scientific reports

Check for updates

OPEN

Improving YOLOv11 for marine water quality monitoring and pollution source identification

Fang Wang

Marine pollution has become an increasingly severe environmental issue, with oil spills, marine debris, and turbid water significantly impacting ecosystems and human health. The You Only Look Once (YOLO) series of target detection has been widely applied in Marine pollution monitoring. However, in complex underwater environments, challenges such as irregular pollutant shapes, varying scales, and background interference limit detection accuracy and robustness. To address these issues, this study proposes an improved YOLOv11 model that integrates Deformable Convolutional Networks version 4 (DCNv4) to enhance adaptability to deformable pollutants, improving detection precision. The Marine Fusion Loss (MFL) mechanism optimizes detection weight allocation among different pollutant categories, reducing false positives. Additionally, Multi-scale Feature Fusion (MFF) combines Convolutional Neural Networks (CNN) and Transformer-based feature extraction to enhance robustness in complex environments. Furthermore, instance segmentation is incorporated to refine boundary detection of pollutants. Experiments show that the improved YOLOv11 model outperforms the most advanced methods such as YOLOv8 and YOLOv10, with an average accuracy of 90.2% when 50% intersection exceeds union (mAP50) and an inference speed of 3.5ms, ensuring high precision and high efficiency. The results validate the effectiveness of the proposed method in enhancing marine pollution detection, providing a high-performance solution for intelligent environmental monitoring.

Keywords Marine pollution detection, YOLOv11, Deformable convolution (DCNv4), Multi-scale Feature Fusion (MFF), Marine Fusion Loss (MFL)

The ocean is the largest ecosystem on Earth, playing a crucial role in maintaining global climate balance, providing biological resources, and supporting sustainable human development[1]. However, with rapid industrialization, marine pollution has become increasingly severe, particularly due to oil spills, marine debris, eutrophic sedimentation, and turbid waters, posing significant threats to marine ecosystems and human health. Oil spills can devastate marine habitats, causing mass mortality of fish, birds, and marine mammals, while also negatively impacting coastal economies such as fisheries and tourism[2]. Marine debris, especially plastic waste, is a global pollution crisis due to its persistence in the environment[3,4]. Additionally, abnormal proliferation of marine plankton (e.g., sea snot) and the formation of turbid waters alter the ocean's optical properties, affecting the accuracy of marine ecological monitoring and remote sensing technologies[5]. Consequently, the development of efficient and accurate intelligent detection methods for marine water quality monitoring and pollution source identification has become a key research focus in environmental protection.

Recent advancements in deep learning-based object detection techniques, such as the YOLO (You Only Look Once) series[6,7], Mask R-CNN[8], and SSD (Single Shot MultiBox Detector)[9], have been widely applied to marine pollution detection[10]. However, challenges persist in underwater environments, including low illumination, suspended particle interference, and object deformation due to water currents, which limit the detection accuracy and generalization ability of conventional models. Optimizing existing object detection frameworks to enhance their adaptability for marine water quality monitoring and pollution source identification remains a critical research challenge[11,12].

YOLOv11[13], the latest iteration in the YOLO series, introduces improvements in the Cross Stage Partial (CSP) structure and a novel loss function to enhance small-object detection accuracy[14]. However, it still suffers from false positives and reduced robustness in complex underwater scenes. Mask R-CNN, an extension of Faster R-CNN[15] with instance segmentation capabilities, improves marine debris detection by identifying object contours but incurs high computational costs, making real-time applications challenging. EfficientDet[16], utilizing a Bidirectional Feature Pyramid Network (BiFPN)[17], excels in multi-scale object detection but struggles

School of Computer and Software Engineering, ZhengZhou Sias University, Zhengzhou 451100, China. email: wangfang2024113@126.com

with variable lighting conditions in underwater imagery. Transformer-based DETR (Detection Transformer)[18] leverages self-attention mechanisms to model long-range dependencies, improving detection robustness for marine debris. However, its high computational demands make it impractical for resource-constrained environments. A lightweight variant, YOLOv8-GST-YOLO[19], optimizes the network structure to achieve high detection speed and improved accuracy for underwater debris detection, making it suitable for embedded systems and Autonomous Underwater Vehicles (AUVs).

To overcome these limitations, this study proposes an improved YOLOv11 framework for marine water quality monitoring and pollution source identification. The model integrates Deformable Convolutional Networks v4 (DCNv4)[20] to enhance adaptability to deformable objects and improve detection accuracy. Additionally, a Marine Fusion Loss (MFL)[21] function dynamically adjust loss weights for different pollutant categories during training, enhancing small-object detection (e.g., plastic debris, oil films). A Multi-scale Feature Fusion (MFF)[22] mechanism is also introduced to effectively utilize both low- and high-level features, enhancing detection robustness in complex underwater environments.

The main contributions of this paper are as follows:

- The improved YOLOv11 model integrates DCNv4 to enhance adaptability to deformable marine pollutants, improving detection accuracy and robustness;
- The proposed MFL dynamically adjust loss weights for different pollutant categories, enhancing small-object detection such as plastic debris and oil films;
- MFF improves detection robustness in complex underwater environments, enabling better recognition of various marine pollution sources.

The remainder of this paper is structured as follows: Section Related work reviews related work, including current marine pollution detection techniques and the application of deep learning in this domain. Section Experiment details the improved YOLOv11 model, including network architecture, loss function, and optimization strategies. Section Results presents experimental results, comparing detection accuracy, computational efficiency, and robustness across various methods. Finally, Section Conclusion summarizes key findings and outlines future research directions.

Related work

Recent advancements in Transformer architectures have provided new solutions for marine pollution monitoring. Compared to traditional convolutional neural networks (CNNs), Transformer-based models can capture long-range dependencies and maintain a global receptive field, making them particularly effective for pollutant detection and tracking in complex environments. Marine pollutants often exhibit irregular shapes, color similarities, and complex backgrounds, which pose challenges for conventional object detection methods[23,24]. Additionally, surface wave fluctuations and lighting variations further compromise detection accuracy. Transformer models, leveraging self-attention mechanisms, effectively extract distinctive pollutant features, enhancing both detection accuracy and robustness.

Swin Transformer[25] is a hierarchical vision Transformer that employs sliding window attention to reduce computational complexity while maintaining high feature extraction efficiency. It has been utilized in high-resolution remote sensing imagery for detecting marine pollutants such as floating debris, oil spills, and algal blooms. While its hierarchical structure enhances detection precision in large-scale remote sensing analysis, its high computational cost limits its practical application in real-time scenarios. DETR[18] (Detection Transformer), the first Transformer-based object detection model, eliminates the need for region proposal networks (RPNs) by directly predicting object classes and locations using self-attention. In marine pollution monitoring, DETR demonstrates robust detection capabilities for floating oil, marine debris, and sea snot, effectively mitigating sea surface reflections and wave disturbances. However, its performance on small object detection is limited, as it requires extensive training time and significant computational resources.

PVT (Pyramid Vision Transformer)[26] combines the advantages of CNNs and Transformers using a pyramid structure for multi-scale feature extraction, making it suitable for hierarchical pollutant detection. It effectively distinguishes between large-scale oil spills, small floating oil patches, and plastic debris, demonstrating adaptability across different resolutions. Despite its superior multi-scale feature extraction, PVT also incurs high computational demands, which restricts its deployment in real-time or resource-constrained applications. DINO[27], an improved DETR variant, integrates denoising training and Hungarian matching to enhance detection accuracy and convergence speed. It excels in dense pollution areas, such as coastal waste clusters and abandoned fishing nets, but remains unsuitable for embedded or low-power devices due to its heavy computational overhead. MaxViT[28], a hybrid CNN-Transformer model, combines local window attention and global attention mechanisms, improving pollutant detection across different scales[29].

In addition to Transformer models, DCNv4[20] enhances deformable object adaptability, making it particularly beneficial for marine pollution detection[30]. Floating pollutants often exhibit varying shapes and orientations, which traditional convolutions struggle to capture. DCNv4 dynamically adjust sampling positions to better align with pollutant contours, significantly improving detection performance in complex marine environments. Table 1 summarizes several existing methods for marine pollution detection, including their advantages, disadvantages, and the research gaps they address. This table highlights the strengths and limitations of current approaches, and clarifies how our proposed model, with its integration of DCNv4 and MFF, improves upon these methods in terms of accuracy and real-time applicability.

Method	Advantages	Disadvantages	Research gaps addressed by this study
YOLOv4[6]	Fast real-time detection, high accuracy for large objects	Struggles with small object detection and cluttered backgrounds	Improves small-object detection with DCNv4 and MFF
Mask R-CNN[8]	Accurate instance segmentation, works well on large objects	High computational cost, not suitable for real-time applications	Enhances real-time performance by optimizing model architecture
EfficientDet[16]	Efficient multi-scale feature extraction, lightweight	Struggles with underwater lighting variations	Improves robustness to environmental variations by incorporating DCNv4
DETR[18]	Robust detection in complex backgrounds, self-attention mechanism	High computational cost, long training time	Optimizes for real-time performance and small-object detection
Transformer-based models[25]	Powerful feature extraction, effective for large datasets	Requires high computational resources	Reduces computational load with lightweight optimizations (e.g., MFF)

Table 1. Summary of Existing Methods for Marine Pollution Detection.

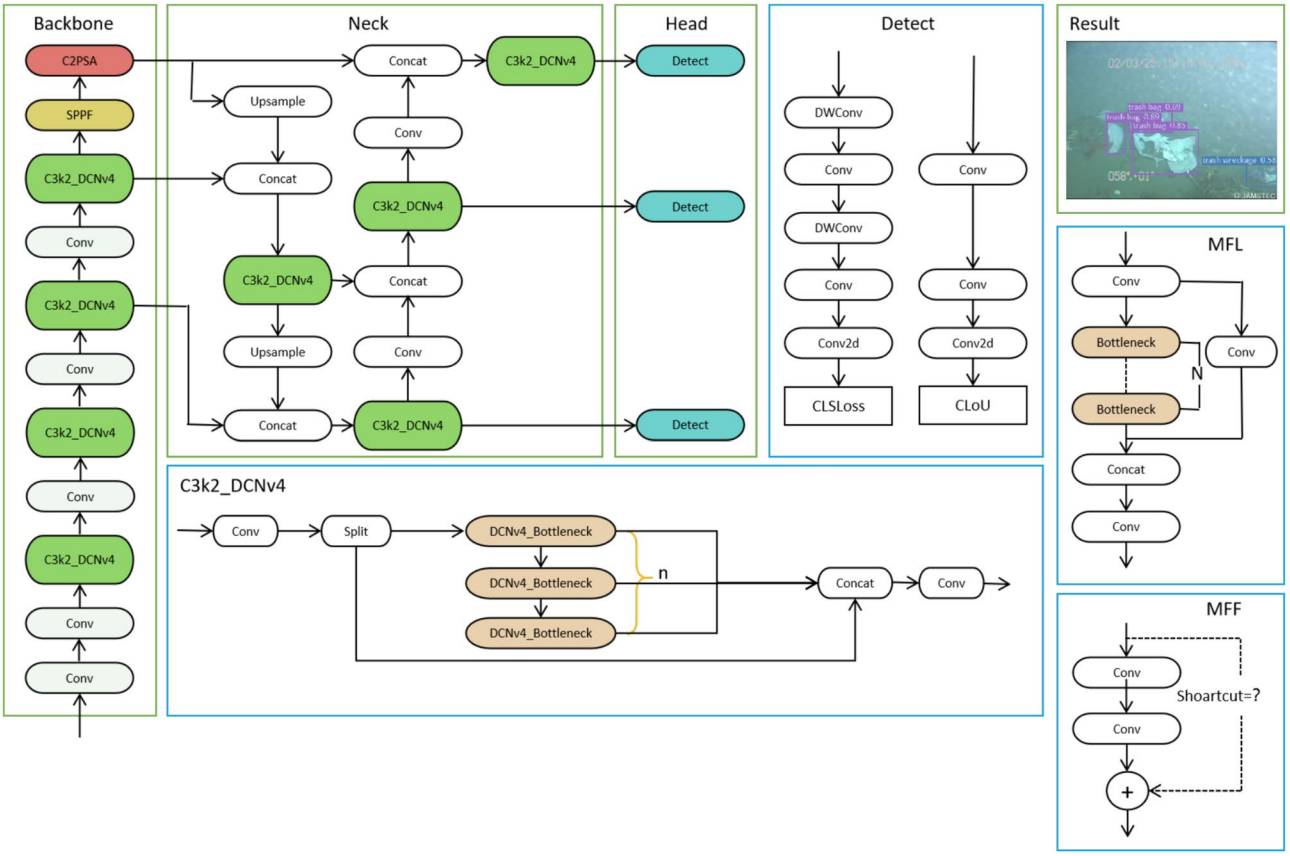

Fig. 1. Improved YOLOv11 network architecture, where the green box represents our improved DCNv4 architecture.

Method

This study proposes an enhanced YOLOv11n-based model for marine pollution monitoring and pollution source identification, focusing on pollutants such as oil spills, marine debris, turbid water, foam, and sea snot, as illustrated in Fig. 1. Marine pollutants often exhibit irregular shapes, varying scales, and low contrast with their backgrounds, which makes conventional object detection methods prone to interference and limits their detection accuracy in complex environments. To address these challenges, we introduce several novel enhancements to the YOLOv11 model. These include Deformable Convolutional Networks version 4 (DCNv4), which improves adaptability to deformable pollutants like diffused oil slicks and floating debris by using deformable convolution operations that refine small-object and boundary detection accuracy. We also incorporate MFL, which uses an adaptive loss-weighting strategy to assign different loss weights to various pollutant categories during training, increasing detection precision and reducing background false positives. Additionally, MFF combines the feature extraction capabilities of CNNs and Transformer architectures, enabling efficient multi-scale pollutant detection, particularly for distant and complex pollution sources. Together, these innovations enhance the model's ability to detect pollutants in a range of underwater environments, improving both detection accuracy and generalization ability.

Enhancing adaptability to deformable pollutants

DCNv4 is an advanced deformable convolution module designed to improve the detection of deformable objects in complex environments. Unlike standard convolution (which uses fixed, regularly spaced sampling points), DCNv4 introduces dynamic sampling points that adapt based on the feature map, significantly improving detection accuracy for irregularly shaped pollutants like oil spills, floating debris, and turbid water. This flexibility allows the model to adjust to shape variations, offering substantial improvements over previous versions of deformable convolutions (e.g., DCN, DCNv2, DCNv3)[31]. Figure 2 illustrates the difference between DCNv3 and DCNv4.

In standard convolution, the sampling points are fixed on a regular grid, which limits the detection performance for spatially deformed objects. For example, in detecting irregular pollutants such as diffused oil slicks, fixed sampling points fail to capture the true contours of these pollutants. In contrast, DCNv4 uses a dynamic offset mechanism, allowing the convolution kernel's sampling points to move according to the shape of the object. This ensures that the kernel can better cover the pollutant regions, significantly improving detection accuracy.

Additionally, DCNv4 introduces a modulation mechanism, which adjust the influence of each sampling point by learning an amplitude modulation weight for each. This modulation enhances the feature extraction capability of the network, particularly in challenging underwater environments where pollutants may be faint or distorted due to water currents or lighting conditions. The dynamic sampling expression of DCNv4 is shown as Eq. 1:

$$y(p) = \sum\_k = 1^K w\_k \cdot x(p + \Delta p\_k) \cdot \Delta m\_k \tag{1}$$

In this equation, $y(p)$ represents the output feature map at position p, while K is the kernel size (e.g., 3×3). The term w_k refers to the standard convolution kernel weights, and $x(p + \Delta p_k)$ is the input feature at the dynamically offset position $p + \Delta p_k$. The offset value Δp_k is calculated by DCNv4, enabling adaptive sampling that is crucial for detecting deformable objects. Additionally, Δm_k represents the modulation scalar, which controls the weight of each sampled point.

In standard convolution, the offset $\Delta p_k = 0$, meaning sampling points are fixed on a regular grid. However, in DCNv4, the offsets Δp_k are dynamic, allowing the model to learn the most relevant sampling points for each pollutant. This flexibility makes DCNv4 especially effective for detecting complex-shaped marine pollutants.

The introduction of shared aggregation weights further improves computational efficiency by allowing multiple channels to process the same features concurrently. This results in faster processing and reduced memory overhead, making DCNv4 an excellent fit for real-time marine pollution monitoring tasks.

By integrating DCNv4 into YOLOv11, the model significantly enhances detection accuracy for deformable pollutants such as diffused oil spills and fragmented marine debris, which are difficult to detect using traditional convolution methods. The modulation mechanism also helps reduce the impact of irrelevant background noise, such as sea waves, on the detection accuracy.

Marine fusion loss for adaptive pollutant feature learning

In marine pollution detection, pollutants such as oil spills, marine debris, foam, and sea snot exhibit characteristics like blurry boundaries, complex shapes, and varying sizes, leading to frequent false positives and false negatives in detection. Moreover, class imbalance (e.g., large-scale oil spills vs. small floating debris) and background interference (e.g., sea waves and shadows) further increase detection complexity. To address these challenges, we introduce MFL, an adaptive loss-weight optimization strategy that improves detection accuracy across various marine pollutants.

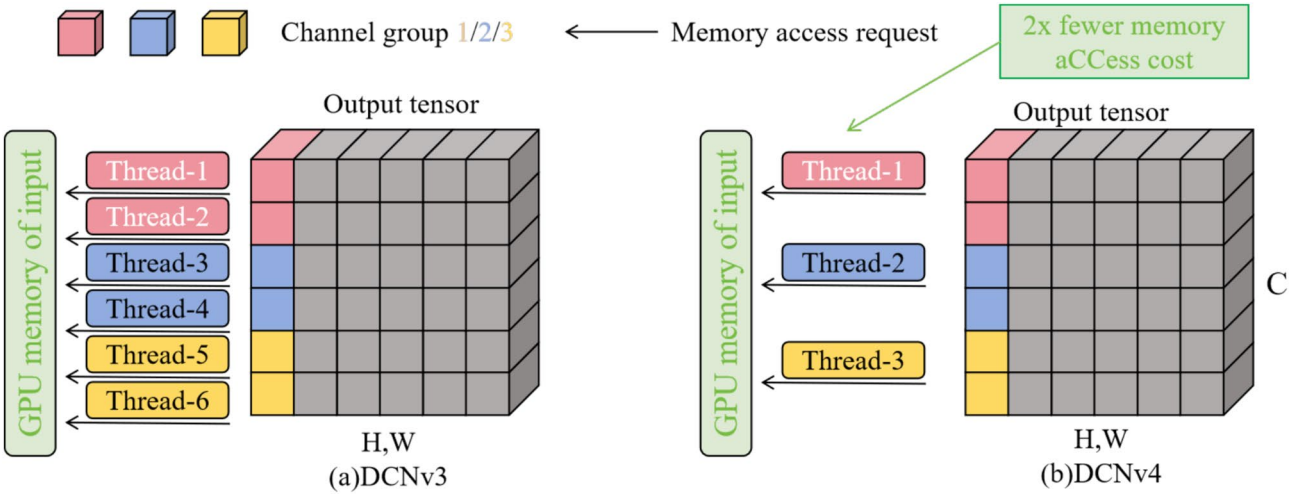

Fig. 2. Structure comparison between DCNv3 and DCNv4.

MFL consists of four core loss components: Localization Loss (L_{loc}), based on Generalized IoU (GIoU), optimizes bounding box prediction accuracy and improves target region coverage. Classification Loss (L_{cls}) utilizes Focal Loss to address class imbalance, enhancing recognition of low-frequency pollutants such as microplastics. Confidence Loss (L_{conf}), implemented using Binary Cross-Entropy (BCE), ensures the reliability of object presence predictions. Attention Loss (L_{att}) refines the attention mechanism to focus on pollutant regions while mitigating background noise interference.

By combining these loss components with adaptive weight adjustments, MFL balances optimization across different pollutant scales and categories, allowing the model to generalize effectively in complex marine environments while improving detection accuracy and robustness.

To enhance performance in marine pollution detection, we integrate MFL into the loss computation module of YOLOv11. By dynamically adjusting loss weights for different pollutant categories, MFL optimizes class-specific detection performance, reducing false positives and false negatives.

GIoU-based L_{loc} enhances bounding box accuracy, improving the detection of diffused oil spills and floating debris. Focal Loss-based L_{cls} strengthens the model's ability to classify hard-to-detect pollutants, such as microplastics and sea foam. Attention Loss L_{att} optimizes the model's attention weights, enabling it to focus on pollution regions and mitigate interference from waves and lighting variations.

MFL integrates multiple weighted loss components for optimal bounding box accuracy. Localized loss is expressed by generalized IoU (GIoU), as shown in Eq. 2:

$$L_{loc} = 1 - GIoU(B_p, B_t) \tag{2}$$

where B_p represents the predicted bounding box, B_t is the ground truth bounding box, and $GIoU$ is the generalized IoU function. To handle classification imbalance, Focal Loss is used for classification loss, such as Eq. 3:

$$L_{cls} = -\alpha(1 - p_t)^{\gamma} \log(p_t) \tag{3}$$

where p_t is the predicted probability for the ground truth class, α is the class balance factor, and γ is the focusing parameter that reduces the impact of easy samples while increasing the weight of hard samples. Optimize the confidence loss by using binary cross-entropy (BCE), such as Eq. 4:

$$L_{conf} = -(y \log(p) + (1 - y) \log(1 - p)) \tag{4}$$

where y indicates the presence of an object (1 for presence, 0 for absence), and p is the predicted confidence score. To further refine the contaminant localization, an attention loss as in Eq. 5 is introduced:

$$L_{att} = \sum_{i=1}^{N} (A_i - A_i^*)^2 \tag{5}$$

where A_i represents the predicted attention map, A_i^* is the ground truth attention map, and N is the total number of attention maps. Final loss Eq. 6:

$$L = \lambda_{loc} L_{loc} + \lambda_{cls} L_{cls} + \lambda_{conf} L_{conf} + \lambda_{att} L_{att} \tag{6}$$

where the weight parameters are defined as follows: λ_{loc}(Localization Loss Weight) ranges from 0.5 to 2.0, λ_{cls}(Classification Loss Weight) ranges from 1.0 to 2.5, λ_{conf}(Confidence Loss Weight) ranges from 0.7 to 1.5, and λ_{att}(Attention Loss Weight) ranges from 0.5 to 1.5.

These hyperparameters can be optimized through grid search to achieve optimal detection performance. MFL enhances marine pollution detection by assigning adaptive loss weights to different pollutant classes, improving detection accuracy and reducing false positives. It employs GIoU-based bounding box optimization to refine pollutant region prediction while mitigating class imbalance effects to increase detection recall. Additionally, it optimizes the attention distribution, ensuring the model focuses on pollutant regions rather than irrelevant background areas.

Multi-scale Feature Fusion

To enhance detection accuracy, we introduce MFF, which integrates the advantages of CNN and Transformer architectures to optimize multi-scale feature extraction in YOLOv11 for marine pollutant detection. MFF enables the model to detect both large-scale pollutants (e.g., oil spills) and small-object pollutants (e.g., plastic bottles, foam) simultaneously, as illustrated in Fig. 3. Additionally, MFF is combined with DCNv4 to refine pollutant boundary detection, reduce background interference, and improve segmentation precision. While CNN extracts local features, Transformer captures global features, allowing the model to better perceive distant pollutants and enhance pollution source tracking.

In YOLOv11's feature extraction process (Backbone and Neck), MFF is incorporated to optimize feature fusion, enhancing detection accuracy and model generalization. In the Backbone, DCNv4 is employed for convolutional operations, improving the model's adaptability to pollutant deformation and enhancing feature extraction flexibility. In the Neck, Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) are integrated to fuse multi-scale features, thereby improving small-object detection. Additionally, Transformer

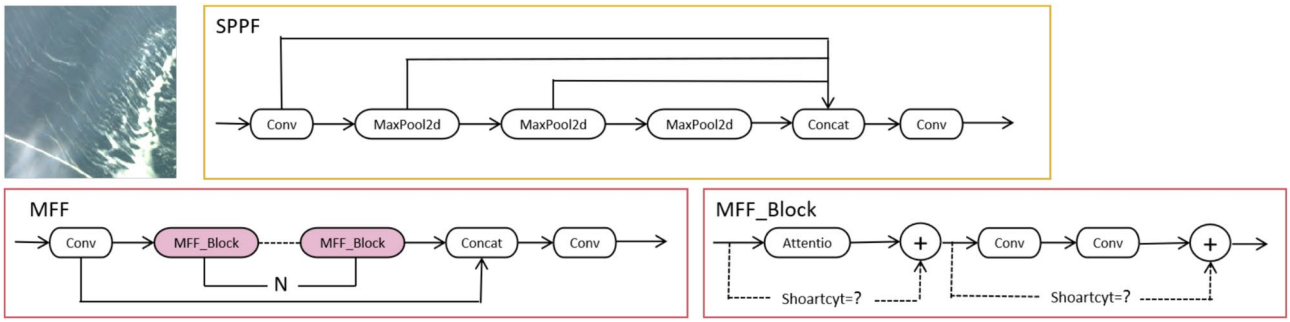

Fig. 3. MFF feature structure diagram.

modules are applied to high-level semantic feature maps, facilitating long-range pollutant detection and making the model well-suited for large-scale remote sensing and underwater monitoring tasks.

We enhance the detection ability by fusing multi-scale feature maps $F_1, F_2, ..., F_n$, as shown in Eq. 7:

$$F_{MFF} = \sum_{i=1}^{n} \alpha_i F_i \tag{7}$$

where F_i represents features from different hierarchical levels, and α_i is the fusion weight, which adjust the contribution of features from different scales. During feature extraction, Self-Attention is incorporated to optimize feature fusion via the attention mechanism, as shown in the formula 8:

$$Z_i = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{8}$$

where Q, K, V represent the Query, Key, and Value matrices, d_k is the scaling factor (typically set as the feature dimension), and Z_i is the final fused feature representation. Finally, we perform a weighted fusion of local features extracted by CNN (F_{CNN}) and global features extracted by Transformer ($F_{Transformer}$), as shown in the formula Eq. 9:

$$F_{final} = \lambda_1 F_{CNN} + \lambda_2 F_{Transformer} \tag{9}$$

where λ_1, λ_2 are feature fusion weights, which are optimized experimentally.

MFF leverages DCNv4 to enhance shape-adaptive feature extraction, improving detection of deformable pollutants such as diffused oil spills and marine debris. It further provides more precise multi-scale features, serving as an enhanced optimization target for MFL, thus improving overall detection accuracy. Additionally, multi-scale features improve instance segmentation, refining pollutant boundary detection and reducing false positives.

Experiment

Dataset

The dataset used to train the proposed network model is a combination of the MADOS dataset[32] and the Trash-ICRA19 dataset[33]. Figure 4 presents a series of sample images.

The MADOS dataset is a high-resolution remote sensing dataset specifically designed for marine pollution monitoring, covering various pollutant types, including oil spills, marine debris, foam, sea snot, and turbid water. Derived from Sentinel-2 satellite imagery, it includes multispectral band information, allowing deep learning models to leverage spectral features for enhanced pollutant detection accuracy. The key characteristics of MADOS include multi-class pollutant annotations, high-resolution imagery, and multi-temporal data acquisition, providing a robust foundation for training deep learning detection models such as YOLOv11. The dataset covers a broad geographical area, with images from various regions, including coastal zones, inland seas, and open ocean areas. This geographical diversity helps the model generalize to different marine environments, though there may still be biases introduced by the specific regions sampled. Additionally, the dataset includes varying imaging conditions such as different weather conditions, time of day, and cloud cover, which can impact detection accuracy. Additionally, MADOS incorporates pixel-level semantic annotations, enabling semantic segmentation tasks to improve pollutant boundary detection. To enhance the model's detection robustness, various data augmentation techniques, such as illumination adjustments, color normalization, and scale variations, were applied. Class distribution within the MADOS dataset is somewhat imbalanced, with certain pollutant types (e.g., oil spills) being more frequent than others (e.g., foam or sea snot), which may require balancing techniques during model training. Furthermore, the dataset was optimized using MFF to improve the model's performance across different pollutant scales.

The Trash-ICRA19 dataset is a multi-scene marine debris detection dataset, encompassing surface debris, underwater trash, and artificial object remnants. It is primarily used for training and evaluating drones, remote

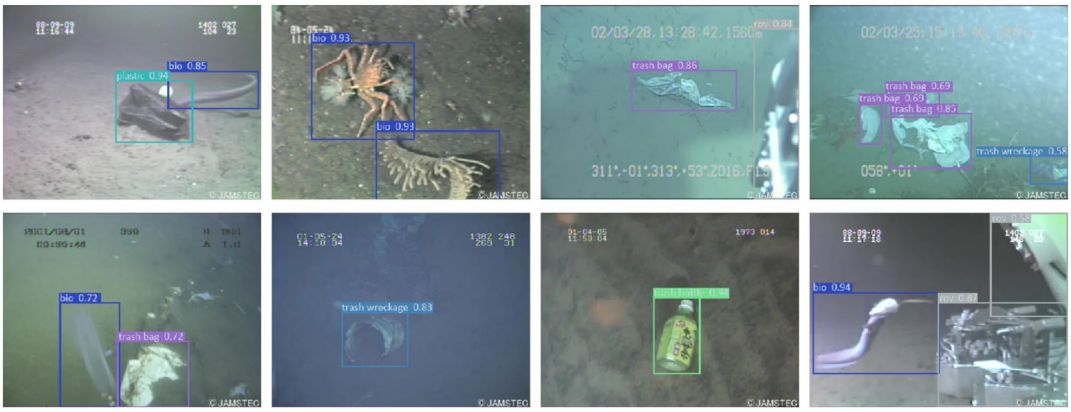

Fig. 4. Examples of the different categories.

Experimental Environment	Value
Operating System	Ubuntu 20.04
Deep Learning Framework	Pytorch 2.0
Programming Language	Python 3.9
CPU	AMD EPYC 7742 64-Core
GPU	NVIDIA A100 80GB GPU
RAM	1TB

Table 2. Experimental environmental parameters.

Hyperparametric	Value
Learning Rate	0.001
Batch Size	32
Epoch	150
Optimizer	AdamW

Table 3. Hyperparameter configuration.

sensing equipment, and underwater robots in tasks such as garbage detection, classification, and tracking. This dataset consists of diverse pollutant images captured from drone aerial photography and underwater robotic cameras, providing real-world scenarios with varied underwater lighting conditions and different pollutant scales, ensuring the model's adaptability to complex marine environments. This diversity in imaging conditions (e.g., variations in water clarity, lighting, and depth) further enhances the robustness of the model but could also introduce biases related to the types of equipment used for image capture. Trash-ICRA19 contains five major categories: Remotely Operated Vehicles (ROV), trash wreckage, trash bags, plastic bottles, and biological waste. To enhance YOLOv11's performance in marine debris detection, the dataset underwent bounding box format conversion, instance segmentation mask generation, and data augmentation. The class distribution in Trash-ICRA19 is also imbalanced, with some categories (e.g., plastic bottles) being more prevalent than others (e.g., biological waste), which may lead to challenges in training models that generalize well across all categories. Additionally, DCNv4 convolutions were integrated to improve the model's ability to detect deformable trash, while MFL was implemented to optimize the loss function, increasing the model's detection accuracy and recall for marine debris.

Implementation details

This study conducted experiments under the Ubuntu 20.04 operating system, mainly using the PyTorch 2.0 deep learning framework and GPU-accelerated training based on CUDA 11.7. All experiments were run on an NVIDIA A100 80GB GPU server equipped with an AMD EPYC 7742 64-Core CPU, 1TB RAM, and Python 3.9 environment. The experimental environment parameters are shown in Table 2. We used the AdamW optimizer for training, with the initial learning rate set to 0.001, and used the Cosine Annealing scheduler for dynamic adjustment. The batch size was set to 32, the number of training rounds was 150, and mixed precision (AMP) training was used to improve computational efficiency. The experimental hyperparameter configuration is shown in Table 3. The experiment uses the MADOS dataset (marine pollution detection) and the Trash-ICRA19 dataset

(marine garbage detection) for training and testing. We mix the two datasets and divide the mixed dataset into training, validation, and test sets in a ratio of 80:10:10. To ensure the generalization ability of the model, we performed 5-fold cross-validation. Each fold uses a different subset of the data for training and testing, which allows for a more thorough evaluation of the model's performance and reduces the risk of overfitting. Additionally, the split is randomized to prevent any potential biases.

Although the model is optimized for better performance, it should be noted that the training of large-scale datasets, particularly the high-resolution images used for marine pollution detection, is computationally expensive. The experiments were run on an NVIDIA A100 80GB GPU, which is capable of handling such tasks, but the training process required considerable computational resources, especially with mixed precision training and the integration of DCNv4 and MFF.

In addition, we conducted a series of ablation studies to evaluate the contribution of DCNv4 and MFF to detection accuracy, and compared them with YOLOv8 and YOLOv10 to verify the advantages of the improved YOLOv11 model proposed in this study in terms of pollutant boundary recognition, small target detection, and adaptability to complex environments. While the 5-fold cross-validation significantly enhanced the generalization of the model, it also led to increased computational cost. Each fold requires separate training and evaluation, which demands additional processing time and memory. These factors should be considered when deploying the model in real-time or resource-constrained environments. To address these concerns, future work will focus on optimizing the model's computational efficiency without compromising detection accuracy. The performance evaluation is measured using mAP (Mean Average Precision), Precision, Recall, and F1-score, and the inference time is calculated to evaluate the real-time performance of the model.

Evaluation metrics

To comprehensively evaluate the performance of the improved YOLOv11 in marine pollutant detection, this study employs the following key metrics: Precision, Recall, Average Precision (AP), Mean Average Precision (mAP), and Detection Time. These metrics assess the model's classification capability across different pollutant categories, detection accuracy, recall rate, and inference speed, ensuring its adaptability to complex marine environments. True Positive (TP) represents correctly detected pollutant objects, False Positive (FP) refers to incorrectly detected pollutants, and False Negative (FN) denotes undetected pollutants.

(1) Precision

Precision measures the proportion of correctly identified pollutants among all detected instances. The calculation method is as shown in Eq. 10:

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

A high precision score indicates a low false positive rate, reducing misclassifications such as mistaking ocean waves for oil spills.

(2) Recall

Recall measures the proportion of actual pollutants successfully detected by the model. A high recall score suggests a low false negative rate, which is particularly crucial for detecting small pollutants such as plastic debris. The definition is as shown in Eq. 11:

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

(3) Average Precision (AP)

$P(r)$ represents the precision at recall level r. AP is obtained by computing the precision at various recall levels and averaging them. It evaluates detection performance for individual pollutant classes, making it particularly useful for assessing different categories in the MADOS and Trash-ICRA19 datasets. The AP calculation is as shown in Eq. 12:

$$AP = \int_0^1 P(r)dr \tag{12}$$

(4) Mean Average Precision (mAP)

mAP reflects the overall detection performance across the dataset and is one of the most widely used metrics in object detection tasks. Here, N represents the total number of pollutant categories, and AP_i is the AP score for the i^{th} category. The calculation method of mAP is as shown in Eq. 13:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{13}$$

(5) Detection Time

Detection time evaluates the time required to process each image, including both forward inference time and post-processing time. It directly affects the real-time performance of the model, and the calculation is as shown in Eq. 14:

$$Detection\ Time = \frac{Total\ Inference\ Time}{Number\ of\ Images} \tag{14}$$

Results

Figures 5 and 6 illustrate the loss function variation over 500 and 600 training epochs, respectively. The trend of the loss curve indicates that as training progresses, the loss value rapidly decreases and gradually converges. In the early training phase (first 50 epochs), the loss decreases sharply, suggesting that the model quickly learns effective features. As training continues (100-300 epochs), the loss decline slows, and the network enters a steady optimization phase. After 400 epochs, the loss stabilizes and reaches a low value at 500 epochs, indicating that the model has effectively learned the dataset features. The curve remains smooth without significant fluctuations, demonstrating a stable training process with no gradient explosion or vanishing gradient issues. Additionally, the final converged loss value is small, confirming that the model achieves a good fit without significant underfitting or overfitting. To determine whether 500 epochs are sufficient, the training was extended to 600 epochs. The loss curve remains stable after 500 epochs, with no significant decrease, indicating that model optimization has reached its limit and further training does not substantially improve performance. The difference between the loss curves of 500 and 600 epochs is minimal, confirming that 500 epochs are sufficient for achieving optimal performance. Additional training beyond this point does not provide noticeable benefits and may lead to overfitting. Therefore, considering computational efficiency and model optimization, 500 epochs is identified as

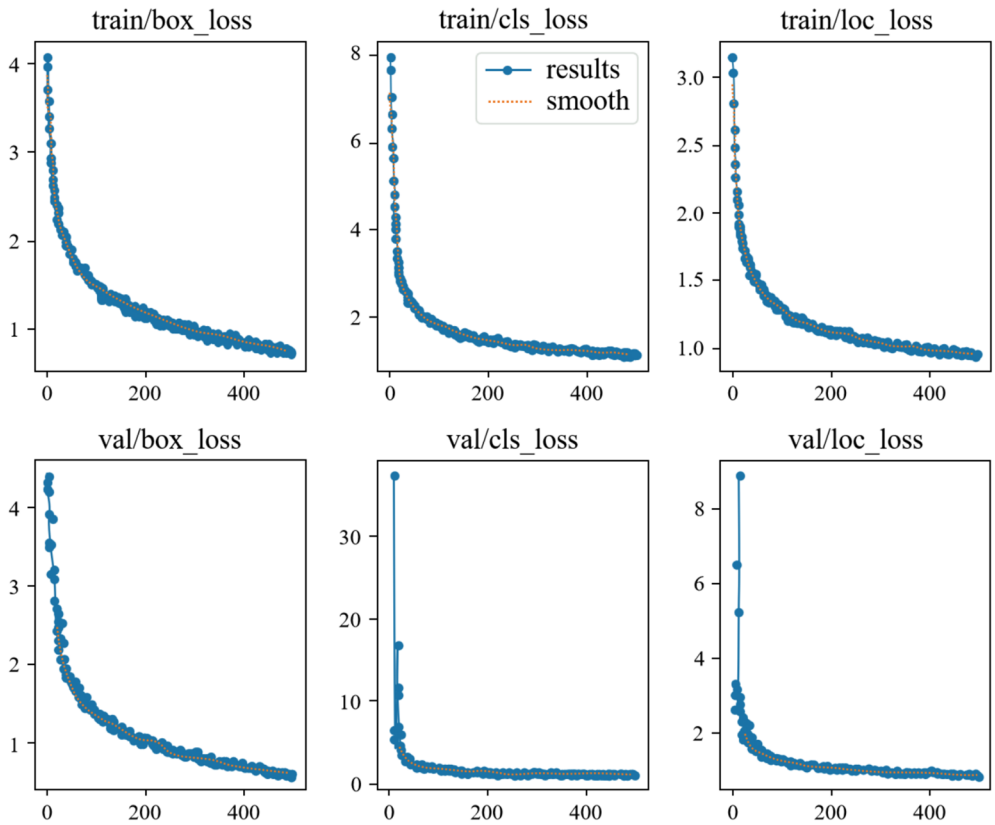

Fig. 5. Loss function curve after 500 rounds of training.

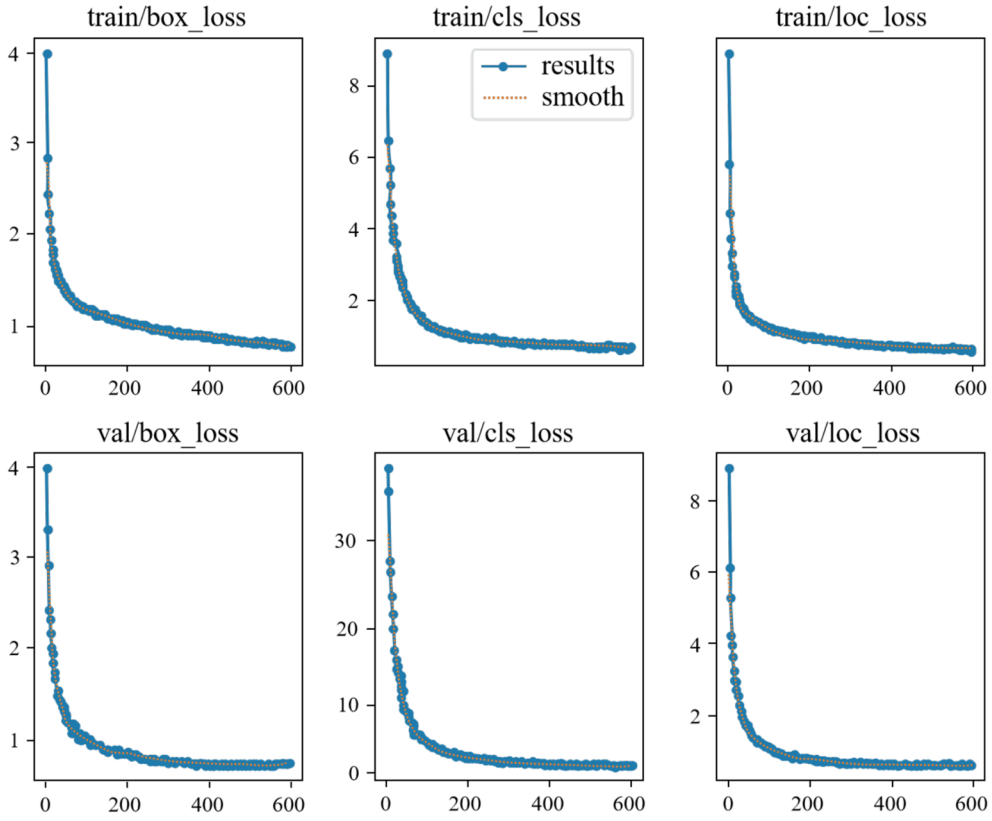

Fig. 6. Loss function curve after 600 rounds of training.

the optimal training duration, balancing performance improvement while avoiding unnecessary computational resource consumption.

In object detection tasks, the Precision-Recall (PR) curve is an essential tool for evaluating model performance. It visually illustrates the trade-off between precision and recall at different thresholds, with curves closer to the upper-right corner indicating superior model performance. Figure 7 presents the PR curves for different pollutant categories and the overall performance under the mAP@0.5 metric. The proposed YOLOv11n+DCNv4 model demonstrates high detection precision and recall across most categories. Specifically, categories such as Bio (0.954), Trash Bottle (0.948), and Foam (0.946) exhibit the best detection performance, indicating that the model can accurately identify these pollutants. Oil Spill (0.938), Trash Wreckage (0.924), and Trash Bag (0.922) also show high detection accuracy, confirming the model's ability to handle complex pollutants like oil spills and marine debris. Additionally, categories like ROV (0.885), Shallow Water (0.885), and Sea Snot (0.884) maintain stable PR curves, demonstrating strong generalization capabilities.

However, more complex water environments, such as Turbid Water (0.882), Waves & Wakes (0.878), and Natural Organic Material (0.842), result in slightly lower PR curves, suggesting a higher false detection rate in these cases. This can be attributed to challenges such as feature overlap, water surface reflections, and lighting variation. Despite this, the overall detection performance remains high. Sediment-Laden Water (0.836) exhibits the lowest PR curve, likely due to its feature similarities with other water pollution categories, which makes it a more challenging detection task. Limited training data for this category also contributes to the reduced performance. The overall mAP@0.5 metric reaches 0.902, confirming that the improved YOLOv11n model achieves high detection accuracy and robust generalization in marine pollutant detection tasks. The PR curve trend suggests that the model effectively balances precision and recall, maintaining high detection accuracy while minimizing missed detections, making it well-suited for real-world marine pollution monitoring applications.

Table 4 presents a comparison of various state-of-the-art object detection models on the mixed dataset of MADOS and Trash-ICRA19. The results indicate that the proposed YOLOv11-based model outperforms all competing models in key metrics. Compared to EfficientDet, Fast R-CNN, YOLOv5s, YOLOv8n, YOLOv10n, and YOLOv11n, the proposed model achieves an 87.1% precision, 84.3% recall, 90.2% mAP@0.5, and 78.6% mAP@0.5-0.75, while maintaining an inference speed of 3.5ms per image. Compared to YOLOv11n (precision 85.8%, mAP@0.5 89.5%), the improvements stem from several key optimizations, including DCNv4, which enhances adaptability to pollutant shape variations, enabling better detection of large-scale pollutants (e.g., oil spills) and small-scale pollutants (e.g., plastic debris). Additionally, the model optimizes classification loss, improving focus on pollution regions, reducing false detections, and increasing accuracy. Furthermore, MFF integrates CNN and Transformer structures to enhance multi-scale pollutant detection in complex marine environments, particularly for long-range pollutant recognition. Although the inference speed is slightly

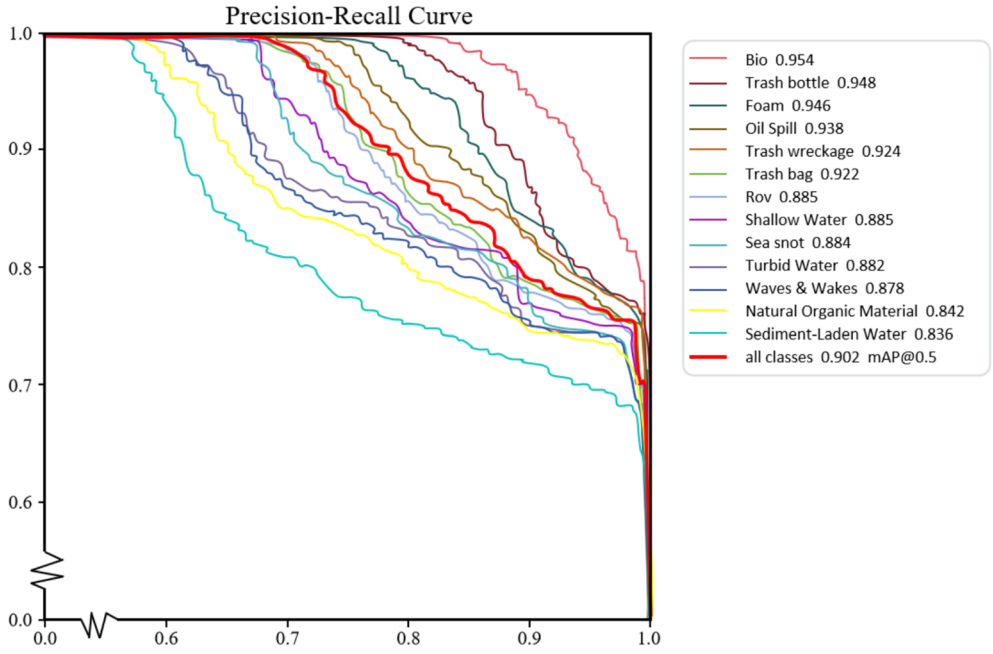

Fig. 7. PR curve graph.

Model	Precision(%)	Recall(%)	mAP50(%)	mAP50-75(%)	Detection Time(ms)
EfficientDet[16]	82.3	52.6	68.5	58.7	3.5
Fast RCNN[15]	78.5	75.6	73.6	67.2	**3.3**
YOLOv5s[34]	79.2	81.4	78.6	73.1	3.4
YOLOv8n[35]	81.2	83.7	79.5	74.8	3.6
YOLOv10n[36]	84.3	83.9	87.3	75.4	3.8
YOLOv11n[13]	85.8	84.2	89.5	76.1	3.9
Ours	**87.1**	**84.3**	**90.2**	**78.6**	3.5

Table 4. Contrast the experimental results. Bold values indicate the best performance.

higher than YOLOv10n and YOLOv11n, the substantial improvement in mAP validates the effectiveness of these optimization strategies, positioning the proposed model as the most suitable solution for marine pollutant detection. Figure 8 further illustrates the mAP@50 values for each category across different models. In summary, the improved YOLOv11 model delivers superior accuracy in detecting marine pollutants such as oil spills, plastic debris, turbid water, and sea snot. It demonstrates strong application potential in drone-based surveillance, satellite remote sensing, and underwater robotic inspections, providing an efficient and precise solution for intelligent marine pollution monitoring.

Table 5 shows the ablation experiment results of this study on the YOLOv11n baseline model, which primarily evaluates the contribution of different modules to the detection performance. It compares the performance of the YOLOv11n baseline model, the final model after adding the marine pollution feature adaptive loss function, adding MFF, and adding both MFF and the loss function. The experimental results show that the YOLOv11n baseline model demonstrates reasonable pollutant detection capability, achieving a precision of 84.3%, a recall of 81.5%, an mAP50 of 85.4%, and an mAP50-75 of 73.5%. However, there is still room for further optimization.

When DCNv4 is added to the baseline model, the precision increases to 84.9%, the recall rises to 82.6%, the mAP50 improves to 86.5%, and the mAP50-75 increases to 75.3%. This demonstrates that DCNv4 can effectively optimize the classification loss, making the model focus more on the pollutant area while reducing background interference and improving classification accuracy. When only MFF is added, the precision improves significantly to 86.3%, the recall increases to 84.2%, the mAP50 rises to 89.5%, and the mAP50-75 increases to 77.2%. This shows that MFF enhances the model's ability to detect pollutants across different scales by combining CNN and Transformer structures, particularly in complex marine environments. The recognition of long-distance pollutants is notably better.

Finally, when both the marine pollution feature adaptive loss function and MFF are added, the model's precision increases further to 87.1%, the recall reaches 84.3%, the mAP50 achieves 90.2%, and the mAP50-75 rises to 78.6%, reaching the best performance. This illustrates the synergistic effect of these two components: the adaptive loss function improves the classification accuracy of pollutant areas and reduces false detections,

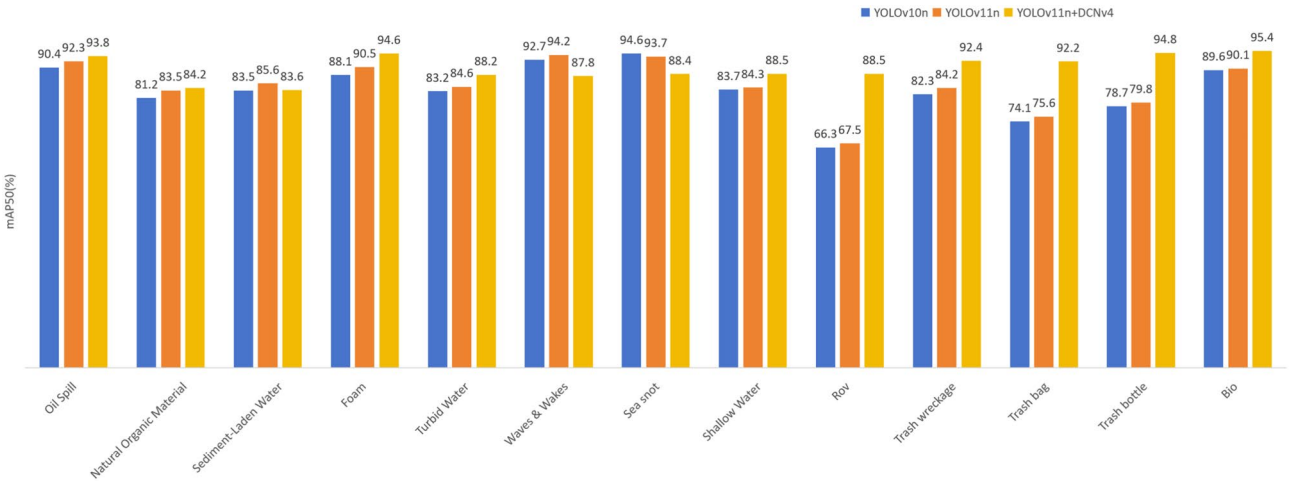

Fig. 8. Comparison of mAP 50 values across different models for each category in the validation set of the mixed dataset.

Model	Precision(%)	Recall(%)	mAP50(%)	mAP50-75(%)
YOLOv11n(Baseline)	84.3	81.5	85.4	73.5
YOLOv11n+	84.9	82.6	86.5	75.3
YOLOv11n+MFF	86.3	84.2	89.5	77.2
YOLOv11n++MFF	**87.1**	**84.3**	**90.2**	**78.6**

Table 5. Results of ablation experiments. Bold values indicate the best performance.

Model	Precision(%)	Recall(%)	mAP50(%)	mAP50-75(%)	Detection Time(ms)
YOLOv11n+SEAMhead	78.5	71.9	72.4	65.8	3.8
YOLOv11n+LSCD	79.4	72.1	72.6	68.4	3.7
YOLOv11n+EfficientHead	80.8	73.4	73.1	69.5	**3.3**
YOLOv11n+Dyhead	81.4	75.6	78.6	73.5	3.8
YOLOv11n+DCNv4	**87.1**	**84.3**	**90.2**	**78.6**	3.5

Table 6. Comparison of YOLOv11n+DCNv4 with other detection heads. Bold values indicate the best performance.

while MFF integrates a multi-scale feature fusion mechanism, which helps the model extract target features from different scales and complex backgrounds, thereby improving detection robustness. In summary, the improved YOLOv11n model proposed in this study has shown significant performance improvements in pollutant detection tasks, especially in complex marine environments. The detection of targets such as oil pollution, marine debris, turbid water, and marine mucus has been further optimized, providing a more accurate and efficient solution for practical applications.

Table 6 presents the ablation experiment results of the YOLOv11n baseline model with different detection heads, aiming to analyze the impact of different detection head structures on small target detection. The YOLOv11n+DCNv4 scheme achieved the best performance across all indicators, demonstrating the effectiveness of DCNv4 in enhancing small target detection. In the baseline YOLOv11n model, the SEAMhead structure performed relatively poorly, with a precision of 78.5%, recall of 71.9%, an mAP50 of 72.4%, and an mAP50-75 of 65.8%, indicating its limited ability to detect small targets in complex backgrounds. The LSCD detection head showed slight improvement, with precision rising to 79.4% and mAP50 reaching 72.6%, but it still failed to significantly improve small target detection robustness. In contrast, the EfficientHead structure showed an advantage in detection time, requiring only 3.3ms, but its mAP50 was only 73.1%, indicating that although it has high computational efficiency, it still has limitations in detection accuracy. The Dyhead structure performed moderately, with mAP50 of 78.6% and mAP50-75 of 73.5%, indicating that its multi-scale feature learning ability helps to improve small target detection, but it is not optimal.

In contrast, the YOLOv11n+DCNv4 solution proposed in this study achieved the best performance across all core indicators, with precision reaching 87.1%, recall of 84.3%, mAP50 rising to 90.2%, and mAP50-75 reaching 78.6%. This is primarily due to the introduction of the DCNv4 structure. This module enhances the model's

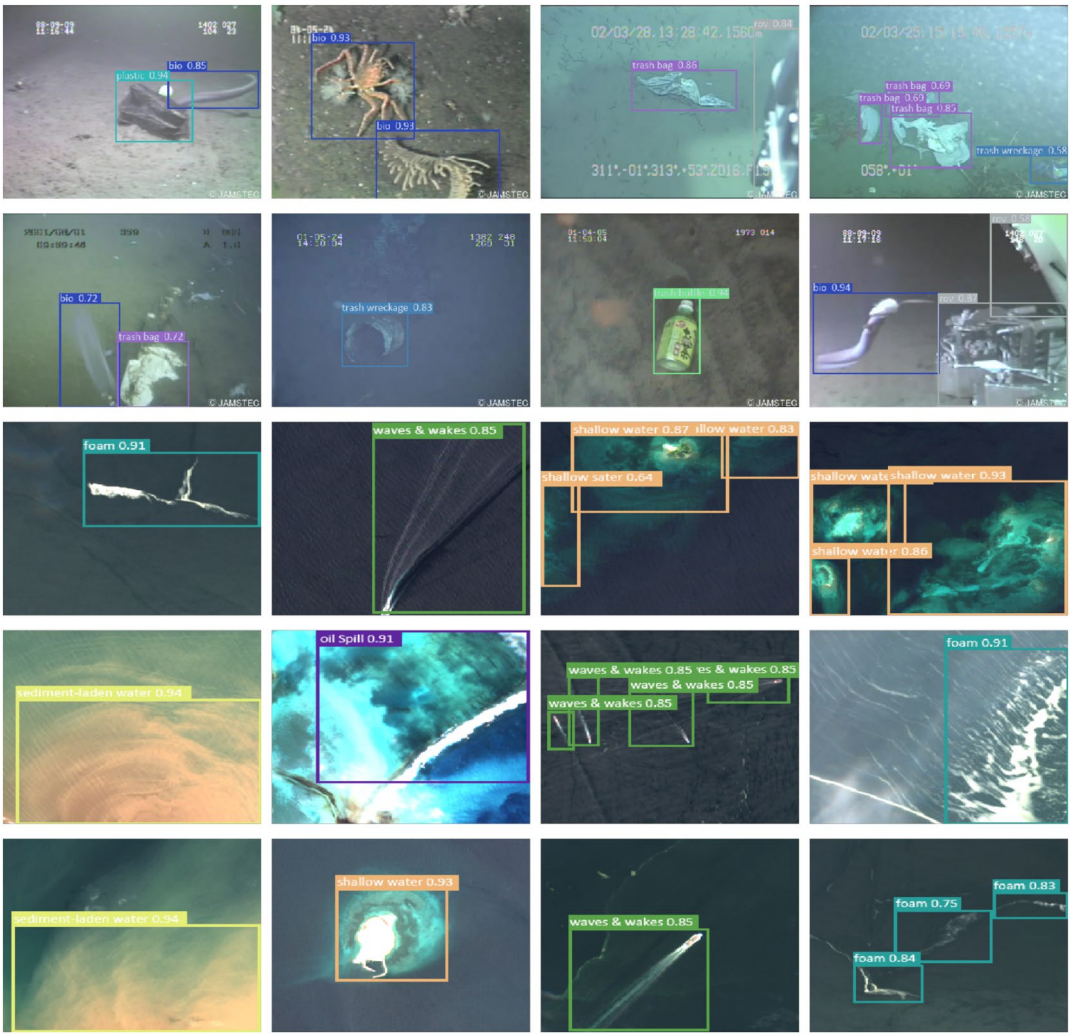

Fig. 9. A Schematic representation of the detection results of the improved model for different categories of target images.

adaptability to small and deformed targets through dynamic sampling offsets and modulation mechanisms, enabling the detection network to more accurately capture pollutant boundary information. Additionally, DCNv4 removes the fixed window limitation of traditional convolutions, adaptively adjusting the receptive field to ensure that the model can detect pollutants of different scales, especially when detecting small pollutants at a distance (e.g., floating plastics and small oil spills), with higher detection accuracy and stronger robustness.

In terms of detection time, the YOLOv11n+DCNv4 solution takes 3.5ms, which is slightly longer than EfficientHead (3.3ms), but more balanced compared to Dyhead (3.8ms) and SEAMhead (3.8ms). This demonstrates that the method proposed in this study not only enhances small target detection capabilities but also maintains high computational efficiency, meeting the needs of real-time applications. In summary, the ablation experiments show that the YOLOv11n+DCNv4 solution significantly outperforms other detection heads in precision, recall, and mAP, effectively detecting small pollutant targets in complex marine environments, such as floating garbage, turbid water, and marine mucus. This validates the superiority of DCNv4 as a detection head and provides a more accurate and efficient solution for marine pollution monitoring.

To illustrate the validity of the proposed algorithm, the visual results of the hybrid dataset test are shown in Fig. 9. As seen from the figure, the improved YOLOv11n can successfully identify multiple small and overlapping targets with high recognition accuracy.

Conclusion

This study presents an improved model for marine pollution monitoring, focusing on the challenges posed by irregular pollutant shapes, large-scale variations, and complex environmental conditions. The proposed approach is built on the YOLOv11n architecture, enhanced with Deformable Convolutional Networks version 4 (DCNv4) to improve adaptability to deformable pollutants, and Multi-scale Feature Fusion (MFF) to combine the strengths of CNN and Transformer-based feature extraction. Additionally, Marine Fusion Loss (MFL) is introduced to optimize detection weights and enhance pollutant classification accuracy. Instance segmentation

is integrated to improve boundary recognition, contributing to more precise detection. Experimental results on the MADOS and Trash-ICRA19 datasets show that the proposed model outperforms existing methods in key performance metrics, achieving an mAP50 of 90.2% and an inference time of 3.5ms, demonstrating both high accuracy and computational efficiency.

Despite these advancements, there are two main areas for future improvement: (1) the model's high computational resource requirements, particularly due to the integration of DCNv4 and Transformer features, which, although enhancing detection performance, present challenges for real-time inference on embedded systems; and (2) the detection of small target pollutants, such as microplastics and floating debris, which remain prone to false positives and missed detections due to their similarity to background features. Future work should focus on optimizing the model's architecture for better efficiency on low-power devices and exploring advanced techniques such as super-resolution reconstruction or self-supervised learning to further enhance small object detection.

In conclusion, the proposed model provides an efficient and accurate solution for marine pollution detection, with potential applications in drone-based inspections, satellite remote sensing, and underwater robotic monitoring. This work contributes valuable technical advancements to the field of marine environmental protection, offering a practical approach for real-world monitoring of marine pollution.

Data availability

The datasets used in this study are publicly available, with the MADOS dataset accessible at https://marine-pollution.github.io/ and the Trash-ICRA19 dataset at https://conservancy.umn.edu/items/c34b2945-4052-48fa~-b7e7-ce0fba2fe649. The code supporting this study is available upon request from the corresponding author and will be publicly released on GitHub upon acceptance of the manuscript.

Received: 27 March 2025; Accepted: 29 May 2025
Published online: 01 July 2025

References

1. Topouzelis, K., Papageorgiou, D., Suaria, G. & Aliani, S. Floating marine litter detection algorithms and techniques using optical remote sensing data: A review. *Mar. Pollut. Bull.* **170**, 112675. https://doi.org/10.1016/j.marpolbul.2021.112675 (2021).
2. Cai, Y., Chen, L., Zhuang, X. & Zhang, B. Automated marine oil spill detection algorithm based on single-image generative adversarial network and yolo-v8 under small samples. *Mar. Pollut. Bull.* **203**, 116475 (2024).
3. Liu, H., Wang, M., Tang, H. & Zhang, H. Progress in research on marine litter-related monitoring technologies. *J. Phys: Conf. Ser.* **2679**, 012055. https://doi.org/10.1088/1742-6596/2679/1/012055 (2024).
4. Zhou, W., Zheng, F., Yin, G., Pang, Y. & Yi, J. Yolotrashcan: a deep learning marine debris detection network. *IEEE Trans. Instrum. Meas.* **72**, 1–12 (2022).
5. Samuriwo, T. N., Babalola, O. P., Sparks, C. A. J., Davidson, I. E. & Dieng, L. Comparative review of remote sensing methods for ocean plastic litter detection. *IEEE Access* **12**, 166126–166161. https://doi.org/10.1109/ACCESS.2024.3494660 (2024).
6. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016).
7. Poloju, N. & Rajaram, A. Transformation with yolo tiny network architecture for multimodal fusion in lung disease classification. *Cybernetics and Systems* 1–22 (2024).
8. He, K., Gkioxari, G., Dollar, P. & Girshick, R. Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (2017).
9. Liu, W. et al. Ssd: Single shot multibox detector. In Leibe, B., Matas, J., Sebe, N. & Welling, M. (eds.) *Computer Vision – ECCV 2016*, 21–37 (Springer International Publishing, Cham, 2016).
10. Gomaa, A. & Saad, O. M. Residual channel-attention (rca) network for remote sensing image scene classification. *Multimedia Tools and Applications* 1–25 (2025).
11. Wang, J., Li, F., Lv, S., He, L. & Shen, C. Physically realizable adversarial creating attack against vision-based bev space 3d object detection. *IEEE Transactions on Image Processing* (2025).
12. Wang, J., Li, F. & He, L. A unified framework for adversarial patch attacks against visual 3d object detection in autonomous driving. *IEEE Transactions on Circuits and Systems for Video Technology* (2025).
13. Khanam, R. & Hussain, M. Yolov11: An overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725 (2024).
14. Gomaa, A. Advanced domain adaptation technique for object detection leveraging semi-automated dataset construction and enhanced yolov8. In *2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, 211–214 (IEEE, 2024).
15. Ren, S., He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031 (2017).
16. Tan, M., Pang, R. & Le, Q. V. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10781–10790 (2020).
17. Doherty, J., Gardiner, B., Kerr, E. & Siddique, N. Bifpn-yolo: One-stage object detection integrating bi-directional feature pyramid networks. *Pattern Recogn.* **160**, 111209. https://doi.org/10.1016/j.patcog.2024.111209 (2025).
18. Carion, N. et al. End-to-end object detection with transformers. In Vedaldi, A., Bischof, H., Brox, T. & Frahm, J.-M. (eds.) *Computer Vision – ECCV 2020*, 213–229 (Springer International Publishing, Cham, 2020).
19. Jiang, L., Liu, F., Lv, J., Liu, B. & Wang, C. Gst-yolo: A lightweight visual detection algorithm for underwater garbage detection. *J. Real-Time Image Proc.* **21**, 114 (2024).
20. Xiong, Y. et al. Efficient deformable convnets: Rethinking dynamic and sparse operator for vision applications. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5652–5661 (2024).
21. Liu, J., Shen, X., Wang, J., Jiang, L. & Zhang, H. An intelligent defect detection approach based on cascade attention network under complex magnetic flux leakage signals. *IEEE Trans. Industr. Electron.* **70**, 7417–7427 (2022).
22. Peng, J. et al. Mlsa-yolo: A multi-level feature fusion and scale-adaptive framework for small object detection. *J. Supercomput.* **81**, 528 (2025).
23. Hao, M. et al. Coarse to fine-based imagepoint cloud fusion network for 3d object detection. *Inf. Fusion* **112**, 102551. https://doi.org/10.1016/j.inffus.2024.102551 (2024).
24. Wang, T. et al. Multidimensional fusion of frequency and spatial domain information for enhanced camouflaged object detection. *Information Fusion* 102871 (2024).
25. Liu, Z. et al. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 10012–10022 (2021).

26. Yin, W., Lingxin, S., Maohuan, L., Qianlai, S. & Xiaosong, L. Pv-yolo: Lightweight yolo for photovoltaic panel fault detection. *IEEE Access* **11**, 10966–10976 (2023).
27. Zhang, H. et al. Dino: Detr with improved denoising anchor boxes for end-to-end object detection. arXiv preprint arXiv:2203.03605 (2022).
28. Shen, A., Zhu, Y., Angelov, P. & Jiang, R. Marine debris detection in satellite surveillance using attention mechanisms. *IEEE J. Selected Top. Appl. Earth Observ. Remote Sens.* **17**, 4320–4330 (2024).
29. Salem, M., Gomaa, A. & Tsurusaki, N. Detection of earthquake-induced building damages using remote sensing data and deep learning: A case study of mashiki town, japan. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, 2350–2353 (IEEE, 2023).
30. Gomaa, A. & Abdalrazik, A. Novel deep learning domain adaptation approach for object detection using semi-self building dataset and modified yolov4. *World Electric Veh. J.* **15**, 255 (2024).
31. Ahn, S., Chang, J.-W. & Kang, S.-J. An efficient accelerator design methodology for deformable convolutional networks. In *2020 IEEE International Conference on Image Processing (ICIP)*, 3075–3079 (IEEE, 2020).
32. Kikaki, K., Kakogeorgiou, I., Hoteit, I. & Karantzalos, K. Detecting marine pollutants and sea surface features with deep learning in sentinel-2 imagery. *ISPRS J. Photogramm. Remote. Sens.* **210**, 39–54. https://doi.org/10.1016/j.isprsjprs.2024.02.017 (2024).
33. Fulton, M. S., Hong, J. & Sattar, J. Trash-icra19: A bounding box labeled dataset of underwater trash. *International Conference on Robotics and Automation (ICRA)* 5752–5758 (2019).
34. Wen, G. et al. Yolov5s-ca: A modified yolov5s network with coordinate attention for underwater target detection. *Sensors* **23**, 3367 (2023).
35. Liu, L., Chu, C., Chen, C. & Huang, S. Marineyolo: Innovative deep learning method for small target detection in underwater environments. *Alex. Eng. J.* **104**, 423–433. https://doi.org/10.1016/j.aej.2024.07.126 (2024).
36. Wang, A. et al. Yolov10: Real-time end-to-end object detection. In Globerson, A. *et al.* (eds.) *Advances in Neural Information Processing Systems*, vol. 37, 107984–108011 (Curran Associates, Inc., 2024).

Acknowledgements

This study was supported by Henan Intelligent Manufacturing Digital Twin Engineering Research Center, Henan Province's 2021 Funding Project for Discipline and Specialty Construction in Private Regular Higher Education Institutions (Document No. 179 of the Education Office of Henan Province on Political and Legal Affairs, Software Engineering Specialty).

Author contributions

F.W. Data curation, Methodology, Formal analysis, Visualization, Writing-original draft, Writing–review & editing;

Additional information

Correspondence and requests for materials should be addressed to F.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com