



OPEN ACCESS

EDITED BY

Miao Zhang,
Shaanxi Normal University, China

REVIEWED BY

Vladimir Razlutskiy,
State Scientific and Production Amalgamation
"Scientific-practical center of the National
Academy of Sciences of Belarus for biological
resources", Belarus
Sebastian Kutyla,
Institute of Environmental
Protection – National Research Institute,
Poland

*CORRESPONDENCE

Ganggang Zuo,
✉ zgg@xaut.edu.cn
Jiancang Xie,
✉ jcxie@xaut.edu.cn

RECEIVED 02 July 2025

ACCEPTED 08 September 2025

PUBLISHED 23 September 2025

CITATION

Song X, Zuo G, Wang X and Xie J (2025)
Automatic recognition of environmental
hazards in river and lake ecosystems using
deep learning.
Front. Environ. Sci. 13:1657930.
doi: 10.3389/fenvs.2025.1657930

COPYRIGHT

© 2025 Song, Zuo, Wang and Xie. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction in
other forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Automatic recognition of environmental hazards in river and lake ecosystems using deep learning

Xueying Song¹, Ganggang Zuo^{2*}, Xiaofeng Wang³ and
Jiancang Xie^{2*}

¹School of Civil Engineering and Architecture, Xi'an University of Technology, Xi'an, Shaanxi, China,

²School of Water Resources and Hydro-Electric Engineering, Xi'an University of Technology, Xi'an, Shaanxi, China, ³Apparel & Art Design College, Xi'an Polytechnic University, Xi'an, Shaanxi, China

Introduction: Accurate identification of environmental issues in river and lake ecosystems is essential for the protection, management, and sustainable use of water resources. Traditional inspection-based approaches are limited by their extensive spatial scope, high labor demands, prolonged execution time, and increased likelihood of overlooking hazards.

Methods: To overcome these limitations, this study investigates intelligent methods for detecting environmental hazards in river and lake settings. Images representing 12 common types of water-related hazards were collected. Using image augmentation techniques, including rotation, transformation, and annotation, a dataset comprising over 1,500 samples of river and lake environmental hazards was constructed. An intelligent recognition model was then developed based on the YOLOv11 algorithm, incorporating transfer learning techniques to enable the detection of pollution categories, pollutant types, sewage outfalls, and shoreline encroachments.

Results: The experimental results demonstrate that, with adequate training data, appropriate categorization, and accurate annotation, the proposed method achieves reliable performance, yielding a balanced F1 score of 0.72.

Discussion: This approach can be deployed on devices such as smartphones, cameras, and unmanned aerial vehicles, offering practical tools for water pollution surveillance, shoreline monitoring, and the broader management of aquatic ecosystems.

KEYWORDS

water environment, hazard identification, deep learning, object detection, transfer learning, YOLOv11 algorithm

1 Introduction

Water resources are of utmost importance for human survival and social development. Currently, problems such as industrial and domestic sewage discharge and non-point source pollution originating from arable land and artificial surfaces located in waterbody catchments have severely endangered their quality (Shakuli, 2021; Kutyla et al., 2024). These pressures accelerate natural eutrophication (Heathcote, 2013). Additional stressors, including overfishing, the construction of hydraulic engineering projects, and inappropriate operation and scheduling practices, have disrupted the ecological balance of river and lake systems, further compromising both the quantity and quality of water

resources (Lai et al., 2024). Disasters such as floods, droughts, sudden water surges, and bank collapses not only inflict substantial losses on nearby populations but also complicate the safe operation and regulation of water conservancy infrastructure, thereby posing risks to overall water resource security.

Accurate identification of existing issues in river and lake water systems is therefore essential to ensure the sustainable development, utilization, and protection of water resources. It also contributes significantly to the ecological security of aquatic environments, the clarification and enforcement of regulatory responsibilities, comprehensive monitoring of watershed pollution, and informed decision-making by local authorities. As human activity and urbanization continue to accelerate, river and lake environmental hazards are becoming more frequent, widespread, and severe. This growing urgency necessitates the development of effective technologies for the identification of such hazards to support resource management and ecological governance (Yan et al., 2025).

Due to the extensive spatial coverage of river and lake systems, effective inspection and monitoring remain challenging, and environmental hazards often go undetected. Rapid identification of these hazards continues to be a major technical bottleneck (Shi, 2017). Traditional detection methods typically rely on laboratory analysis, which demands highly skilled personnel and maintenance of sophisticated instrumentation. In field-based inspections, issues such as high professional thresholds, intensive workloads, and inadequate responsiveness persist (Wei et al., 2022; Liu, 2019; Ren et al., 2022; Lin, 2024).

With the rapid advancement of artificial intelligence (AI) and computer vision, numerous sectors have begun applying these technologies to address domain-specific problems (Aziz et al., 2020). For instance, image recognition in autonomous driving (Xu et al., 2024; Du et al., 2024; Li, 2022; Cai et al., 2020), image-based diagnostics in healthcare (Guzel et al., 2024; Zhou et al., 2023; Liu et al., 2024; Chen et al., 2022; Lee et al., 2024; Li et al., 2024; Ronneberger et al., 2015; Xu et al., 2023), and agricultural applications such as fruit classification and detection (Anand et al., 2019; Guzel et al., 2024; Zhou et al., 2023; Liu et al., 2024) have yielded notable results. In this context, the integration of AI and computer vision presents new opportunities for enhancing the efficiency and effectiveness of environmental hazard identification in river and lake ecosystems.

Image recognition and object detection technologies are central to feature extraction, classification, and pattern matching in visual data. These technologies significantly enhance the efficiency and accuracy of identifying environmental hazards in river and lake systems, while also accelerating computational processes and lowering operational costs (Wei et al., 2022; Zeng, 2024). Within the context of environmental monitoring, artificial intelligence (AI) holds considerable potential for analyzing the color of water bodies (Gao, 2023) and for detecting the types, locations, pollution levels, and discharge characteristics of contaminants with scientific precision and timeliness.

Since the introduction of AlexNet by Krizhevsky et al. (2012), convolutional neural networks (CNNs) have undergone substantial advancements. A notable milestone was the development of R-CNN in 2014, which marked significant progress in object detection algorithm research (Girshick et al., 2014). This was followed by Fast R-CNN (Girshick, 2015), which integrated classification and

detection using a multi-task loss function, thereby improving computational efficiency. In the same year, Ren et al. (2015) enhanced detection performance by sharing weights between Region Proposal Networks (RPN) and convolutional layers. Subsequent developments included the Single Shot MultiBox Detector (SSD) introduced by Liu et al. (2016), which offered improved speed over Faster R-CNN, and the one-stage detection algorithm You Only Look Once (YOLO) proposed by Redmon et al. (2016), which further increased real-time detection capabilities. The YOLO algorithm has since evolved through multiple iterations, reaching YOLOv11 as of 2024. This latest version provides faster detection speeds and higher accuracy, supporting a wide range of industry applications through advanced computer vision solutions.

Object detection algorithms based on computer vision have also been extensively applied in the water conservancy sector. Huang (2022) investigated defect detection in transmission lines of water infrastructure using unmanned aerial vehicle (UAV) imagery. Lu and Gao (2023) examined the role of image recognition in monitoring construction quality within water conservancy projects. Ding et al. (2024) proposed an improved underwater image recognition model based on the EfficientNet architecture, while Ge et al. (2024) developed five models for dam defect recognition. In the context of river and lake hazard identification, Shen (2022) employed deep learning to detect the “four disorder” problems, unauthorized occupation, mining, dumping, and construction, based on satellite remote sensing imagery. This approach enabled large-scale identification of issues within aquatic environments. Nonetheless, current research remains limited in two key areas. First, local-scale detection of water problems requires further investigation. Second, the scope of detectable targets in previous studies has been narrow, focusing mainly on structures such as greenhouses and buildings, and thus requires expansion.

The objective of this study is to assess the effectiveness of a deep learning-based object detection model (YOLOv11) in automatically identifying environmental hazards from visual data of river and lake ecosystems, thereby supporting surface water monitoring and evaluation. Photographic data depicting typical environmental hazards were collected and annotated to construct a benchmark dataset containing over 1,500 images. These images include representative cases such as discoloration from water pollution, garbage accumulation, floating fish, drainage outlets, sand yards, and unauthorized buildings. Utilizing the YOLOv11 algorithm in conjunction with transfer learning, a model for hazard identification was developed and evaluated for its performance. This study offers both a standard dataset and an intelligent detection approach applicable to *in situ* monitoring of river and lake water environments.

2 Data and methods

The identification of environmental hazards in river and lake systems involves the recognition, classification, and analysis of various ecological and hydrological issues. These hazards typically pertain to both “quantity” and “quality” aspects of water bodies, encompassing indicators such as water quality parameters, hydrological volumes, and forms of visible pollution.

TABLE 1 Categories and descriptions of target objects for river and lake environmental hazard detection.

Serial number	Identification target	Monitoring description	Number of samples
1	Algae Pollution	Green discoloration in river or lake water, with visible floating green algae	150
2	Oil Film Pollution	Surface signs of oil contamination in river or lake water	80
3	Red Pollution	Reddish coloration observed in the water body	150
4	Yellow Pollution	Yellowish discoloration present in the water	150
5	Foam Pollution	Presence of surface foam indicating pollution	150
6	Sewage	Blackish or grayish water coloration associated with sewage contamination	150
7	Garbage	Floating or shoreline debris in rivers and lakes	150
8	Fish	Floating dead fish on the water surface, typically in large numbers	80
9	Leaves	Floating leaves observed on the water surface	80
10	Sewage Outlet	Visible sewage or drainage pipes along the waterbody banks	150
11	Sand Yard	Evidence of sand mining along river or lake banks	150
12	Building	Presence of buildings adjacent to river or lake banks	150

2.1 Construction of the river and lake water environmental hazards dataset

2.1.1 Data collection

To support the task of environmental hazard detection, a dataset titled WATER-DET was compiled using relevant image data. The sources of these images included field photography, surveillance camera footage, and UAV-based aerial imaging. To ensure diversity within the dataset, images were collected across multiple types of water bodies (including rivers, lakes, and reservoirs), various categories of environmental hazards (e.g., pollutant types, sand mining operations, unauthorized structures), differing lighting conditions, and across multiple seasons.

Online sources were also used to supplement the dataset. Keyword-based searches were conducted to retrieve news articles, reports, and field imagery depicting relevant scenes. The selection process involved evaluating both the content of reports and expert input to identify representative instances of water-related hazards. These images were then annotated to ensure both representativeness and labeling accuracy, providing a reliable foundation for model training and validation.

2.1.2 Target classification

Target classification depends on the application context. This paper proposes three main scenarios. First, monitoring water color as a quality indicator that helps preliminarily identify the type of pollution. Second, monitoring floating objects on the surface of rivers and lakes to assess the ecosystem’s condition and detect pollution on the water and along the banks. Third, monitoring river courses in accordance with national regulations and water protection policies.

Following a systematic analysis, the identification tasks for common river and lake water hazards were categorized into four main detection types.

- Water Pollution Detection: Identifies pollution types based on water color and detects visible pollutants such as oil films, algae, and floating waste.

- Surface Floating Object Detection: Identifies objects floating on the water surface, including branches, leaves, debris along the banks, and dead fish.
- Sewage Outlet Detection: Detects pipelines and discharge outlets to assess their location and operating conditions.
- Detection of Regulatory Violation Targets (related to “Four Hazards” Elimination): Includes illegal sand mining within protected river zones, soil excavation in management areas, and unauthorized construction that obstructs flood pathways.

These target types were selected based on their direct environmental impact, alignment with regulatory priorities, and the technical feasibility of automated identification. A detailed breakdown of the classification scheme is presented in Table 1, which outlines the 12 selected target categories: (1) Algae pollution, (2) Oil film pollution, (3) Red-colored pollution, (4) Yellow-colored pollution, (5) Foam pollution, (6) Sewage (black or grey), (7) Garbage (on the water surface or along the bank), (8) Floating fish, (9) Floating leaves, (10) Sewage outlets, (11) Sand yards, and (12) Buildings (adjacent to riverbanks). These categories are mutually exclusive and precisely defined to ensure both the rationality and scalability of image classification.

2.1.3 Data annotation

The labelImg software was used for manual annotation of the collected images. Annotation included both the classification of environmental hazard types (e.g., pollution, physical obstructions) and the spatial information of targets, marked using rectangular bounding boxes. All annotations followed a standardized protocol to ensure quality. A cross-verification approach involving multiple annotators was employed to validate the labeling consistency and accuracy. Representative annotation samples are displayed in Figure 1, showing data collected from various water environments.

The annotation process was organized as a multi-step, collaborative workflow. Initially, the annotation framework and classification schema were established. Data were then named and organized according to predefined conventions. The

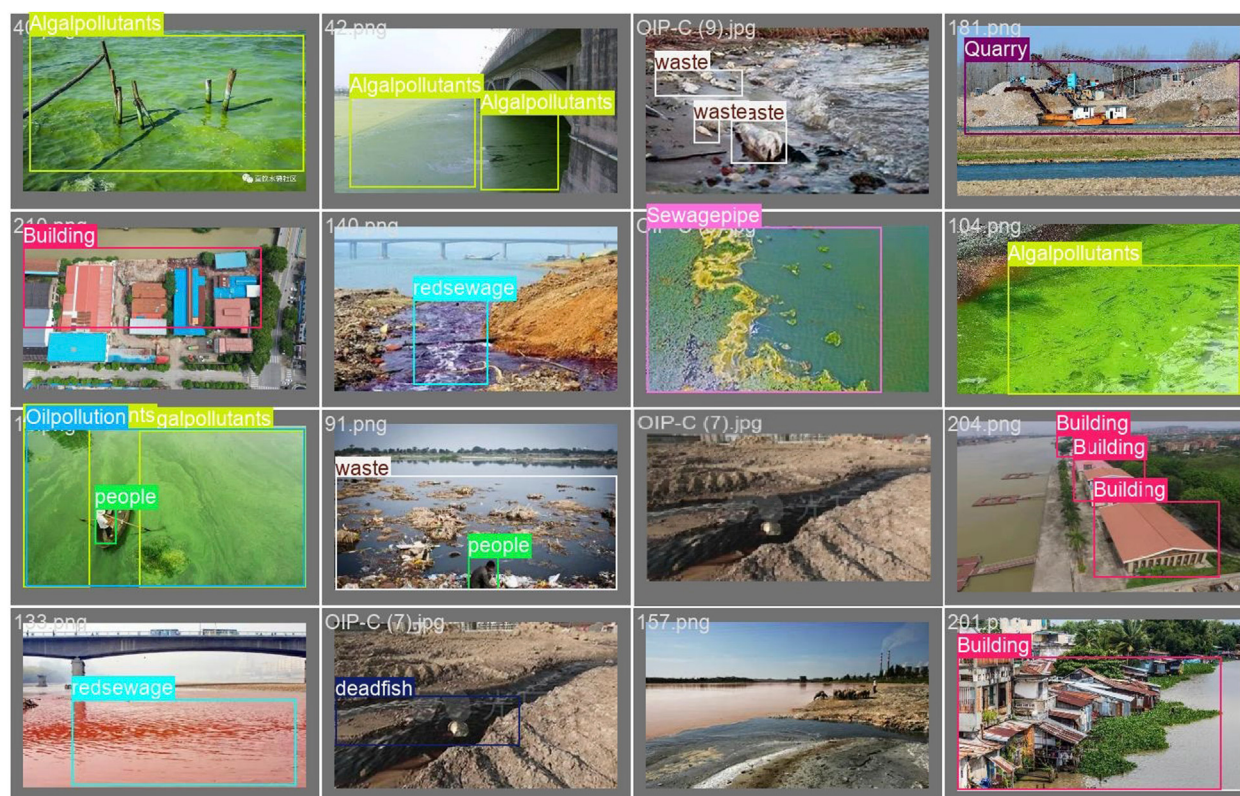


FIGURE 1
Sample images after data annotation.

workflow included task assignment by an administrator, annotation by trained annotators, and review by dedicated reviewers. This was followed by data augmentation and automated annotation processes, with a final round of review and dataset partitioning. Once validated, the annotated data were exported for use in model training.

The final dataset, named WATER-DET, comprises 1,500 RGB images depicting various river and lake environmental hazards. To accommodate the diversity of target categories, the test set size was slightly increased. The dataset was divided into training, validation, and test sets in a 7:1:2 ratio. Specifically, the training set includes 1,050 images, the validation set contains 150, and the test set includes 300 images. This partitioning strategy was designed to maintain data diversity and support the effective application of transfer learning in model development, validation, and evaluation.

2.2 Development of a transfer learning model for environmental hazard detection in river and lake systems based on YOLOv11

This study employs transfer learning using the YOLOv11 architecture to construct a model for detecting environmental hazards in river and lake ecosystems. The algorithm's core principles, network structure, and implementation process are detailed in the following subsections.

2.2.1 YOLOv11 algorithm principle

YOLOv11 incorporates an enhanced backbone and neck architecture to improve feature extraction and increase object detection accuracy. Through refined architectural design and advanced training strategies, the algorithm delivers faster processing speeds while maintaining detection precision. It also reduces parameter complexity, thereby improving computational efficiency. In addition to object detection, YOLOv11 supports a range of computer vision tasks, including instance segmentation, image classification, pose estimation, and oriented object detection.

Compared with earlier iterations such as YOLOv5 and YOLOv8, YOLOv11 introduces several notable innovations. It integrates the newly developed C3K2 module to enhance feature extraction efficiency and incorporates the Cross Stage Partial with Spatial Attention (C2PSA) mechanism to better emphasize relevant spatial regions. While retaining optimized elements such as the Spatial Pyramid Pooling-Fast (SPPF) module, YOLOv11 achieves superior accuracy, faster inference, and a reduction in parameter count. Its multi-task capability enables the execution of diverse computer vision functions within a unified and efficient framework.

2.2.1.1 Network architecture

YOLOv11 adopts a redesigned network architecture composed of three primary components: a backbone network, a neck network, and a head network. The backbone network is responsible for extracting features from the input image through a sequence of convolutional and pooling operations. These operations gradually

reduce the spatial resolution of the image while increasing the level of feature abstraction. The extracted features capture various aspects of objects in the image, such as edges, textures, and color patterns. Through continuous learning, the backbone network is capable of representing both low-level and high-level image features, thereby providing essential information for the subsequent detection process.

The neck network functions as an intermediary between the backbone and head networks. It further processes and integrates the features extracted by the backbone. This component applies specialized convolutional and pooling layers to adjust feature dimensions and scales, allowing for the effective combination of features across different hierarchical levels. As a result, the model can leverage both detailed spatial information and high-level semantic context, which improves overall detection accuracy.

The head network generates predictions for target objects based on the refined features provided by the neck. It outputs the classification, location, and confidence scores of detected objects. For each prediction, the model estimates the probability of each class along with the spatial coordinates of the object, typically represented by the bounding box parameters. These parameters include the center point, width, and height, all of which are computed relative to the original image dimensions.

The detection process in YOLOv11 includes grid division, bounding box prediction, and category classification. The input image is first divided into a set of grid cells, with each cell responsible for detecting objects that fall within its area. This approach enables localized image analysis, reduces computation, and enhances detection speed. For each grid cell, the model predicts multiple bounding boxes, each containing positional information relative to the cell. These predictions are transformed into absolute coordinates using predefined formulas, and refined during training to better approximate ground truth values. In addition, each grid cell estimates the object category based on local features and learned classification patterns. The model then outputs class probabilities, and the category with the highest probability is selected as the final classification result.

2.2.1.2 Loss function

YOLOv11 employs a composite loss function that integrates three components: bounding box regression loss (Box Loss), classification loss (CSL), and confidence loss (CFL). The Box Loss (BL) component optimizes the spatial discrepancy between the predicted bounding boxes and the corresponding ground truth boxes. The formula for BL is expressed as [Formula 1](#):

$$\begin{aligned} \text{BL} = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{\text{obj}} \left((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right) \\ & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{\text{obj}} \left(\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right) \end{aligned} \quad (1)$$

Here, S denotes the grid size, B represents the number of bounding boxes predicted per grid cell, and l_{ij}^{obj} indicates whether the j -th bounding box in the i -th grid cell is responsible for predicting the target. The variables x, y correspond to the center coordinates of the bounding box, while w, h represent its width and height. The coefficient λ_{coord} serves as a balancing factor to manage the relative contributions of different loss components.

The CSL quantifies the discrepancy between the predicted probability distribution and the actual class labels. It is calculated as [Formula 2](#):

$$\text{CSL} = \sum_{i=0}^{S^2} l_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \quad (2)$$

In this equation, $p_i(c)$ denotes the predicted probability that the object in the i -th grid cell belongs to category c , while $\hat{p}_i(c)$ represents the true label. The other variables follow the same definitions as above.

CFL evaluates the accuracy of the model's predictions regarding the presence or absence of a target within each grid cell. A high confidence value is expected when a target exists, while a low confidence is preferred when no object is present. This component enhances the model's ability to distinguish between object and background regions and reduces the likelihood of false positives. CFL is typically implemented using a variation of the cross-entropy loss, which is adjusted to focus more on difficult-to-classify samples. Its formulation is given by [Formula 3](#):

$$\begin{aligned} \text{CFL} = & - \sum_{i=1}^N \sum_{c=1}^C y_{ic} (\alpha (1 - p_{ic})^\gamma \log(p_{ic}) \\ & + (1 - \alpha) p_{ic} \log(1 - p_{ic})) \end{aligned} \quad (3)$$

Here, N is the number of training samples, C is the number of categories, y_{ic} represents the one-hot encoded ground truth label for the i -th sample, and p_{ic} is the model's predicted probability that the i -th sample belongs to class c . The parameter α is a balance factor controlling the emphasis on positive and negative samples, while γ is the focusing parameter used to emphasize hard examples during training.

2.2.2 Transfer learning strategy

This study applies a transfer learning approach based on fine-tuning to adapt a general-purpose YOLOv11 model for detecting environmental hazards in river and lake systems. Transfer learning methods generally fall into two categories: feature-based and model-based learning. The fine-tuning strategy used here belongs to the latter category. Its core principle involves retraining the final layers of a pre-trained model, typically the fully connected layers, while retaining the parameters of the earlier layers. This selective retraining enables the model to adapt more effectively to the new task by leveraging general feature representations learned from large-scale datasets.

2.2.3 Technical framework and model training

An intelligent identification model for detecting environmental hazards in river and lake water systems was developed using the deep learning framework TensorFlow. The model training and validation were implemented in Python within the Anaconda environment. The overall technical framework is illustrated in [Figure 2](#).

The transfer learning process consists of several stages. First, the collected samples of river and lake environmental hazards were reconstructed and divided into training, validation, and test subsets. The training set was used to fine-tune the parameters of the YOLOv11 model, while the validation set was employed to

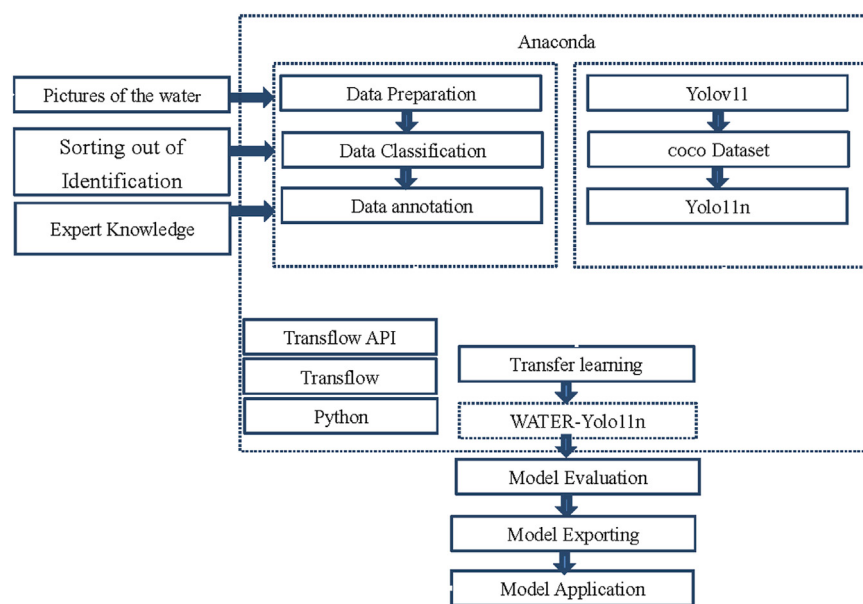


FIGURE 2
Technical framework and model training process.

optimize the hyperparameter configuration. Next, the test set was used to evaluate the fine-tuned YOLOv11 model and assess its detection precision. Finally, the trained model was exported, localized, and deployed within a practical water hazard identification system.

2.2.4 Model training strategy

Selection of Pre-trained Model: The YOLO11n model pre-trained on the COCO (Common Objects in Context) dataset was selected as the foundation for transfer learning. These pre-trained models provide robust feature representations learned from large-scale visual data, offering effective initial weights for the river and lake environmental hazard detection task.

Fine-tuning Strategy: To adapt the model to the specific classification requirements of this study, the final classification layer was replaced to match the number of target categories. A fine-tuning strategy was applied wherein the weights of lower convolutional layers were frozen, and only the upper layers, along with the newly added classification layer, were retrained. This approach preserves the general feature extraction capacity of the base model while reducing training time and computational cost.

Setting of Training Parameters: Hyperparameters such as the learning rate, batch size, and number of training epochs were configured using the YAML settings file provided by the YOLOv11 framework.

2.2.5 Model validation

The dataset was divided into training, validation, and test sets in a 7:1:2 ratio. Manual adjustments were applied to ensure an even distribution of various hazard categories across the subsets, thereby enhancing the generalizability of the model evaluation.

Performance metrics used in this study include Precision (P), Recall (R), F1-score, and Average Precision (AP). Precision

measures the proportion of correctly identified positive samples relative to all samples predicted as positive. Recall evaluates the proportion of true positives among all actual positive cases. Since an improvement in one metric may lead to a decline in the other, the F1-score, defined as the harmonic mean of precision and recall, provides a balanced assessment. Average Precision offers a summarization of model performance across different recall thresholds. The corresponding formulas are as follows as [Formula 4–7](#):

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \quad (5)$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (6)$$

$$AP = \frac{1}{N} \sum_{k=1}^N AP_k \quad (7)$$

In the equations above, TP (True Positives) represents correctly identified positive samples, FP (False Positives) indicates negative samples incorrectly classified as positive, and FN (False Negatives) denotes positive samples that were not identified by the model.

Model Adjustment and Optimization: Model adjustments and optimization were conducted iteratively based on performance metrics obtained from the validation set.

Adjusting Parameters Based on Model Performance Evaluation Indicators: Model performance was evaluated using indicators such as Average Precision (AP), recall, and precision. AP provides a comprehensive assessment by integrating both precision and recall across different confidence thresholds, thus reflecting overall detection effectiveness. Recall indicates the proportion of true positives successfully identified, while precision refers to the proportion of correct detections among all predicted positives.

These metrics directly reflect the model's ability to recognize various target types. If the model exhibits low recall for a particular category, it suggests that many true positives are being missed, requiring parameter adjustments to improve sensitivity to that category. Conversely, low precision indicates a high rate of false positives, suggesting that optimization is needed to reduce misclassification. For instance, in this study, the detection accuracy for floating fish was relatively low, likely due to the visual complexity of such targets. To improve performance, additional annotated samples of floating fish were introduced, enabling the model to learn more representative features and enhance prediction accuracy.

Adjusting Parameters Based on Overfitting and Underfitting Conditions: The training process also involved monitoring for signs of overfitting or underfitting by analyzing loss curves and validation performance. Overfitting was identified when the training loss continued to decline while validation loss increased and accuracy fluctuated. This indicated that the model was too complex and had begun memorizing noise and fine-grained details, reducing its generalization capacity. In such cases, regularization techniques such as L1 and L2 were applied to constrain model complexity, or the network architecture was simplified by reducing the number of parameters. Underfitting was diagnosed when both training and validation losses remained high and decreased slowly, implying that the model lacked the complexity to effectively learn from the data. To address this, the model architecture was enhanced by increasing the number of layers or neurons, and training parameters were adjusted, such as increasing the learning rate, to accelerate convergence and improve performance.

Adjusting According to the Influence of Parameters on the Model: Further optimization was carried out by analyzing the impact of training parameters on model behavior. The learning rate, which governs the magnitude of parameter updates, significantly affects convergence. An excessively high learning rate may cause the model to overshoot optimal values, leading to unstable loss patterns, while a rate that is too low slows convergence substantially. In practice, the learning rate was increased when convergence was too slow and reduced when the loss became unstable. Similarly, batch size influenced training efficiency and stability. Larger batch sizes contributed to more stable training dynamics but required greater memory resources, whereas smaller batch sizes allowed faster iterations but increased convergence variability. Therefore, the batch size was selected based on available computational resources and the observed training behavior.

In the experiments conducted in this study, overall model performance was improved by expanding the training dataset, fine-tuning training parameters, and modifying the model architecture to better suit the characteristics of river and lake water environmental hazards.

3 Experimental design and result analysis

3.1 Experimental environment

The experiments were conducted on a system running the Windows 10 operating system, equipped with 16 GB of memory.

The central processing unit (CPU) was an Intel Core i7-10700K with a base clock frequency of 3.8 GHz, and the graphics processing unit (GPU) was an NVIDIA GeForce RTX 3060. The TensorFlow deep learning framework, integrated through Anaconda, was employed within a Python 3.8.16 environment to support the implementation and execution of the YOLOv11 model.

3.2 Experimental setup

In view of computational and memory constraints, and to ensure that the model sufficiently learns the data features, the batch size was set to 16 images per iteration, with an input resolution of 160×160 pixels. The training process was executed over 200 epochs. Following each training cycle, the classification losses for both the training set (train/cls loss) and the validation set (val/cls loss) were monitored. If the loss curves failed to converge or showed instability, the training outcome was considered suboptimal, necessitating dataset adjustment, an increase in training epochs, or reconfiguration of model parameters.

After several training rounds, the optimal results are presented in Figure 3. As illustrated, the bounding box loss for the training set (train/box loss) showed a steady downward trend and reached convergence. Similarly, the bounding box loss for the validation set (val/box loss) also decreased consistently and approached stability. The classification losses for both the training and validation sets were also observed to stabilize, indicating that the model achieved synchronous convergence across both datasets and that the training process was effective.

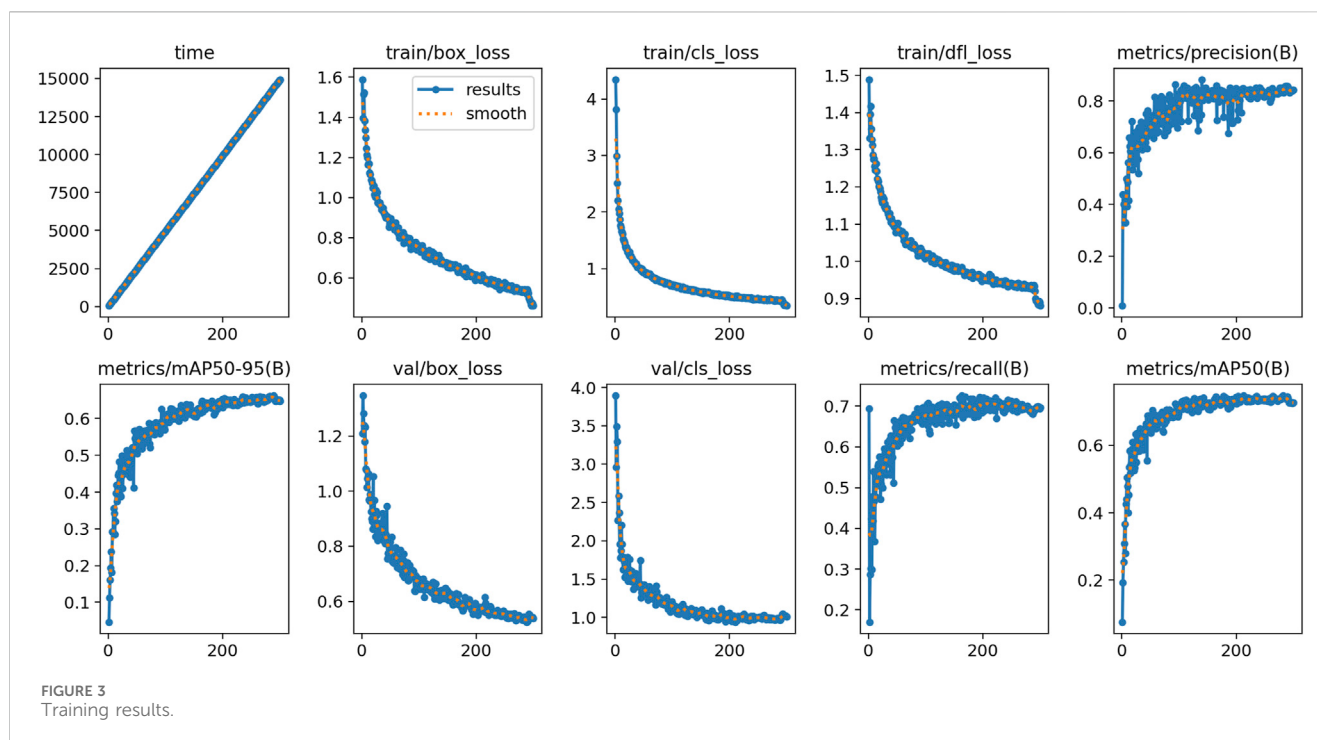
The precision metric (Metrics/Precision) reached approximately 90%, suggesting a high proportion of correct detections among all predicted positive instances. The recall metric (Metrics/Recall) exceeded 0.7, reflecting strong sensitivity to actual positive samples. The mean Average Precision (mAP), a key performance metric in object detection and information retrieval tasks, was also evaluated. Specifically, mAP50–95 refers to the mean of average precision values computed at intersection-over-union (IoU) thresholds ranging from 0.5 to 0.95 in increments of 0.05 (i.e., 0.50, 0.55, 0.60, ..., 0.95). As shown in Figure 3, after 200 training epochs, the mAP50–95 exceeded 70%. The mAP50, which corresponds to an IoU threshold of 0.5, also surpassed 70%, further indicating that the model achieved a relatively high level of prediction accuracy.

3.3 Result analysis

The analysis of results in this study is presented from two perspectives. First, the accuracy of the model is assessed across the training, validation, and test sets. Second, the performance of the model after transfer learning, referred to as Water-YOLO11n, is compared with the baseline YOLO11n model in terms of target classification and detection accuracy across problem categories.

3.3.1 Analysis of experimental results

The confusion matrix for the improved model is displayed in Figure 4, where the horizontal axis indicates the true class labels



and the vertical axis represents the predicted class labels. The matrix shows that most predictions align with the actual labels, demonstrating that the model exhibits strong predictive capability.

Figure 5 illustrates the precision values (y-axis) at varying confidence thresholds (x-axis). This precision–confidence curve allows an assessment of the model’s performance across different threshold settings. At a confidence threshold of 1.0, the model achieves an overall average precision of 0.99, indicating a high level of classification accuracy. However, the precision associated with floating fish is comparatively low. This may be attributed to the variability in appearance between individual fish and fish schools, which introduces complexity in feature learning. To improve performance in this category, further image samples representing diverse fish scenarios are needed.

The P–R curve is presented in Figure 6, showing the mAP@0.5 value for each target category, along with the overall mAP@0.5 for the Water-YOLO11n model following transfer learning. The results indicate that the transfer learning process improved the overall mAP@0.5–74.4%, reflecting a significant gain in detection accuracy. However, the mAP values for floating fish and leaves remain lower than for other categories. This performance gap is likely due to the complex visual characteristics of these targets and limited sample representation. Expanding the dataset with more images of these categories is expected to improve detection accuracy.

The F1-score performance across confidence thresholds is shown in Figure 7. According to the curve, the optimal F1-score of 0.72 is achieved at a threshold of 0.796. This indicates that the model maintains balanced precision and recall at that threshold, with an average F1-score of 0.72 across all categories. The results confirm that the Water-YOLO11n model delivers a high level of detection performance following transfer learning.

3.3.2 Instance-level verification

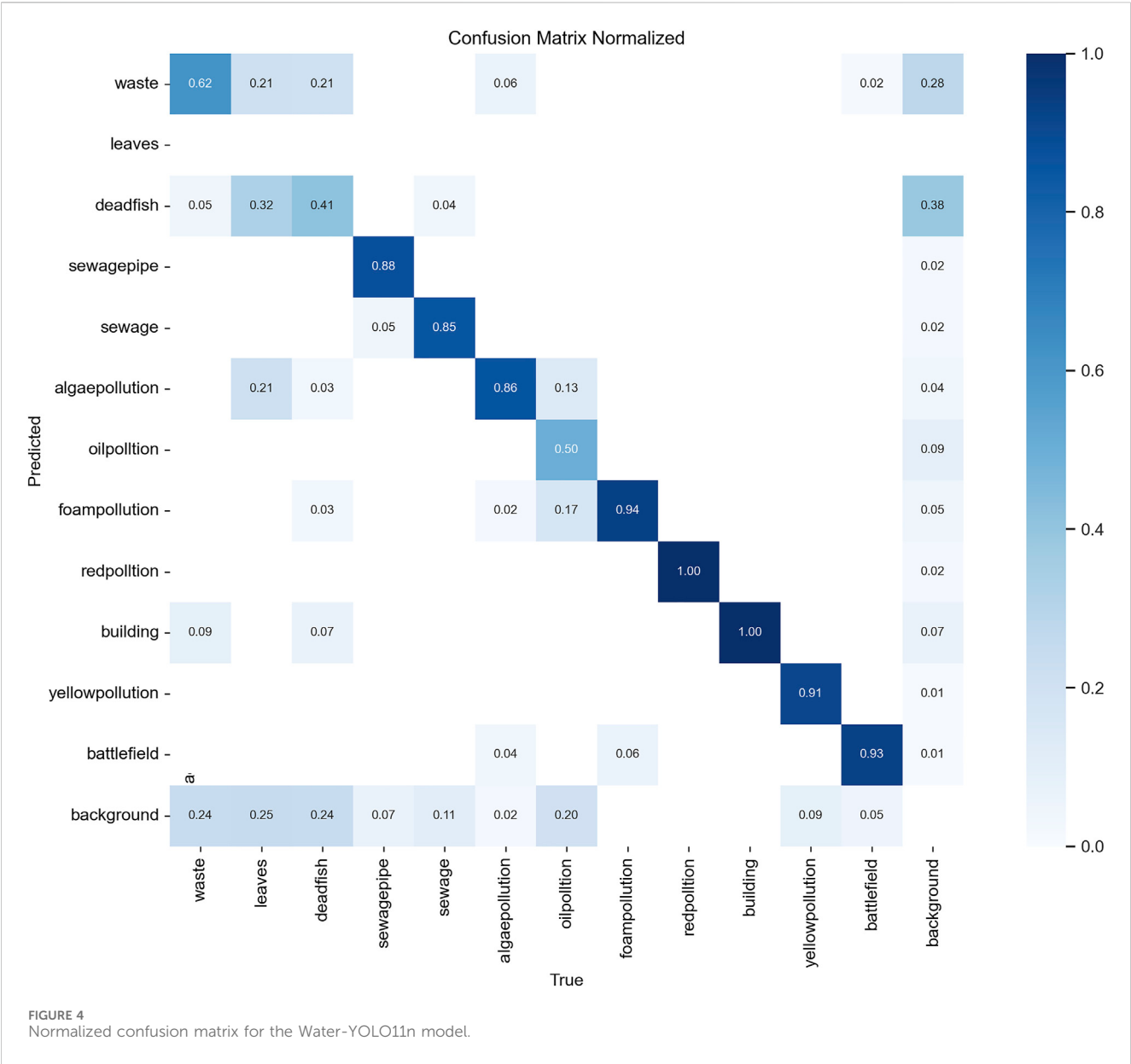
To further assess the effectiveness of the proposed method, the Water-YOLO11n model, developed through transfer learning based on YOLOv11, was compared with the baseline YOLO11n model before transfer learning. Figure 8 illustrates this comparison, where Figure 8A presents the detection results after transfer learning, and Figure 8B shows the results before transfer learning.

The post-transfer model demonstrated significantly improved recognition capabilities for categories such as garbage, floating objects on the water surface, sand yards, buildings, and drainage pipes. In contrast, the original YOLO11n model was unable to generate detection boxes for these categories, indicating that it failed to recognize such environmental hazards. The baseline model retained relatively high detection performance for general object types such as people and boats. However, further analysis revealed that while the original model performed well in identifying individuals on riverbanks, it exhibited lower recognition accuracy for individuals located in the water. To address this, future dataset expansions should include additional annotated samples of people working within water bodies, thereby improving the model’s ability to detect such instances.

The results confirm that transfer learning significantly enhances the model’s capability to identify diverse water-related environmental hazards. Furthermore, because annotation of specialized categories requires considerable domain knowledge, it may be practical to reduce or omit annotations for common object classes to minimize annotation workload. Ultimately, comprehensive target coverage may be achieved through model fusion strategies.

3.3.3 Model comparison analysis

To evaluate the performance of the Water-YOLO11n model in a broader context, its results were compared against several baseline



models, including YOLOv8s, YOLOv10n, and YOLOv11n, both prior to and after transfer learning. The comparison outcomes are summarized in [Table 2](#).

The Water-YOLO11n model achieved an F1-score of 0.72, with a precision of 0.848 and a recall of 0.69. Its mAP@0.5 reached 74.4%, while its mAP50–95 was recorded at 65.4%, indicating strong performance across a range of confidence thresholds. These metrics represent a substantial improvement over the original YOLO11n model before transfer learning, which achieved a mAP@0.5 of only 0.7% and a precision of 0.013. The enhanced model also demonstrated improved recall and F1-score, showing its ability to identify more relevant targets while reducing the incidence of false positives.

When compared with Water-YOLOv8s and Water-YOLOv10n, the Water-YOLO11n model yielded the highest mAP@0.5, while maintaining fewer parameters than Water-YOLOv8s. Although its inference speed (frames per second, FPS = 5.3) was slightly lower,

this trade-off is considered acceptable given the notable gain in detection accuracy. Among all the evaluated models, Water-YOLO11n exhibited the most balanced performance across precision, recall, and mAP, rendering it a suitable candidate for practical deployment in river and lake water environmental monitoring applications.

4 Discussion

Unlike the study by [Shen \(2022\)](#), which focuses on detecting the “four pests” issues from a broad perspective using satellite imagery, the present work emphasizes localized monitoring using drones or ground-based cameras. The model developed in this study is designed as a comprehensive tool for detecting a wide range of aquatic environmental hazards across multiple application scenarios. These include recognition of water discoloration to

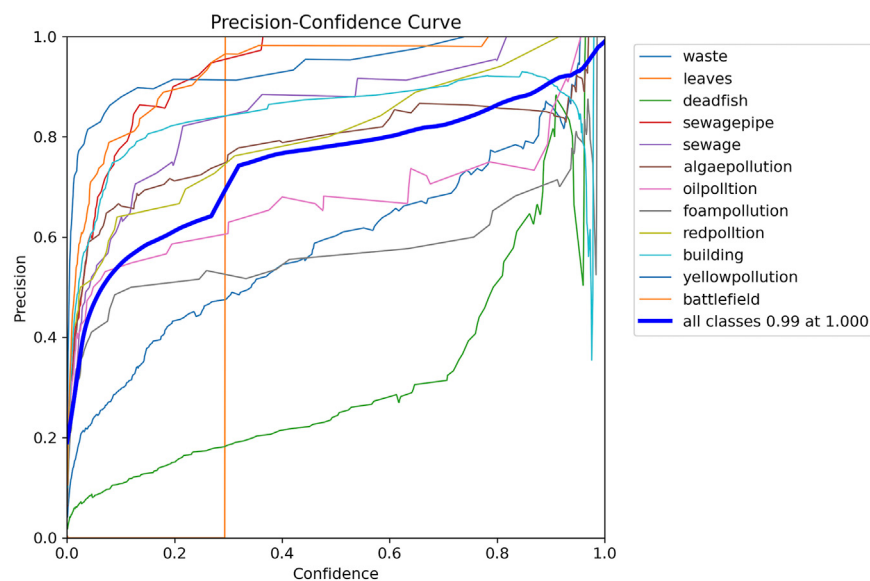


FIGURE 5
Prediction-confidence curve.

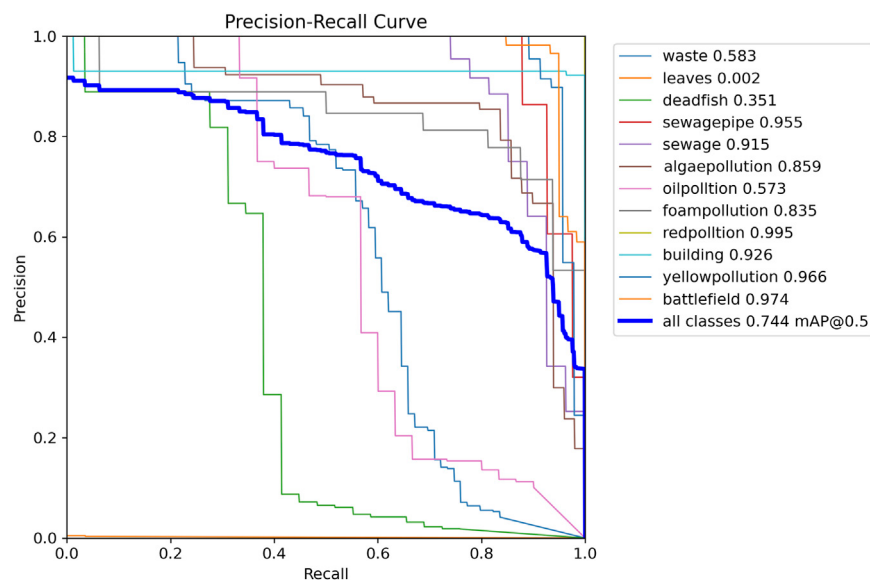


FIGURE 6
Prediction-recall curve.

support pollution type classification, monitoring of floating objects for assessing impacts on ecological indicators, and identification of sewage discharge outlets and structural intrusions as part of river management practices.

While the primary objective of this study was to develop a deep learning-based model for automatic detection of environmental hazards in river and lake ecosystems, the model also holds potential for use in broader environmental assessment tasks. In Europe, hydromorphological assessment of rivers and lakes is a crucial requirement for implementing the objectives of the Water

Framework Directive (European Commission, 2000). To address this need, member states of the European Union have developed methodologies for evaluating the hydromorphological status of lakes (e.g., Kutyla et al., 2021; Carriere et al., 2024) and rivers (e.g., Kamp et al., 2007; Szoszkiewicz et al., 2020). However, current approaches often depend on labor-intensive field surveys and visual assessments to evaluate shoreline structure, identify pollution sources, and track anthropogenic alterations. The model proposed in this study could augment these manual efforts by automatically detecting visible hazards such as

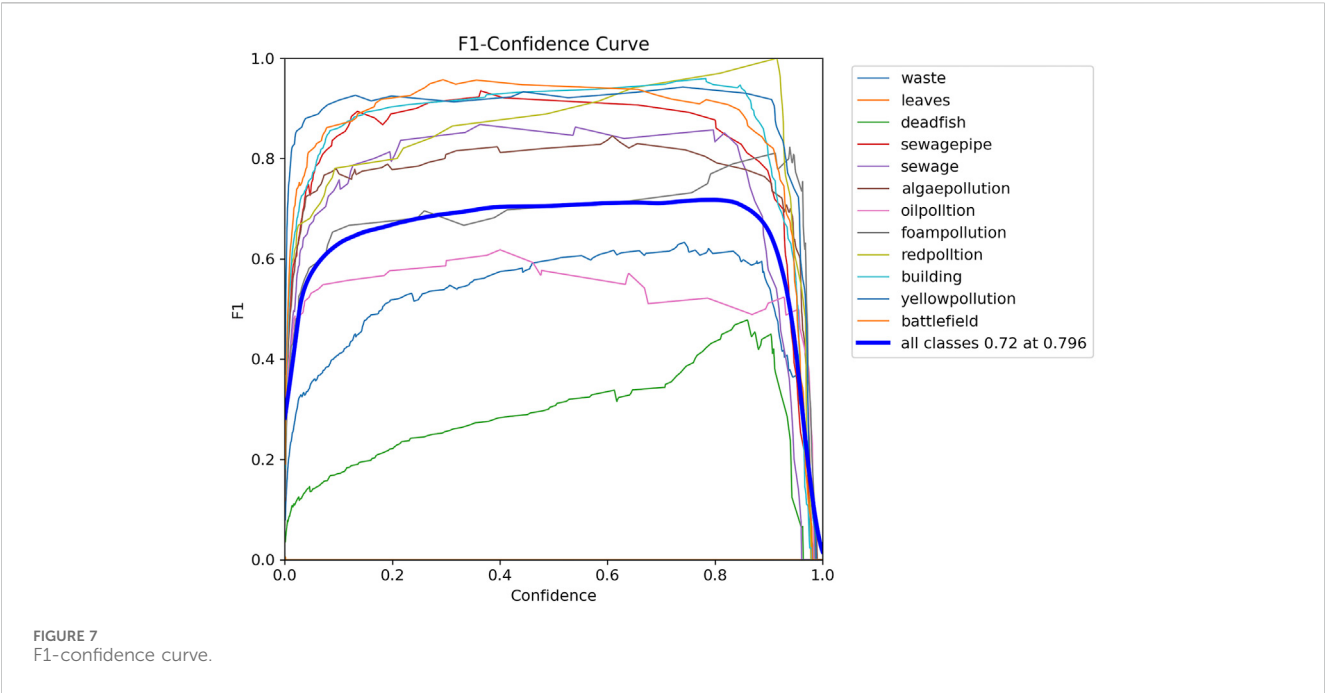


TABLE 2 Comparative performance of Water-YOLO11n and other YOLO series models.

Models	F1-score	P	R	mAP50	mAP50-95	Parameters	FPS
YOLO8s	—	0.00667	0.0740	0.00439	0.00241	11,156,544	11.5
YOLOV10n	—	0.01300	0.0532	0.00777	0.00453	2,762,608	6.8
YOLO11n	—	0.01300	0.0532	0.00777	0.00453	2,762,608	5.7
Water-YOLO8s	0.72	0.85100	0.6990	0.73900	0.68100	11,130,228	10.9
Water-YOLOV10n	0.69	0.79400	0.6840	0.72900	0.65700	2,699,876	6.2
Water-YOLO11n	0.72	0.84800	0.6900	0.74300	0.65400	2,584,492	5.3

floating debris, algal blooms, and unauthorized shoreline modifications using imagery captured by UAVs, mobile phones, or surveillance cameras. This integration would reduce field workload, enable more frequent assessments, and provide continuous visual documentation.

Beyond deep learning, several image-based analytical methods, as summarized in the review by [Manfreda et al. \(2024\)](#), may be integrated with the proposed model to enhance its utility in ecosystem monitoring. For instance, Large-scale Particle Image Velocimetry (LSPIV) can be applied to estimate surface flow

velocities for discharge estimation; Spectral Angle Mappers (SAMs) are effective for detecting and classifying macroplastics in both RGB and multispectral images; and structure-from-motion/multi-view stereo (SfM-MVS) photogrammetry techniques are used to reconstruct river morphology, enabling derivation of key hydromorphological parameters such as bankline positions, sandbar migration, and shallow water depth measurements.

The model also offers potential applications in environmental remediation and public engagement initiatives. For example, it can be used to monitor changes in surface-level pollution over time or integrated into citizen science platforms aimed at encouraging public participation in water quality surveillance. A mobile or web-based interface could be developed to allow volunteers, students, or local environmental groups to submit images or short video clips of water bodies captured via smartphones or drones. The model would then automatically analyze the media to detect features such as debris, algal accumulation, or shoreline alterations, providing immediate analytical feedback. This approach, drawing inspiration from initiatives like CrowdWater and Plastic Pirates, can expand the spatial and temporal reach of environmental monitoring while also fostering public environmental literacy and promoting a shared commitment to the protection of freshwater ecosystems.

Considering potential applications within China, the proposed model could be integrated into the national water quality assessment framework by aligning with existing infrastructure such as the National Surface Water Quality Monitoring Network and the River and Lake Chief System. For instance, it could enhance routine on-site inspections by assisting in the detection of illegal sewage outlets, enable real-time identification of macroplastics in urban rivers, such as those in the Yangtze and Pearl River basins, and function as an early warning tool for algal blooms in major reservoirs, including key drinking water sources like the Danjiangkou Reservoir and Taihu Lake. These applications are consistent with the objectives of the Yangtze River Protection Law and the National Ecological Civilization Strategy and may significantly reinforce compliance monitoring and promote broader community participation in the protection of aquatic ecosystems.

Despite the model's overall favorable performance, certain limitations remain, particularly in detecting specific categories such as floating fish and leaves. A contributing factor is likely the limited number of training samples for these categories, which constrains the model's ability to generalize. To address this, future research should focus on expanding the dataset with more varied and representative samples, incorporating feature enhancement techniques during preprocessing, and exploring the use of attention mechanisms to emphasize subtle distinguishing features. The visual complexity of these categories, including overlapping outlines, partial submergence, and low contrast with the background, also presents a challenge. These characteristics complicate the model's ability to distinguish between individual instances and grouped formations. Further development of datasets and recognition methods that support both individual-level and group-level detection will be necessary.

In addition to dataset-related factors, environmental conditions, such as lighting, weather variability, and seasonal changes, can impact detection accuracy. For example, glare on the water surface, low illumination during dawn or dusk, and seasonal shifts in leaf coloration may interfere with object visibility and lead to false positives or missed detections. To improve

performance under such conditions, future work should consider data augmentation strategies that simulate diverse environmental scenarios. Another promising direction is the use of multimodal fusion, which involves integrating visual data with auxiliary information such as time stamps, weather data, or Global Positioning System (GPS) coordinates. Such enhancements could improve the model's robustness and adaptability across a wider range of real-world application environments.

5 Conclusion

To address the demand for rapid and accurate identification of environmental hazards in river and lake systems, this study constructed the WATER-DET dataset, comprising over 1,500 annotated images spanning 12 categories of water-related hazards. This dataset serves as a specialized benchmark resource for research in aquatic hazard detection. Based on the YOLOv11 architecture and leveraging transfer learning through the YOLO11n model, a detection model named Water-YOLO11n was developed. The model achieves a high level of accuracy in identifying true hazards when the model is highly confident— at a confidence score of 1.0, the overall precision of all classes reaches 0.99. At an intersection-over-union threshold of 0.5, the average recall across all categories reached 74.4%, indicating strong detection sensitivity. The model's F1-score of 0.72 reflects a well-balanced trade-off between precision and recall, demonstrating its robust overall performance.

The model showed high accuracy in comparative evaluations, confirming its potential as a reliable tool for the detection of water environmental hazards. As such, it offers valuable technical support for water resource management and protection efforts.

In the future, the model can be exported in widely supported formats and deployed across various operational contexts. It is suitable for implementation on servers, edge computing platforms, or mobile devices, enabling both real-time and offline monitoring according to specific application requirements. Ongoing collection of user feedback and field imagery will support further expansion of the dataset, facilitating continuous refinement of the model. These improvements are expected to enhance the model's detection capability by broadening its classification scope and improving identification accuracy and reliability in practical environmental monitoring scenarios.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

XS: Writing – review and editing, Conceptualization, Investigation. GZ: Writing – original draft, Methodology, Investigation, Writing – review and editing. XW: Data curation, Visualization, Writing – review and editing, Investigation. JX: Conceptualization, Supervision, Writing – review and editing, Investigation.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

I would like to express my gratitude to all those who have helped me during the writing of this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Anand, K., Kerry, W., Wang, Z., and Cheryl, M. (2019). Deep learning–Method overview and review of use for fruit detection and yield estimation. *Comput. and Electron. Agric.* 162, 219–234. doi:10.1016/j.compag.2019.04.017
- Aziz, L., Salam, S. B. H. Md, Sheikh, U. U., and Ayub, S. (2020). Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: a comprehensive review. *IEEE Access* 8, 170461–170495. doi:10.1109/access.2020.3021508
- Cai, C., Gao, S., Zhou, J., and Huang, Z. (2020). Freeway anti-collision warning algorithm based on vehicle-road visual collaboration. *J. Image Graph.* 25 (8), 1649–1657. doi:10.11834/jig.190633
- Carriere, A., Reynaud, N., Gay, A., Baudoin, J. M., and Argillier, C. (2024). LHYMO: a new Water Framework Directive-compliant multimetric index to assess lake hydromorphology and its application to French lakes. *Aquatic Conservation Mar. Freshw. Ecosyst.* 34 (1), e4029. doi:10.1002/aqc.4029
- Chen, J., He, Z., Zhu, D., Hui, B., Yi, R., Li, M., et al. (2022). Mu-Net: Multi-Path upsampling convolution network for medical image segmentation. *Comput. Model. Eng. and Sci.* 131 (4), 73–95. doi:10.32604/cmes.2022.018565
- Ding, Y., Yang, A., and Kang, W. (2024). Underwater image recognition based on improved EfficientNet. *Ship Sci. Technol.* 46 (15), 95–100.
- Du, Y., Zhang, R., Shi, P., Zhao, L., Zhang, B., and Liu, Y. (2024). ST-LaneNet: lane line detection method based on swin transformer and LaneNet. *Chin. J. Mech. Eng.* 1, 14–158. doi:10.1186/s10033-024-00992-z
- European Commission (2000). *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy*. Official Journal of the European Communities L 327/1 of 22.12.2000.
- Gao, Y. (2023). *Research on the identification of water pollution problems and process-oriented countermeasures in the Shaanxi section of the weihe river*. Xi'an: Xi'an University of Technology.
- Ge, C., Chen, J., and Yan, J. (2024). Design and implementation of dam defect recognition system. *Softw. Guide* 22 (5), 84–90.
- Girshick, R. (2015). “Fast R-CNN,” in *Proceedings of the IEEE conference on computer vision and pattern recognition* (Boston, USA: IEEE), 1440–1448. doi:10.1109/CVPR.2015.7298684
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition* (Columbus, Ohio: IEEE), 580–587. doi:10.1109/CVPR.2014.81
- Guzel, M., Turan, B., Kadioglu, I., Basturk, A., Sin, B., and Sadehpour, A. (2024). Deep learning for image-based detection of weeds from emergence to maturity in wheat fields. *Smart Agric. Technol.* 9, 100552. doi:10.1016/j.atech.2024.100552
- Heathcote, A. J. (2013). *Anthropogenic eutrophication and ecosystem functioning in freshwater lakes*. Iowa State University Capstones, 158. PhD dissertation.
- Huang, G. (2022). A method for detecting defects in transmission lines of water conservancy projects based on UAV image recognition technology. *Water Conservancy Sci. Technol. Econ.* 28 (8), 137–141.
- Kamp, U., Binder, W., and Hölzl, K. (2007). River habitat monitoring and assessment in Germany. *Environ. Monit. Assess.* 127, 209–226. doi:10.1007/s10661-006-9274-x
- Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. *Annu. Conf. Neural Inf. Process. Syst.* 25, 84–90. doi:10.1145/3065386
- Kutyla, S., Soszka, H., and Kolada, A. K. (2021). Hydromorphological assessment of Polish lakes: elaborating the Lake Habitat Survey for Polish Lakes (LHS_PL) method and determining ecologically based boundary values for lake classification. *Ecologyhydrology* 14 (6), e2320. doi:10.1002/eco.2320
- Kutyla, S., Kolada, A., and Ławniczak-Malińska, A. (2024). How far from the shoreline? The effect of catchment land use on the ecological status of flow-through lakes. *Ecologyhydrology* 24 (2), 299–310. doi:10.1016/j.ecohyd.2023.08.010
- Lai, Y., Zhang, J., Li, W., and Song, Y. (2024). Water quality monitoring of large reservoirs in China based on water color change from 1999 to 2021. *J. Hydrology* 633, 130988. doi:10.1016/j.jhydrol.2024.130988
- Lee, E., Kim, J.-S., Park, D. K., and Whangbo, T. (2024). YOLO-MR: meta-learning-based lesion detection algorithm for resolving data imbalance. *IEEE Access* 12, 49762–49771. doi:10.1109/access.2024.3384088
- Li, J. (2022). Traffic signal detection based on YOLOv3 in deep learning PyTorch framework. *Automob. Electr. Appliances* 6, 4–7. doi:10.14031/j.cnki.njwx.2022.06.001
- Li, Y., Duan, Y., Duan, L., Xiang, W., and Wu, Q. (2024). “YOLO-TL: a tiny object segmentation framework for? Low quality medical images,” in *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, 143–159. doi:10.1007/978-3-031-66958-3_11
- Lin, K. (2024). Implementation method of real-time automatic patrol inspection for monitoring status of intelligent Hydropower stations. *Eng. Technol. Res.* 9 (14), 42–44.
- Liu, J. (2019). Application and prospect of artificial intelligence (AI) in the water industry abroad. *Water Purif. Technol.* 9, 6–11.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., et al. (2016). “SSD: Single shot multibox detector,” in *Proceedings of the European conference on computer vision* (Springer), 21–37. doi:10.1007/978-3-319-46448-0_2
- Liu, Y., Zhao, Q., Wang, X., Sheng, Y., Tian, W., and Ren, Y. (2024). A tree species classification model based on improved YOLOv7 for shelterbelts. *Front. Plant Sci.* 14, 1265025. doi:10.3389/fpls.2023.1265025
- Lu, H., and Gao, H. (2023). “Research on image recognition and pattern recognition technologies in the construction quality monitoring of modern water conservancy projects,” in *Zhengzhou: 2023 China Water Conservancy Academic Conference*, Zhengzhou, China, 13–14 August 2023.
- Manfreda, S., Miglino, D., Saddi, K. C., Jomaa, S., Eltner, A., Perks, M., et al. (2024). Advancing river monitoring using image-based techniques: challenges and opportunities. *Hydrological Sci. J.* 69 (6), 657–677. doi:10.1080/02626667.2024.2333846
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). “You only look once: unified, real-time object detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 26 June – 1 July 2016, 779–788. doi:10.1109/CVPR.2016.91

- Ren, S., He, K., Girshick, R., and Sun, J. (2015). "Faster R-CNN: towards real-time object detection with region proposal networks," in *International Conference on Advances in Neural Information Processing Systems*, Montreal, Canada, 7-12 December 2015, 91–99.
- Ren, Z., Wang, Q., and Zhu, B. (2022). "Application of image recognition technology in intelligent patrol inspection of a large hydropower station," in *Proceedings of the 2022 annual meeting of automation committee of China hydropower engineering society and national Academic exchange conference on intelligent application of hydropower plants, jiangsu, China*.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-Net: convolutional networks for biomedical image segmentation," in *Proceedings of the 18th international conference on medical image computing and computer - assisted intervention (MICCAI 2015)* (Munich, Germany: Springer), 234–241. doi:10.1007/978-3-319-24574-4_28
- Shakuli, S. (2021). An overview on water pollution. *Int. Multidiscip. Res. J.* 11, 1046–1051. doi:10.5958/2249-7137.2021.02559.3
- Shen, Y. (2022). *Specific object detection in satellite remote sensing images of the Yellow river basin based on deep learning*. North China University of Water Resources and Electric Power. Master's thesis.
- Shi, T. (2017). Analysis of Urban Water Pollution Problems and Their Prevention and Control Measures. *China Venture Capital* 2, 298–326.
- Szozskiewicz, K., Jusik, S., Gebler, D., Achtenberg, K., Adynkiewicz-Piragas, M., Radecki-Pawlik, A., et al. (2020). Hydromorphological index for rivers: a new method for hydromorphological assessment and classification for flowing waters in Poland. *J. Ecol. Eng.* 21 (8), 261–271. doi:10.12911/22998993/126879
- Wei, X., Chen, Y., Zhang, L., Chang, M., and Gao, H. (2022). Research progress on basin water pollution monitoring and source tracing technologies. *Environ. Monit. China* 38 (5), 27–31. doi:10.19316/j.issn.1002-6002.2022.05.04
- Xu, J., Ren, H., Cai, S., and Zhang, X. (2023). An improved faster R-CNN algorithm for assisted detection of lung nodules. *Comput. Biol. Med.* 153 (Suppl. C), 106470. doi:10.1016/j.combiomed.2022.106470
- Xu, Y., Wen, M., He, W., Wang, H., and Xue, Y. (2024). An improved multi-scale and knowledge distillation method for efficient pedestrian detection in dense scenes. *J. Real-Time Image Process.* 21, 126–8200. doi:10.1007/s11554-024-01507-8
- Yan, D., Lei, M., and Shi, Y. (2025). A hybrid approach to advanced NER techniques for AI-driven water and agricultural resource management. *Front. Environ. Sci.* 13, 1558317. doi:10.3389/fenvs.2025.1558317
- Zeng, L. (2024). Estimation of water quality in Korattur Lake, Chennai, India, using Bayesian optimization and machine learning. *Front. Environ. Sci.* 12, 1434703. doi:10.3389/fenvs.2024.1434703
- Zhou, H., Jin, S., Zhou, L., Guo, Z., Sun, M., and Shi, M. (2023). Classification and recognition of camellia oleifera fruit in the field based on transfer learning and YOLOv8n. *Trans. Chin. Soc. Agric. Eng.* 39 (20), 159–166. doi:10.11975/j.issn.1002-6819.202307244