔMacro-F1
Inception-ResNet	RGB + WCR	0.800	0.722	−0.078	0.612	0.481	−0.131
	RGB + DEM + WCR	0.933	0.933	0.000	0.730	0.730	0.000
ResNet34	RGB + WCR	0.867	0.867	0.000	0.406	0.406	0.000
	RGB + DEM + WCR	0.789	0.767	−0.022	0.540	0.516	−0.024
VGG19	RGB + WCR	0.867	0.844	−0.022	0.525	0.538	+0.013
	RGB + DEM + WCR	0.878	0.878	0.000	0.511	0.511	0.000
AlexNet	RGB + WCR	0.822	0.800	−0.022	0.602	0.614	+0.012
	RGB + DEM + WCR	0.833	0.844	+0.011	0.634	0.680	+0.046
GoogLeNet	RGB + WCR	0.656	0.567	−0.089	0.444	0.359	−0.085
	RGB + DEM + WCR	0.822	0.811	−0.011	0.481	0.467	−0.014
InceptionV3	RGB + WCR	0.800	0.744	−0.056	0.476	0.471	−0.005
	RGB + DEM + WCR	0.889	0.856	−0.033	0.670	0.626	−0.044

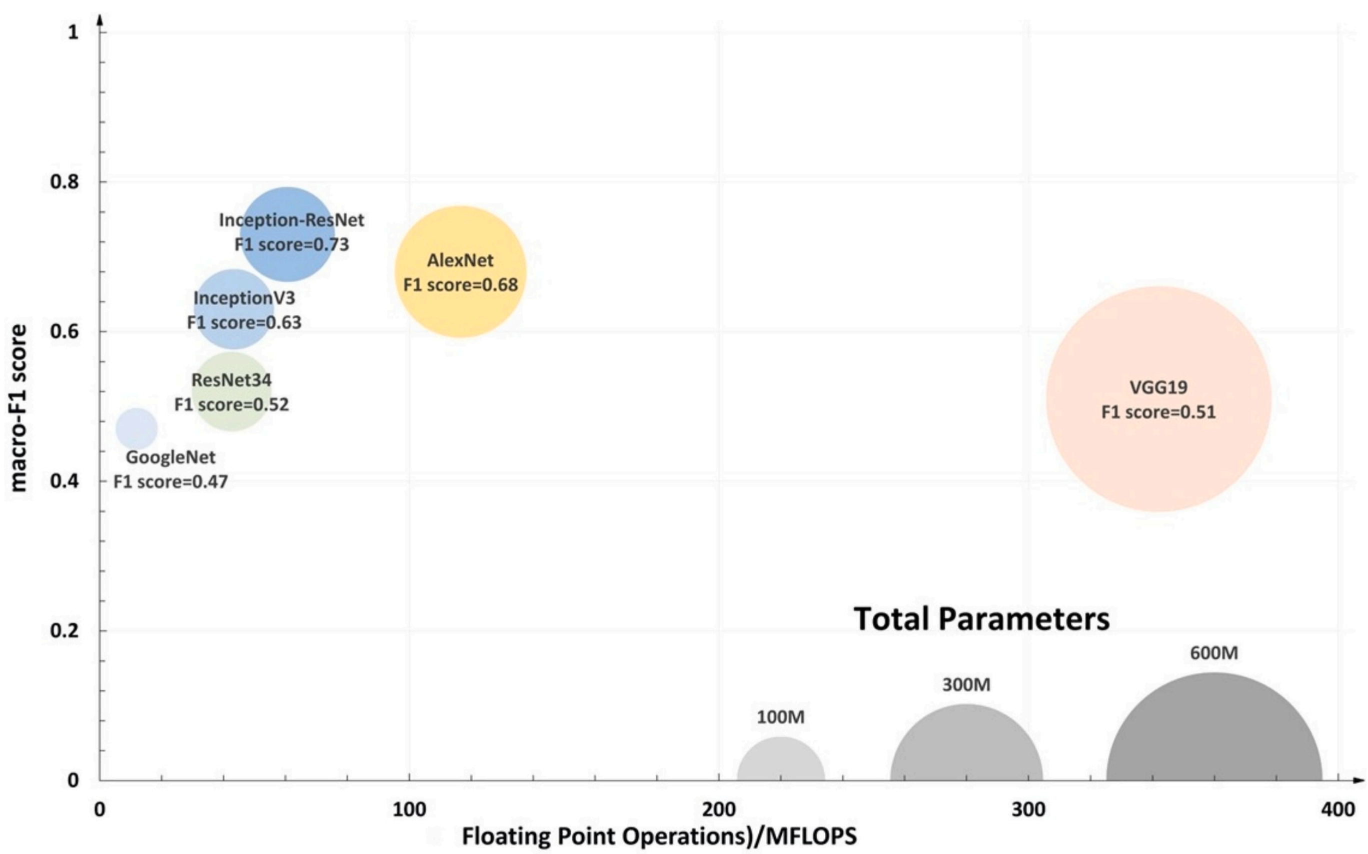

Fig. 10. Comparison of the six CNN architectures adopted for the flood classification stage. Circle size indicates the total number of parameters in each network.

and WCR features to integrate spectral, topographic, and water-connectivity information. Combined with the U-Net segmentation model, the proposed multi-CNN workflow jointly produces flood extent maps and scene-level flood type labels, providing a mechanism-aware extension to conventional satellite-based flood mapping.

6. Discussion and conclusions

6.1. Innovation and advances

Most operational flood mapping systems remain extent-centric, providing spatial delineation of inundation while offering limited insight into the underlying flood-generating mechanisms (Lee et al., 2020; Nelson-Mercer et al., 2025). Although such products are valuable for situational awareness, they do not distinguish explicitly between fluvial, pluvial, and coastal processes, which are associated with different hydrodynamic behaviour, contamination pathways, and recovery trajectories. This research advances beyond extent-only representations by integrating flood type attribution directly into the flood mapping workflow, enabling mechanism-aware flood products derived from remote sensing imagery and ancillary geospatial data. By embedding flood type classification within the mapping pipeline, the proposed framework reduces reliance on hydrometeorological driver data that are often incomplete, delayed, or spatially sparse. The observation-driven attribution strategy provides a complementary perspective to conventional driver-based inference, particularly in mechanism-complex environments such as estuarine deltas, coastal cities, and urban river corridors.

The use of a U-Net backbone provided a data-efficient basis for inundation delineation and supported the generation of flood masks for downstream connectivity analysis. The incorporation of DEM and Water Connectivity Ratio (WCR) features allowed the classification model to use not only spectral information from Sentinel-2 imagery, but also terrain context and the spatial relationship between inundation and permanent inland or marine water bodies. This is important because flood type is rarely determined by image texture alone; it is also shaped by topographic setting and hydrological connectivity. The benchmarking of multiple CNN architectures further clarified how network design influences the recognition of inundation morphology and spatial context. Among the tested models, Inception-ResNet achieved the strongest overall performance, indicating that the combination of multi-scale feature extraction and residual learning is well suited to scene-level flood type classification (Zhao et al., 2020; Hashemi-Beni and Gebrehiwot, 2021).The proposed workflow therefore provides an integrated approach for jointly producing flood extent maps and dominant flood type labels, extending conventional satellite-based flood mapping toward more process-relevant flood information.

6.2. Limitations and future research

Despite the promising results, several limitations warrant further research.

First, although the proposed framework incorporates DEM and water connectivity information, the image-based input is currently limited to Sentinel-2 RGB bands. RGB imagery provides useful spectral and contextual information, but it is vulnerable to cloud cover and atmospheric interference, which commonly accompany extreme flood events. In addition, the use of only three visible bands may limit the ability of the model to capture more detailed water, vegetation, and soil moisture characteristics. Future research could incorporate additional remote sensing data sources, such as near-infrared bands and synthetic aperture radar (SAR), to support more comprehensive multi-source feature fusion and improve robustness under cloudy or low-visibility conditions (Torres et al., 2012; Pulvirenti et al., 2016; Cohen et al., 2019; Zhang et al., 2023; Zhao et al., 2021).

Second, this study used FathomDEM to represent the static surface topographic context of each flood scene. While elevation information is useful for flood type attribution, static DEM data cannot describe dynamic floodwater surface elevation or water depth, both of which are important for understanding flood processes and impacts. Future work could explore the integration of SWOT observations to provide additional information on floodwater surface elevation and support more physically informative flood type attribution (Xu et al., 2026; Yao et al., 2025, Yao et al., 2026).

Third, temporal resolution remains constrained by satellite revisit cycles. Short-duration or rapidly evolving flood events may require near-real-time monitoring at sub-daily scales, particularly for pluvial flooding and compound coastal events where inundation extent can change rapidly. Future research could integrate higher-frequency data sources such as UAV imagery and geostationary satellite observations to improve temporal responsiveness. Furthermore, future work should also explore multi-temporal fusion strategies that combine the spatial detail of UAV or high-resolution satellite imagery with the frequent revisit capability of geostationary observations. Such integration could improve the ability of the model to capture flood dynamics, update flood type attribution during event evolution, and support more timely emergency response. Additionally, leveraging pre-trained vision foundation models may reduce training time and facilitate rapid adaptation to new events.

Fourth, segmentation uncertainty may propagate into flood type classification through the WCR features. Because WCR is calculated from flood extent masks, missed or false flood detections may alter the estimated connectivity between inundation and permanent inland or marine water bodies, thereby affecting flood type attribution. In this study, WCR features used for classifier training were derived from ground-truth flood masks to avoid confounding classifier learning with segmentation errors, while end-to-end workflow inference relies on U-Net-predicted flood masks. The quantitative comparison showed that replacing ground-truth-mask-derived WCR with U-Net-predicted-mask-derived WCR affected some classifier configurations, particularly those without DEM. Future work should evaluate this robustness across larger datasets and a wider range of segmentation qualities.

Fifth, although this study refines flood extent mapping by excluding permanent water bodies and focuses on dominant flood type attribution, it does not explicitly identify the land-cover or exposure classes affected by flooding. Information on whether flooded areas correspond to cropland, built-up areas, roads, vegetation, or other surface classes would be valuable for flood damage assessment and emergency response prioritisation. The refined flood extent layer produced in this study can be overlaid with land-cover, infrastructure, or exposure datasets in future work to identify affected classes and support more impact-oriented flood assessment.

Finally, flood type attribution in this study was conducted at the scene level using majority voting across image patches. While this strategy improves classification stability, it is sensitive to patch delineation and assumes a dominant mechanism within each scene. Compound flooding processes or mixed-driver events may therefore not be fully captured under this formulation. In addition, although scene-level partitioning is appropriate for evaluating classification across local flood scenes, it does not fully test transferability to completely unseen flood events. Future research should implement strict event-level and region-independent validation, as well as pixel-level or region-level mechanism mapping, to better assess generalisation and represent uncertainty in mixed flood processes.

6.3. Conclusions

This study demonstrates the feasibility of attributing dominant flood types at the scene level using remote sensing imagery and ancillary geospatial data, and that flood type attribution can be integrated into a satellite-based flood mapping workflow. By combining U-Net-based flood extent segmentation with CNN-based scene-level flood type classification, the proposed multi-CNN framework produces both inundation extent and flood type information within a unified workflow.

The results show that spatial morphology, terrain context, and water connectivity provide useful cues for distinguishing fluvial, pluvial, and coastal flooding under dominant-mechanism conditions. The WFT dataset, containing 464 flood scenes from 120 flood events across 48 countries, supports systematic evaluation of this task across diverse geographic settings. Among the tested architectures, Inception-ResNet provided the best overall balance between classification performance and model complexity.

Overall, the proposed framework offers a scalable, observation-driven complement to hydrometeorological flood classification approaches. By enriching extent-based flood maps with dominant flood type information, it provides a more informative representation of flood events that may support emergency response, impact interpretation, and post-disaster assessment. Further validation under compound and mechanism-complex conditions, together with multi-sensor and probabilistic extensions, will be important for advancing toward more adaptive and reliable flood intelligence systems.

CRediT authorship contribution statement

Ziming Wang: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Jeffrey Neal:** Writing – review & editing, Visualization, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Peter M. Atkinson:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Ce Zhang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors wish to thank the financial support by the UK Research and Innovation (UKRI) [grant number EP/Y002490/1] and the University of Bristol.

Data availability

The World Flood Type (WFT) dataset generated and analysed in this study has been deposited in Zenodo and is openly available at 10.5281/zenodo.20225841. The preprocessing scripts and model training code used in this study have also been deposited in Zenodo and are openly available at 10.5281/zenodo.20250106.

References

Abid, N., Shahzad, M., Malik, M.I., Schwanecke, U., Ulges, A., Kovács, G., Shafait, F., 2021. UCL: unsupervised curriculum learning for water body classification from remote sensing imagery. Int. J. Appl. Earth Obs. Geoinf. 105, 102568. https://doi.org/10.1016/J.JAG.2021.102568.

Abu, I.-O., Ibebuchi, C.C., 2025. Risk assessment of the 2022 Nigerian flood event using remote sensing products and climate data. Remote Sens. (Basel) 17. https://doi.org/10.3390/rs17111814.

Altenau, E.H., Pavelsky, T.M., Durand, M.T., Yang, X., de Frasson, R.P.M., Bendezu, L., 2021. The surface water and ocean topography (SWOT) mission river database (SWORD): a global river network for satellite data products. Water Resour. Res. 57, e2021WR030054. https://doi.org/10.1029/2021WR030054.

Bai, Y., Yu, Z., Yajing, S., Xinrong, Z., Yuan, X., 2022. Deep learning in different remote sensing image categories and applications: status and prospects. Int. J. Remote Sens. 43, 1800–1847. https://doi.org/10.1080/01431161.2022.2048319.

Barui, I., Bhakta, S., Ghosh, K., 2025. Storm surge-induced soil salinization and its impact on agriculture in the coastal area of the Indian Sundarban. Environ. Dev. 101250. https://doi.org/10.1016/J.ENVDEV.2025.101250.

Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R., 2022. Deep learning methods for flood mapping: a review of existing applications and future research directions. Hydrol. Earth Syst. Sci. 26, 4345–4378. https://doi.org/10.5194/hess-26-4345-2022.

Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., Hostert, P., 2023. In-season and dynamic crop mapping using 3D convolution neural networks and sentinel-2 time series. ISPRS J. Photogramm. Remote Sens. 195, 335–352. https://doi.org/10.1016/j.rse.2021.112831.

Brunner, M.I., Slater, L., Tallaksen, L.M., Clark, M., 2021. Challenges in modeling and predicting floods and droughts: a review. WIREs Water 8, e1520.

Call, B.C., Belmont, P., Schmidt, J.C., Wilcock, P.R., 2017. Changes in floodplain inundation under nonstationary hydrology for an adjustable, alluvial river channel. Water Resour. Res. 53, 3811–3834. https://doi.org/10.1002/2016WR020277.

Chen, W., Zhu, Y., Han, W., Liu, D., Mo, G., Xing, Z., 2026. Sen2GF3Floods: a benchmark multi-source flood dataset with dual-temporal and active learning annotation. Sci. Data 13, 540. https://doi.org/10.1038/s41597-026-06929-6.

Cohen, S., Raney, A., Munasinghe, D., Loftis, J.D., Molthan, A., Bell, J., Rogers, L., Galantowicz, J., Brakenridge, G.R., Kettner, A.J., Huang, Y.-F., Tsang, Y.-P., 2019. The floodwater depth estimation tool (FwDET v2.0) for improved remote sensing analysis of coastal flooding. Nat. Hazards Earth Syst. Sci. 19, 2053–2065. https://doi.org/10.5194/nhess-19-2053-2019.

El Khalki, E.M., Tramblay, Y., Massari, C., Brocca, L., Simonneaux, V., Gascoin, S., Saidi, M.E.M., 2020. Challenges in flood modeling over data-scarce regions: how to exploit globally available soil moisture products to estimate antecedent soil wetness conditions in Morocco. Nat. Hazards Earth Syst. Sci. 20, 2591–2607. https://doi.org/10.5194/nhess-20-2591-2020.

Falconer, R.H., Cobby, D., Smyth, P., Astle, G., Dent, J., Golding, B., 2009. Pluvial flooding: new approaches in flood warning, mapping and risk management. J. Flood Risk Manag. 2, 198–208. https://doi.org/10.1111/j.1753-318X.2009.01034.x.

Farhadi, H., Ebadi, H., Kiani, A., Asgary, A., 2024. Near real-time flood monitoring using multi-sensor optical imagery and machine learning by GEE: an automatic feature-based multi-class classification approach. Remote Sens. Basel 16, 4454. https://doi.org/10.3390/rs16234454.

Fernandes, J.N., 2025. Core challenges for sustainable river flood management. Sustainability 17. https://doi.org/10.3390/su172410981.

Green, J., Haigh, I.D., Quinn, N., Neal, J., Wahl, T., Wood, M., Eilander, D., de Ruiter, M., Ward, P., Camus, P., 2025. Review article: a comprehensive review of compound flooding literature with a focus on coastal and estuarine regions. Nat. Hazards Earth Syst. Sci. 25, 747–816. https://doi.org/10.5194/nhess-25-747-2025.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 770–778.

He, Y., Wang, Z., Cheng, H., Ding, W., 2026. Multi-scale feature fusion and uncertainty quantification in streamflow prediction: a temporal convolutional network approach with hybrid denoising. Environ. Model. Software 198, 106879. https://doi.org/10.1016/J.ENVSOFT.2026.106879.

Hou, J., Li, X., Bai, G., Wang, X., Zhang, Z., Yang, L., Du, Y., Ma, Y., Fu, D., Zhang, X., 2021. A deep learning technique based flood propagation experiment. J. Flood Risk Manag. 14, e12718. https://doi.org/10.1111/jfr3.12718.

Huang, X., Wang, C., Li, Z., 2018. A near real-time flood-mapping approach by integrating social media and post-event satellite imagery. Ann. GIS 24, 113–123. https://doi.org/10.1080/19475683.2018.1450787.

Jo, J., Kim, S., Kim, J., Mori, N., 2025. Coastal and pluvial compound flooding of surge, wave, tide, rainfall-runoff, sewer flow, and seawall failure at coastal urban areas. J. Hydrol. Amst) 661, 133698. https://doi.org/10.1016/j.jhydrol.2025.133698.

Konapala, G., Kumar, S.V., Khalique Ahmad, S., 2021. Exploring Sentinel-1 and Sentinel-2 diversity for flood inundation mapping using deep learning. ISPRS J. Photogramm. Remote Sens. 180, 163–173. https://doi.org/10.1016/j.isprsjprs.2021.08.016.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira, F., Burges, C.J., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems. Curran Associates Inc.

Lane, K., Charles-Guzman, K., Wheeler, K., Abid, Z., Graber, N., Matte, T., 2013. Health effects of coastal storms and flooding in urban areas: a review and vulnerability assessment. J. Environ. Public Health 2013, 913064. https://doi.org/10.1155/2013/913064.

Le Roy, S., Pedreros, R., André, C., Paris, F., Lecacheux, S., Marche, F., Vinchon, C., 2015. Coastal flooding of urban areas by overtopping: dynamic modelling application to the Johanna storm (2008) in Gâvres (France). Nat. Hazards Earth Syst. Sci. 15, 2497–2510. https://doi.org/10.5194/nhess-15-2497-2015.

Lee, J., Perera, D., Glickman, T., Taing, L., 2020. Water-related disasters and their health impacts: a global review. Prog. Disaster Sci. 8, 100123. https://doi.org/10.1016/J.PDISAS.2020.100123.

Löwe, R., Böhm, J., Jensen, D.G., Leandro, J., Rasmussen, S.H., 2021. U-FLOOD – Topographic deep learning for predicting urban pluvial flood water depth. J. Hydrol. (Amst) 603, 126898. https://doi.org/10.1016/J.JHYDROL.2021.126898.

Mateo-Garcia, G., Veitch-Michaelis, J., Smith, L., Oprea, S.V., Schumann, G., Gal, Y., Baydin, A.G., Backes, D., 2021. Towards global flood mapping onboard low cost satellites with machine learning. Sci. Rep. 11, 7249. https://doi.org/10.1038/s41598-021-86650-z.

Merz, B., Hall, J., Disse, M., Schumann, A., 2010. Fluvial flood risk management in a changing world. Nat. Hazards Earth Syst. Sci. 10, 509–527. https://doi.org/10.5194/nhess-10-509-2010.

Moftakhari, H.R., Salvadori, G., AghaKouchak, A., Sanders, B.F., Matthew, R.A., 2017. Compounding effects of sea level rise and fluvial flooding. Proc. Natl. Acad. Sci. 114, 9785–9790. https://doi.org/10.1073/pnas.1620325114.

Muñoz, D.F., Muñoz, P., Moftakhari, H., Moradkhani, H., 2021. From local to regional compound flood mapping with deep learning and data fusion techniques. Sci. Total Environ. 782, 146927. https://doi.org/10.1016/J.SCITOTENV.2021.146927.

Nelson-Mercer, B., Kim, T., Tran, V.N., Ivanov, V., 2025. Pluvial flood impacts and policyholder responses throughout the United States. npj Nat. Hazards 2, 8. https://doi.org/10.1038/s44304-025-00058-7.

Panahi, M., Jaafari, A., Shirzadi, A., Shahabi, H., Rahmati, O., Omidvar, E., Lee, S., Bui, D.T., 2021. Deep learning neural networks for spatially explicit prediction of flash flood probability. Geosci. Front. 12, 101076. https://doi.org/10.1016/J.GSF.2020.09.007.

Peng, B., Huang, Q., Vongkusolkit, J., Gao, S., Wright, D.B., Fang, Z.N., Qiang, Y., 2021. Urban flood mapping with bitemporal multispectral imagery via a self-supervised learning framework. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 2001–2016. https://doi.org/10.1109/JSTARS.2020.3047677.

Portalés-Julià, E., Mateo-García, G., Purcell, C., Gómez-Chova, L., 2023. Global flood extent segmentation in optical satellite images. Sci. Rep. 13, 20316. https://doi.org/10.1038/s41598-023-47595-7.

Pulvirenti, L., Chini, M., Pierdicca, N., Boni, G., 2016. Use of SAR data for detecting floodwater in urban and agricultural areas: the role of the interferometric coherence. IEEE Trans. Geosci. Remote Sens. 54, 1532–1544. https://doi.org/10.1109/TGRS.2015.2482001.

Rentschler, J., Avner, P., Marconcini, M., Su, R., Strano, E., Vousdoukas, M., Hallegatte, S., 2023. Global evidence of rapid urban growth in flood zones since 1985. Nature 622, 87–92. https://doi.org/10.1038/s41586-023-06468-9.

Olafand, Ronneberger, Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. In: Nassir, N., Hornegger, J., WM, W., AF, F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.

Rosenzweig, B.R., McPhillips, L., Chang, H., Cheng, C., Welty, C., Matsler, M., Iwaniec, D., Davidson, C.I., 2018. Pluvial flood risk and opportunities for resilience. WIREs Water 5, e1302.

Roth, F., Bauer-Marschallinger, B., Tupas, M.E., Reimer, C., Salamon, P., Wagner, W., 2023. Sentinel-1-based analysis of the severe flood over Pakistan 2022. Nat. Hazards Earth Syst. Sci. 23, 3305–3317. https://doi.org/10.5194/nhess-23-3305-2023.

Schumann, G., Giustarini, L., Tarpanelli, A., Jarihani, B., Martinis, S., 2022. Flood modeling and prediction using earth observation data. Surv. Geophys. https://doi.org/10.1007/s10712-022-09751-y.

Seydi, S.T., Amani, M., Ghorbanian, A., 2022. A dual attention convolutional neural network for crop classification using time-series sentinel-2 imagery. Remote Sens. (Basel) 14. https://doi.org/10.3390/rs14030498.

Shahapure, S., Eldho, T., Rao, E., 2010. Coastal Urban Flood simulation using FEM, GIS and Remote Sensing. Water Resour. Manag. 24, 3615–3640. https://doi.org/10.1007/s11269-010-9623-y.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2016a. Inception-v4, inception-resnet and the impact of residual connections on learning. ArXiv abs/1602.07261.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016b. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826.

Tai, C.V., Kim, D., Kronenberg, R., Vorobevskii, I., Luong, T.T., 2025. Beneath the surface: exploring relationship between pluvial floods and income disparities for residential basements in Seoul, South Korea. Int. J. Disaster Risk Reduct. 123, 105501. https://doi.org/10.1016/j.ijdrr.2025.105501.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B.Ö., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24. https://doi.org/10.1016/J.RSE.2011.05.028.

Uhe, P., Lucas, C., Hawker, L., Brine, M., Wilkinson, H., Cooper, A., Saoulis, A.A., Savage, J., Sampson, C., 2025. FathomDEM: an improved global terrain map using a hybrid vision transformer model. Environ. Res. Lett. 20, 34002. https://doi.org/10.1088/1748-9326/ada972.

Vázquez-Tarrío, D., Ruiz-Villanueva, V., Garrote, J., Benito, G., Calle, M., Lucía, A., Díez-Herrero, A., 2024. Effects of sediment transport on flood hazards: lessons learned and remaining challenges. Geomorphology 446, 108976. https://doi.org/10.1016/j.geomorph.2023.108976.

Wang, J., Pottier, C., Cazals, C., Battude, M., Sheng, Y., Song, C., Sikder, M.S., Yang, X., Ke, L., Delhoume, M., Gosset, M., Oliveira, R.R.A., Grippa, M., Girard, F., Allen, G.H., Xu, X., Zhu, X., Biancamaria, S., Smith, L.C., Crétaux, J.-F., Pavelsky, T.M., 2025. The surface water and ocean topography mission (SWOT) prior lake database (PLD): lake mask and operational auxiliaries. Water Resour. Res. 61, e2023WR036896. https://doi.org/10.1029/2023WR036896.

Wang, Y., Fang, Z., Hong, H., Peng, L., 2020. Flood susceptibility mapping using convolutional neural network frameworks. J. Hydrol. (Amst) 582, 124482. https://doi.org/10.1016/J.JHYDROL.2019.124482.

Welch, K., Lambert, L.H., Lambert, D.M., Kenkel, P., 2022. Flood-induced disruption of an inland waterway transportation system and regional economic impacts. Water (Basel) 14. https://doi.org/10.3390/w14050753.

Xu, N., Zhang, Z., Xu, H., Yao, J., Lu, H., Cai, W., Ou, Y., Luan, H., Gong, P., Tu, W., Li, Q., 2026. Multifaceted assessment of recent saltwater intrusion along China's coasts. Sci. Bull. (Beijing). https://doi.org/10.1016/J.SCIB.2026.02.021.

Yao, J., Wang, M., Xu, N., Liu, T., Mo, F., Li, X., Cao, Y., Chen, K., Sun, J., Lu, H., 2025. 3-D Flood mapping from SWOT observations during extreme rainfall: a case study of Gangnan Reservoir. Int. J. Digit. Earth 18, 2544916. https://doi.org/10.1080/17538947.2025.2544916.

Yao, J., Zhao, Z., Pan, M., Yang, Y., Li, X., Liu, T., Wen, Z., Xu, N., Lu, H., 2026. SWOT-based water surface elevation observations improve flood modeling of the 25·7 Miyun reservoir basin extreme rainfall event. npj Nat. Hazards 3, 12. https://doi.org/10.1038/s44304-026-00176-w.

Zhang, Q., Wu, Z., Zhang, H., Dalla Fontana, G., Tarolli, P., 2020. Identifying dominant factors of waterlogging events in metropolitan coastal cities: the case study of Guangzhou. China. J. Environ. Manage. 271, 110951. https://doi.org/10.1016/j.jenvman.2020.110951.

Zhang, Y., Liu, P., Chen, L., Xu, M., Guo, X., Zhao, L., 2023. A new multi-source remote sensing image sample dataset with high resolution for flood area extraction: GF-FloodNet. Int. J. Digit. Earth 16, 2522–2554. https://doi.org/10.1080/17538947.2023.2230978.

Zhao, J., Pelich, R., Hostache, R., Matgen, P., Wagner, W., Chini, M., 2021. A large-scale 2005–2012 flood map record derived from ENVISAT-ASAR data: United Kingdom as a test case. Remote Sens. Environ. 256, 112338. https://doi.org/10.1016/J.RSE.2021.112338.

Zhao, S., Tu, K., Ye, S., Tang, H., Hu, Y., Xie, C., 2023. Land use and land cover classification meets deep learning: a review. Sensors 23. https://doi.org/10.3390/s23218966.