Robotica (2025), **43**, pp. 1867–1884
doi:10.1017/S0263574725000475

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

REFNet: reparameterized feature enhancement and fusion network for underwater blur target recognition

Benshun Li[1] and Lei Cai[2]

[1]School of Mechanical and Electrical Engineering, Henan University of Technology, Zhengzhou, P.R. China
[2]School of Artificial Intelligence, Henan Institute of Science and Technology, Xinxiang, P.R. China
Corresponding author: Lei Cai; Email: cailei2014@126.com

Received: 30 September 2024; **Revised:** 15 February 2025; **Accepted:** 7 March 2025; **First published online:** 16 May 2025

Keywords: blurred target recognition; you only look once; information fusion; lightweight architecture; deep learning; attention mechanism

Abstract

The underwater target detection is affected by image blurring caused by suspended particles in water bodies and light scattering effects. To tackle this issue, this paper proposes a reparameterized feature enhancement and fusion network for underwater blur object recognition (REFNet). First, this paper proposes the reparameterized feature enhancement and gathering (REG) module, which is designed to enhance the performance of the backbone network. This module integrates the concepts of reparameterization and global response normalization to enhance the network's feature extraction capabilities, addressing the challenge of feature extraction posed by image blurriness. Next, this paper proposes the cross-channel information fusion (CIF) module to enhance the neck network. This module combines detailed information from shallow features with semantic information from deeper layers, mitigating the loss of image detail caused by blurring. Additionally, this paper replace the CIoU loss function with the Shape-IoU loss function improves target localization accuracy, addressing the difficulty in accurately locating bounding boxes in blurry images. Experimental results indicate that REFNet achieves superior performance compared to state-of-the-art methods, as evidenced by higher mAP scores on the underwater robot professional competitionand detection underwater objects datasets. REFNet surpasses YOLOv8 by approximately 1.5% in $mAP_{50:95}$ on the URPC dataset and by about 1.3% on the DUO dataset. This enhancement is achieved without significantly increasing the model's parameters or computational load. This approach enhances the precision of target detection in challenging underwater environments.

1. Introduction

Underwater target detection plays a crucial role in areas like marine science and environmental monitoring. With the advancement of autonomous underwater vehicle technology [1], the accurate and efficient detection and identification of underwater targets has become a research focus. However, the complexity of underwater environments, including light attenuation, suspended particles, and scattering effects, often results in blurred images. These issues significantly affect the accuracy and robustness of underwater target detection, posing major challenges for such tasks [2].

In recent years, the rapid development of deep learning technologies has facilitated their widespread application across various fields [3, 4], owing to their notable efficiency and accuracy. While most current object detection algorithms perform well in terrestrial applications [5, 6], they require improvements to address the complexities of underwater environments (Figure 1). Some studies tackle the challenge of underwater target detection by using image enhancement techniques to improve detail and contrast, which in turn enhances detection accuracy [7, 8]. Some other researchers focus on designing new multi-scale fusion network architectures to enhance the feature extraction capabilities of networks, thereby

© The Author(s), 2025. Published by Cambridge University Press.

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

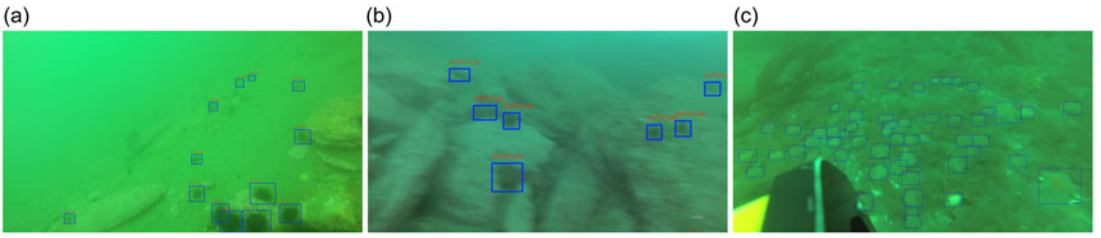


Figure 1. *Challenges in underwater target detection: (a) Fuzziness due to water turbidity; (b) fuzziness caused by movement; (c) low contrast.*

improving the accuracy of underwater target detection [9, 10]. However, using image enhancement methods alone to against underwater blurring and other complexities has limitations, and complex network structures often incur substantial computational complexity and network parameters.

To address these challenges and enhance algorithm efficiency for underwater operations, this paper presents an enhanced YOLOv8-based framework. This method improves feature extraction and fusion mechanisms for complex underwater environments without substantially increasing computational burden, thus enhancing detection accuracy for challenging targets. The key contributions of this work are as follows:

1. To mitigate the impact of image blurriness on network feature extraction, this paper proposes an enhanced backbone network for YOLOv8 incorporating a REG module. This module boosts the network's feature extraction capabilities while simultaneously reducing computational load and the number of parameters. Consequently, it improves detection performance in complex scenarios while preserving the network's lightweight architecture.
2. To address the loss of detailed information during feature extraction in blurry and complex environments, this paper proposes an improved neck network for YOLOv8 using a CIF module. This module strengthens the network's feature representation by improving feature fusion across cross-layer, thereby boosting target detection performance in challenging environments.
3. To overcome challenges in achieving precise bounding box localization due to factors like blurriness, this paper proposes the Shape-IoU loss function for bounding box regression. This loss function enables the network to predict and adjust bounding boxes more accurately, thereby improving detection accuracy in complex environments.

The structure of the paper is as follows: Section 2 provides a review of related work. Section 3 introduces the REFNet detector, detailing its backbone network, feature fusion framework, and novel loss function. Section 4 presents experimental results and discusses performance on benchmark datasets. Finally, Section 5 provides the conclusion.

2. Related work

2.1. Underwater target recognition algorithms

Target detection technologies have made significant strides in recent years, transitioning from handcrafted-feature approaches to approaches based on deep learning. Traditional detection methods primarily consist of feature-based algorithms and template matching algorithms, relying on manually designed features and heuristic rules. However, these approaches are sensitive to variations in lighting and complex environments, which limits their practical applicability.

Rapid advancements in deep learning technologies have notably improved underwater target detection. Contemporary architectures, including YOLO, SSD, and Faster R-CNN leverage enhanced feature extractors and region proposal networks. Additionally, some researchers have introduced multi-scale

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

feature enhancement techniques to improve the network's ability to detect targets of varying scales, further increasing detection precision [11, 12].

The single-stage object detection algorithm YOLOv8 [13] has shown outstanding performance among these advancements. YOLOv8 significantly optimizes its architecture by leveraging the advantages of its predecessors in the YOLO series, resulting in a lighter and faster model. This makes it particularly suitable for real-time target detection on autonomous underwater vehicles. Consequently, many researchers have further developed YOLOv8 to explore its potential in practical underwater target detection applications. For instance, Qu et al. [14] significantly enhanced the accuracy and efficiency of underwater target detection by integrating the lightweight and efficient LEPC module, the AP-FasterNet architecture, and the CARAFE upsampling operation. Similarly, Liu et al. [15] enhanced the YOLO network to improve detection efficiency through innovations such as the multi-path fusion block, lightweight components like GSConv and VoV-GSCSP, and an adaptive power transformation label assignment strategy. Nevertheless, further optimization of the YOLOv8 network is still possible, particularly in addressing challenges posed by complex underwater environments.

2.2. Attention mechanism

Recently, attention mechanisms have been widely used in underwater target detection. Detecting targets in complex underwater environments continues to be a significant challenge. Inspired by human biological systems, attention mechanisms improve detection efficiency and accuracy by focusing on the most relevant information [16]. Fei et al. [17] enhanced the network's sensitivity to small objects through the integration of a lightweight mixed local context attention, thereby improving the accuracy of underwater object detection. Quan et al. [18] proposed a novel network based on YOLOv8, incorporating SE attention and shuffle attention mechanisms to enhance detection performance. Similarly, Zhuo et al. [19] introduced a dual-terminal attention mechanism into the YOLOv3 network, effectively reducing noise in marine biological datasets and thereby improving the accuracy and stability of target detection.

Attention mechanisms are generally classified into spatial attention, channel attention, and mixed attention. STNs [20] dynamically adjust input images by learning spatial transformation parameters, helping the model focus on regions of interest. SENet [21] introduces a channel weighting mechanism that dynamically adjusts each channel's response, amplifying important channels and reducing interference. DAN [22] integrates spatial and channel attention, enabling joint learning of relationships between spatial positions and feature channels, thereby enhancing performance in visual tasks.

While channel attention mechanisms significantly boost the performance of convolutional neural networks (CNNs), complex attention modules may also increase model complexity and computational load. To address this, ECA-Net [23] improves performance with minimal additional parameters. Unlike the channel attention module in SENet, which captures channel relationships via global average pooling and fully connected layers, ECA-Net employs a more nuanced approach by utilizing local interactions between channels. By eliminating fully connected layers, ECA-Net reduces model complexity. The ECA module facilitates inter-channel interaction through one-dimensional convolution, employing adaptively selected kernel sizes to capture dependencies between channels, thereby balancing performance and efficiency.

This paper integrates ECA attention into both the CIF and REG modules to enhance the network's channel interaction capabilities, preserving essential feature information, and improve feature extraction performance.

2.3. Feature fusion network

Early classic object detectors [24, 25] depend on the final feature maps produced by convolutional neural networks for object detection. Shallow features capture fine details, whereas deeper features offer more robust semantic information. To fully leverage both shallow and deep features, Lin et al. [26] introduced the FPN architecture to fuse features across multiple levels.

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

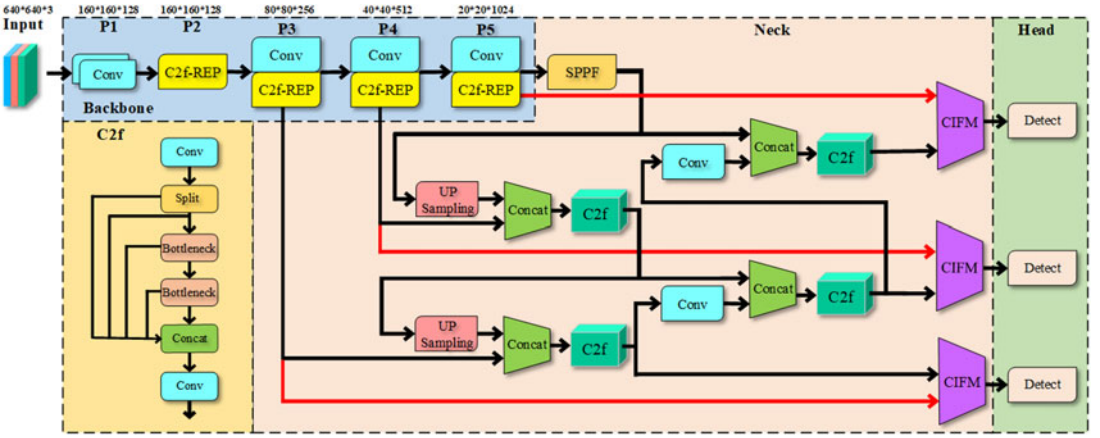


Figure 2. *The network architecture of the REFNet detector.*

FPNs integrate high-level features into low-level features to enable multi-scale feature fusion. However, FPNs have limited localization capabilities for high-level features due to their single top-down pathway. PANet [27] introduced a bottom-up path that shortens the information transmission path and enhances multi-scale feature fusion. NAS-FPN [28] employs neural architecture search to replace the original FPN structure, achieving higher training accuracy at the cost of increased computational complexity. Chen et al. [29] introduced a frequency-aware feature fusion method that enables the network to effectively integrate features from different frequencies, thereby enhancing detection efficiency in dense image prediction tasks. Recently, Sun et al. [30] proposed a multi-level feature fusion network with a multilevel fusion module that combines feature enhancement and interaction to improve feature representation and detection efficiency. Zhang et al. [31] enhanced the network's detection performance by designing a multi-scale adaptive feature fusion module combined with an edge-aware guidance module, which better captures multi-scale information and improves feature learning at blurred boundaries.

This paper presents the CIF module and introduce a novel inter-layer connection method to enhance the network's adaptability for detecting ambiguous targets. This method effectively combines the detailed information from shallow features with the semantic richness of deep features, resulting in enhanced detection precision for fuzzy targets.

3. Method

This paper presents the REFNet for underwater blurred object recognition. REFNet is an optimized version of YOLOv8, and the network architecture of the REFNet detector is illustrated in Figure 2. Specifically, the paper introduces the REG module into the backbone network, replacing the Bottleneck module in the original C2f module. The REG module enhances the network's feature extraction capabilities through reparameterized depthwise convolution, the incorporation of the ECA attention mechanism, a response normalization mechanism, and the use of residual connections. These improvements enhance both the model's ability to interpret complex scenes and its capacity to extract diverse target features. The paper proposes the CIF module prior to the head detection in the neck network. This module enhances the REFNet's ability to integrate contextual information and capture detail features by fusing low-level and high-level features, thereby improving its performance in detecting fuzzy targets. Finally, this paper replaces CIoU with Shape-IoU as the network's loss function, improving the localization and regression capabilities of the bounding box and enabling more accurate detection of underwater blurred targets.

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

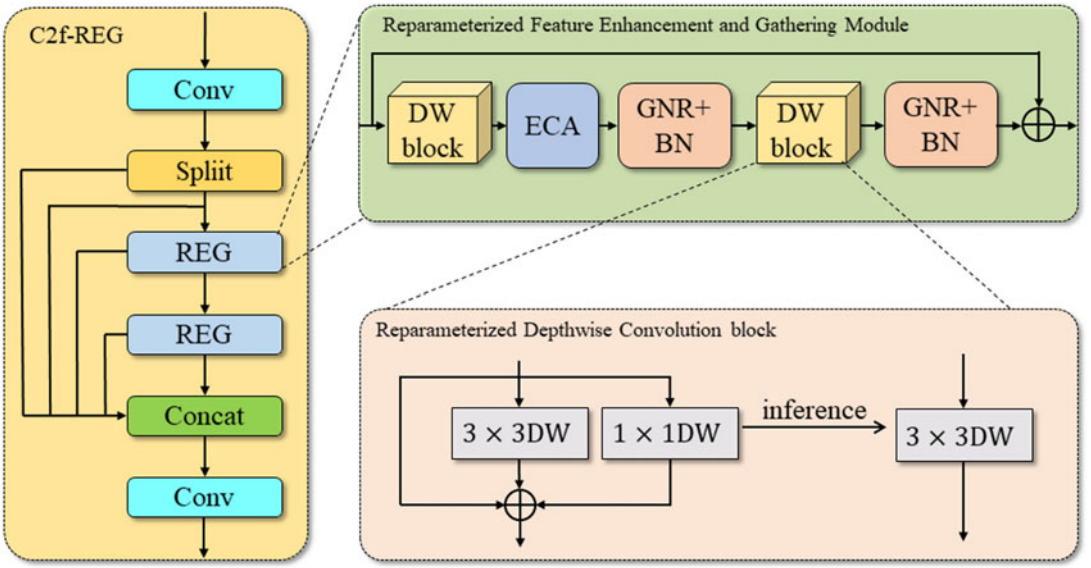


Figure 3. *Detailed structure diagram of the REG module.*

3.1. Reparameterized feature enhancement and gathering module

After preprocessing, the image is input into the backbone network for feature extraction. To enhance the network's feature extraction capabilities, this paper presents architectural optimizations for the backbone, as shown in Figure 3. Additionally, it proposes the REG module as the foundation for constructing the new C2f module. These innovations significantly improve the network's image feature extraction performance.

The REG module incorporates structural reparameterization with deep convolution based on the UniRepvit [32] network, which enhances performance during the training phase through a multi-branch topological structure. This multi-branch structure captures more features during training, thereby improving the model's expressiveness. The multi-branch structure is consolidated into a single-branch configuration during the inference process, significantly reducing computational and memory overhead. This design leverages the advantages of the multi-branch structure during training while ensuring computational and memory efficiency during inference.

$$Conv3(X) = Conv1(X_{i1}) \oplus Conv2(X_{i1}) \oplus X_{i1} \tag{1}$$

The reparameterization process for deep convolutional computations is illustrated in Eq. (1). Conv1(·) and Conv2(·) respective depthwise separable convolutions with kernel sizes of $C\times3\times3$ and $C\times1\times1$,respectively.

Next, the ECA mechanism is integrated within the REG module to better capture inter-channel information. ECA attention enhances the model's response to critical features by assigning varying weights to channels based on their importance, thereby improving overall feature extraction accuracy.

Additionally, the REG module integrates the global response normalization (GRN) [33] mechanism, which further improves the effectiveness of feature extraction. GRN consists of three sequential steps: First, global feature aggregation condenses spatial features into a vector using global average pooling or similar methods, summarizing the spatial information of the image into a global feature representation. Next, normalization is applied to these global features, calculating the relative importance of each channel, thereby enhancing the module's response to key channels. Finally, feature calibration applies these normalized weights to enhance channel contrast and feature selectivity, effectively improving the network's discriminative power.

$$U_1 = \gamma(\beta(Conv3(\beta(\gamma(\alpha(Conv3(X_{i1})))))))) \tag{2}$$

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

$$U_2 = \gamma(\beta(Conv1(\beta(\gamma(\alpha(Conv1(X_{i1})))))))) \tag{3}$$

The training and inference workflows of the REG module are represented in Eq. (2) and Eq. (3), respectively. Among them, α, β, and γ correspond to ECA attention, GRN, and batch normalization (BN) operations, respectively.

$$Y_{i1} = U_1 \oplus X_{i1} \tag{4}$$

$$Y_{i2} = U_2 \oplus X_{i1} \tag{5}$$

The output processes for training and inference in the REG module are represented by Eq. (4) and Eq. (5), respectively.

The proposed advancements significantly enhance the performance and diversity of feature extraction within the REG block, optimizing the trade-off among accuracy, processing speed, and computational efficiency. Specifically, the reparameterization technique facilitates the capture of finer-grained details during the training phase, improving the quality of feature extraction. ECA focuses the model on critical features by appropriately weighting channels, thereby enhancing the precision of feature representation. GRN optimizes feature extraction through global feature aggregation and normalization, improving contrast and selectivity. Furthermore, the incorporation of residual connections in the REG module strengthens feature extraction and enhances the model's generalization ability. Collectively, these enhancements enable the model to more effectively tackle complex image tasks, achieving a balance between superior performance and practical application.

3.2. Cross-channel information fusion module

Deep neural networks possess formidable learning capabilities, particularly excelling at capturing high-level semantic information. Nonetheless, they face significant challenges in detecting targets within complex environments. As convolutional neural networks are deepened, they become increasingly proficient at extracting clear target features and enhancing semantic information. Detection becomes challenging in the presence of underwater blurred targets, as targets features are often obscured by complex and variable environmental conditions. To address this issue, the CIF module addresses this challenge by improving the detail expression capability of deep networks.

Low-level features encompass a wealth of information regarding target positions and details, which is pivotal for recognizing indistinct target objects. To further refine feature representation, this paper proposes a spatial refinement attention fusion module. This module aggregates channel context features at each spatial location, leveraging the spatial detail present in low-level features to enrich high-level features, thereby enhancing the recognizability of the target. In the spatial refinement attention fusion module, the ECA attention mechanism is first employed to effectively preserve important feature information. Subsequently, a convolutional network processes these features to learn the importance weights for each channel. Finally, these weights are normalized and multiplied with the deep features to achieve effective embedding of low-level detail information.

$$W(X) = \mu(\varepsilon(Conv1(\theta(\varepsilon(Conv1(\delta(X_{i1})))))))) \tag{6}$$

$$out(X) = W(X) \otimes Y_{i1} \tag{7}$$

Spatial refinement focuses on feature refinement and fusion processes within the fusion module, as formulated in Eq. (6) and Eq. (7). Among them, Conv1$(\cdot)$ denote the 1D convolution with kernel sizes of C $\times$ 3 $\times$ 3. δ, ε, θ, and μ respetively represent ECA attention, BN, and ReLU. and Sigmoid functions.

In addition, deep features encode richer semantic information, facilitating enhanced comprehension of target characteristics of target characteristics. However, relying exclusively on deep features remains insufficient for effective detection. Therefore, this paper introduces a channel attention fusion module to improve feature representation. This module adjusts the importance of each channel, integrating shallow

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

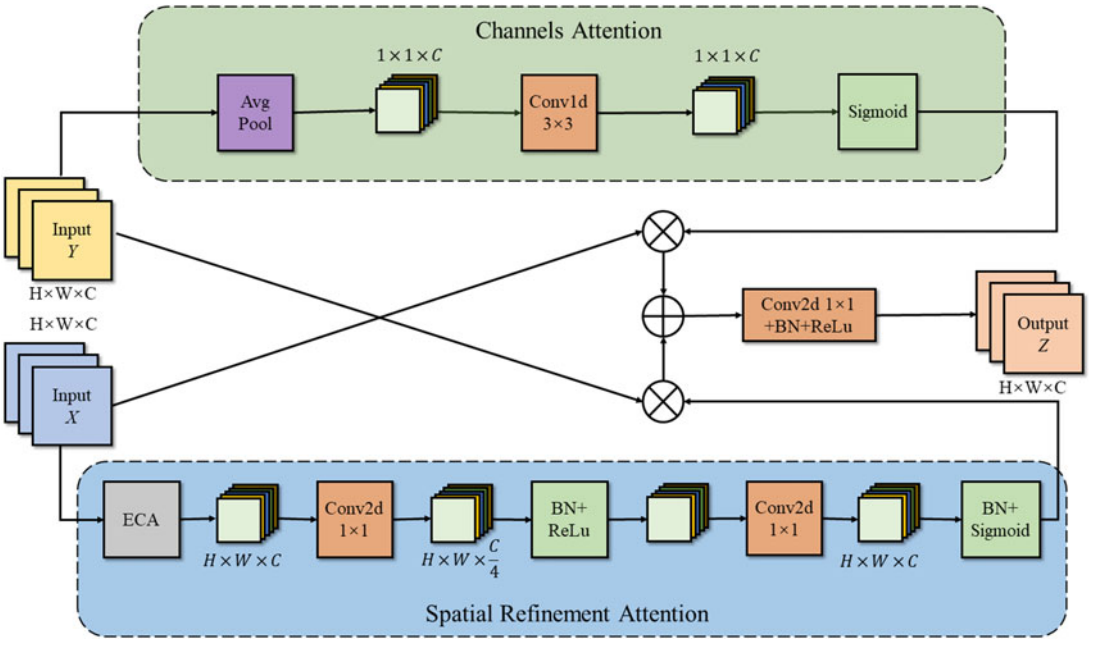


Figure 4. *Detailed structure of the CIF module.*

and deep features to improve feature representation. Specifically, the channel attention fusion module first obtains a global representation of each channel through global average pooling. These descriptors are then processed by a convolutional network to assess the significance of each channel. After normalization, these weights are applied to the shallow features, thereby embedding semantic information from deep features and optimizing feature representation.

$$W(Y) = \mu(Conv2(Avgpool(Y_{i1}))) \tag{8}$$

$$out(Y) = W(Y) \otimes X_{i1} \tag{9}$$

The fusion mechanism of the channel attention fusion module is mathematically defined in Eq. (8) and Eq. (9). Among them, Conv2(·), respectively, represents the 2D convolution of C × 3 × 3.

$$Z = \sum \theta(\varepsilon(Conv1(W(Y) \otimes X_{i1} \oplus W(X) \otimes Y_{i1}))) \tag{10}$$

The fusion output process of the CIF module is illustrated in Eq. (10). Among them, input X_{i1} and input Y_{i1} have the same shape and number of channels.

In summary, the CIF module effectively integrates detailed information from low-level features into deep features through the synergistic integration of spatial refinement attention fusion and channel attention fusion. Simultaneously, it propagates semantic information from high-level features to low-level features. Figure 4 illustrates the detailed structure of the CIF module. This approach not only enhances the representation of fine details in deep features but also improves the semantic understanding of low-level features, capturing more accurate and rich feature information for underwater target detection tasks. Consequently, the CIF module significantly strengthens strengthens the network's capability to detect fuzzy targets in complex environments, addressing the limitations of deep networks in detecting targets under challenging fuzzy conditions.

3.3. Loss function

In object detection frameworks, the design of bounding box loss function is essential for enabling the network to achieve accurate target detection results. Its primary purpose is to optimize the error between

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

the predicted bounding box and the actual object position, making sure that the predicted box aligns as closely as possible with the true boundaries of the object. In the factors scenario of underwater object detection, environmental complexity often leads to a loss of image details, increasing the likelihood of deviations or errors in target localization by detection networks. To address this, this paper adopts Shape-IoU [34] serving as the bounding box loss function, which enhances localization precision in challenging environments.

The key advantage of Shape-IoU lies in its focus on the shapes and sizes of bounding boxes, rather than solely on the area of overlap. This method facilitates more accurate bounding box regression, thereby improving the model's performance in complex environments. The Shape-IoU loss function explicitly models geometric factors including the width, height, and position of the bounding boxes, allowing for a more detailed depiction of discrepancies between predicted and actual boxes. Specifically, the calculation of Shape-IoU follows a formula that reflects these characteristics. The formula for calculating Shape-IoU is as follows:

$$IOU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \tag{11}$$

$$ww = 2 \times \frac{(w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \tag{12}$$

$$hh = 2 \times \frac{(h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \tag{13}$$

$$distance^{shape} = hh \times \frac{(x_c - x_c^{gt})^2}{c^2} + ww \times \frac{(y_c - y_c^{gt})^2}{c^2} \tag{14}$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-w_t})^{\theta}, \theta = 4 \tag{15}$$

$$\begin{cases} w_w = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})} \\ w_h = ww \times \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases} \tag{16}$$

Among them, 'Scale' reflects the target's relative size distribution within the dataset. Meanwhile, ww and hh are coefficients that adjust for the width and height of the bounding box, respectively, based on the shape of the ground truth box. The formula for calculating the bounding box regression loss is as follows:

$$L_{Shape-IoU} = 1 - IoU + distance^{shape} + 0.5 \times \Omega^{shape} \tag{17}$$

Integrating Shape-IoU as the bounding box loss function enables the model to demonstrates enhanced robustness the complexities of underwater environments, improving both the accuracy and robustness of the predicted bounding boxes. The Shape-IoU metric is designed to account not only for the overlap between boxes but also to incorporate factors such as shape and size, leading to superior performance across diverse datasets. In summary, Shape-IoU provides a geometrically aware loss formulation for object detection, making it particularly suitable for complex applications like underwater target detection.

4. Results and discussion

In this section, the paper provides a comprehensive description of the experimental setup, including the datasets used, evaluation metrics, and specific training procedures. Additionally, ablation studies are conducted to verify the contribution of each module, and the performance of REFNet is compared

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

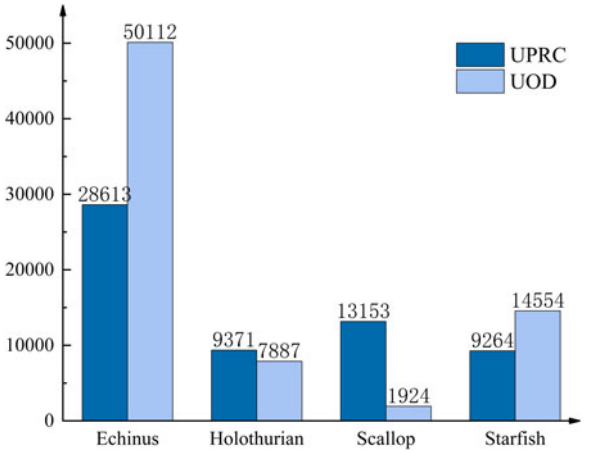


Figure 5. *Category distribution diagram of UPRC dataset and DUO dataset.*

to existing methods. These experiments thoroughly assess the model's performance, demonstrating its superiority and improvements under different conditions.

4.1. Dataset

The study utilizes two public ocean datasets, the DUO dataset and the URPC dataset, to compare and validate the proposed algorithm. The URPC dataset consists of 7543 images, including detection targets such as sea urchins, sea cucumbers, scallops, and starfish. Each image has a corresponding annotation file that records the species type and its location. The DUO dataset contains 7782 images, which include four categories of benthic marine species: sea cucumbers, sea urchins, scallops, and starfish. The category distributions for both datasets are shown in Figure 5. For the experiments, the datasets were split into training, validation, and test sets with a ratio of 7:2:1, ensuring a balanced distribution for evaluating the model's performance.

4.2. Training details

This study was conducted on a Linux-based platform equipped with an AMD EPYC 7282 processor, 250 GB of RAM, and an NVIDIA A100 80 GB PCIe GPU. For training, this paper employed a batch size of 16 and trained for 300 epochs. Hyperparameters were set as follows: an initial learning rate of 0.01, a momentum of 0.937 with SGD, and a weight decay of 0.0005. The model was trained on the URPC and DUO datasets, generating the corresponding weight files. Figs. 6 and 7 present visual results from the training process of the REFNet on these datasets. The training loss curves show a consistent decrease in both training and validation losses, with no observed increase in validation loss toward the end of training, indicating good generalization and effective performance without overfitting.

4.3. Evaluation indicators

The effectiveness of the algorithm was analyzed using evaluation metrics such as mean average precision (mAP), F1-score, precision (P), and recall (R). The following equations define true positives (TP) and false positives (FP) as the number of accurate and inaccurate positive predictions, respectively. True negatives (TN) and false negatives (FN) represent the number of accurate and inaccurate negative predictions. In multi-class classification, only the predictions of the current class are treated as positive.

$$Precision = \frac{TP}{(TP + FP)} \times 100\% \tag{18}$$

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

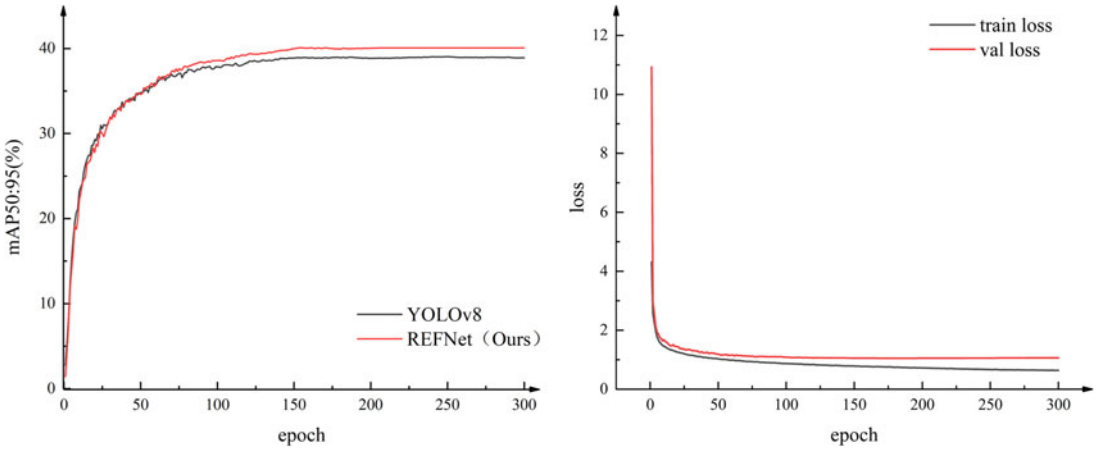


Figure 6. *Evaluation of validation accuracy and training loss for YOLOv8 versus REFNet (Ours) on the URPC dataset.*

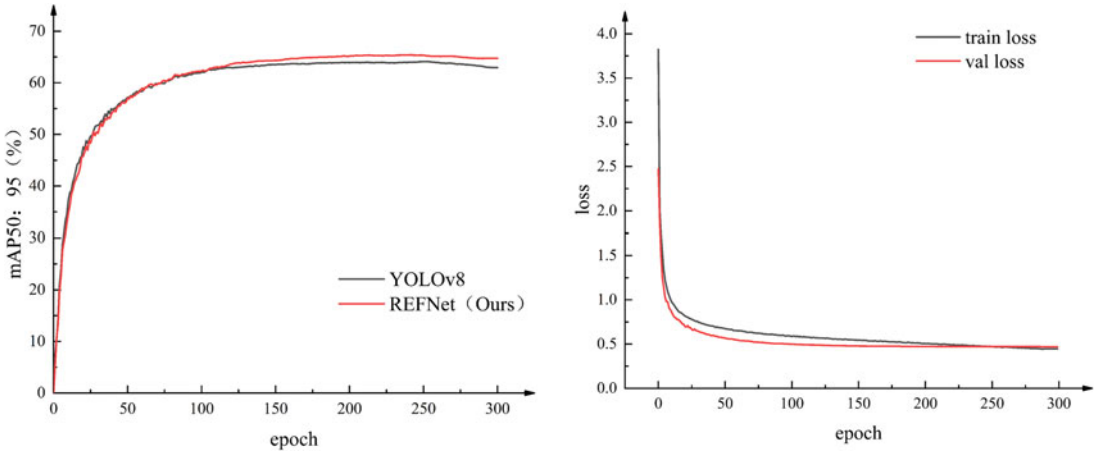


Figure 7. *Evaluation of validation accuracy and training loss for YOLOv8 versus REFNet (Ours) on the DUO dataset.*

Precision measures the algorithm's ability to correctly detect positive examples, the ratio of true positives to the total number of predicted positives.

$$Recall = \frac{TP}{(TP + FN)} \times 100\% \tag{19}$$

Recall evaluates the extent to which the algorithm misses positive examples, the proportion of true positives out of all actual positives.

$$F1 - score = \frac{2TP}{(2TP + FP + FN)} \times 100\% \tag{20}$$

F1-score quantifies the harmonic mean of precision and recall, used for a comprehensive evaluation of the algorithm's performance.

$$mAP = \frac{\sum_{x=i}^{n} (Precision_i)}{n} \times 100\% \tag{21}$$

mAP represents the average precision value across all classes within the dataset, used to measure overall detection performance.

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

Table I. *Effectiveness of the REFNet module: ablation study results on the URPC dataset.*

Model	REG	CIF	Shape-IoU	Precision(%)	Recall(%)	mAP_{50}(%)	$mAP_{50:95}$(%)
YOLOv8	×	×	×	78.2	63.7	70.7	39.1
Improve-1	×	×	√	78.1	64	71	39.5
Improve-2	×	√	×	**80.6**	62.7	71.3	39.6
Improve-3	√	×	×	79.9	63.4	71.6	40.1
Improve-4	√	√	×	80.5	63.2	71.9	40.4
REFNet	√	√	√	80.3	**64.2**	**72.2**	**40.6**

4.4. Ablation experiment

To evaluate the impact of the proposed enhancements on underwater target detection, this study conducts ablation experiments by progressively integrating the enhanced modules into the YOLOv8 detection model. The effects of each module on detection accuracy are analyzed. As shown in Table I, all improvements contribute positively to the model's overall performance.

First, this paper examines the impact of each individual improvement made to YOLOv8. Compared with the original YOLOv8, improving the backbone network by adding the REG module increased mAP_{50} and $mAP_{50:95}$ by 0.9% and 1.0%, respectively, indicating that the REG module enhances the network's feature extraction capability. Employing the CIF module to improve the neck of the YOLOv8 network increased mAP_{50} and $mAP_{50:95}$ by 0.6% and 0.5%, respectively, indicating that the CIF module enhances the network's feature fusion capability. Integrating Shape-IoU as a loss function improved mAP_{50} and $mAP_{50:95}$ by 0.3% and 0.4%, respectively, showing that Shape-IoU improves bounding box regression and localization accuracy. Additionally, combining the REG and CIF modules in YOLOv8 further increased mAP_{50} and $mAP_{50:95}$ by 1.2% and 1.3%, respectively. These ablation experiments validate the contribution of each module to enhancing the network's performance under challenging underwater conditions.

The proposed REFNet, which integrates all these improvements including the REG module for the backbone, the CIF module for feature fusion, and Shape-IoU for bounding box regression loss achieves a 1.5% improvement in both mAP_{50} and $mAP_{50:95}$ compared with the original YOLOv8. The results indicate that the REFNet significantly enhances detection accuracy in complex underwater environments, demonstrating superior performance compared with the baseline YOLOv8 model.

The superior performance of REFNet can be attributed to three key architectural innovations. First, the REG module in the backbone network employs depth-wise separable convolutions with asymmetric kernels, which effectively captures fine-grained features of blurred targets while maintaining computational efficiency. This is particularly crucial for distinguishing low-contrast underwater objects from complex backgrounds. Second, the CIF module implements cross-scale attention mechanisms during feature fusion, enabling better integration of contextual information from different receptive fields. This addresses the scale variation challenge commonly encountered in underwater scenes. Third, the adoption of Shape-IoU loss introduces geometric constraints during bounding box regression, which regularizes the prediction of distorted objects typical in underwater imagery.

4.5. Comparison of experimental results

In this comparative experiment, this paper compares the performance of the proposed REFNet algorithm with that of several state-of-the-art object detection models, including Faster R-CNN [35], CenterNet [36], DETR [37], YOLOv5 [38], YOLOv6 [39], YOLOv7 [40], YOLOv8 [41], and YOLOv10 [42], using the URPC and DUO datasets. All models were trained at a resolution of 640×640, and the performance was assessed using various evaluation metrics, including the backbone architecture, FPS (frames per second),parameter count and computational complexity. Additionally, FPS was used to assess detection speed. The comparison results, as summarized in Table II, reveal that the REFNet algorithm is among the top performers across these metrics.

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

Table II. *REFNet performance metrics compared to leading methods.*

Model	Backbone	Param(M)	FLOPs(G)	FPS
Faster-Rcnn(FPN) [35]	ResNet50	41.4	134.3	17
CenterNet [36]	ResNet50	34.5	56.7	77
DETR [37]	ResNet50	41.3	36.8	45
YOLOv5 [38]	CSPDarknet53	**1.8**	**4.2**	**144**
YOLOv6 [39]	EfficientRep	4.3	11.1	114
YOLOv7 [40]	ELAN	3.4	7.6	90
YOLOv8 [41]	CSPNet	3.0	8.2	139
YOLOv10 [42]	CSPNet	2.7	8.4	104
REFNet	CSPNet-REG	3.1	8.1	101

Table III. *Performance metrics of REFNet compared to leading methods on the URPC dataset.*

Model	Precision(%)	Recall(%)	F1-score(%)	mAP_{50}(%)	$mAP_{50:95}$(%)
Faster-Rcnn(FPN) [35]	75	64	69	70.4	36.5
CenterNet [36]	72.8	64	68	70.1	35.9
DETR [37]	76	58	66	67.3	33.8
YOLOv5 [38]	79	63.1	69	70.5	38
YOLOv6 [39]	80.3	62.8	**70**	70.7	39
YOLOv7 [40]	77.2	**64.8**	**70**	71.5	38.6
YOLOv8 [41]	78.2	63.7	69	70.7	39.1
YOLOv10 [42]	78.2	62.3	69	70.2	39
REFNet	**80.3**	64.2	**70**	**72.2**	**40.6**

Table III presents the performance metrics of REFNet compared to leading methods on the URPC dataset. The evaluation criteria include P, R, F1-score, mAP_{50}, and $mAP_{50:95}$. The comparative results in Table II and Table III highlight performance differences, with REFNet achieving the highest $mAP_{50:95}$ on the URPC dataset compared to other models. While REFNet may not lead in all evaluation metrics, it demonstrates competitive performance overall. Notably, the trade-off between accuracy and speed aligns with the requirements of underwater robotics applications where detection reliability often takes precedence over real-time performance. The 2.6% improvement in $mAP_{50:95}$ over YOLOv5 primarily stems from our enhanced feature extraction pipeline, which reduces false negatives in turbid water conditions as evidenced by the heatmap analysis. For example, despite being slightly larger in model size and computational complexity, as well as slower in detection speed compared to YOLOv5, REFNet surpasses it with a 2.6% higher $mAP_{50:95}$. REFNet also outperforms Faster R-CNN, CenterNet, and DETR across all evaluation metrics. When compared to YOLOv6, YOLOv7, and YOLOv10, REFNet shows comparable performance in terms of model size, computational load, and detection speed. However, it demonstrates superior accuracy, with $mAP_{50:95}$ values 1.6% higher than YOLOv6, 2.0% higher than YOLOv7, and 1.6% higher than YOLOv10. Although YOLOv7 has a slightly higher recall rate, REFNet's improved accuracy in complex underwater environments makes it a more suitable choice for such scenarios. In contrast to YOLOv8, REFNet sacrifices some detection speed and slightly increases model complexity but achieves a 1.5% improvement in $mAP_{50:95}$. This performance gap highlights the importance of task-specific optimization: while YOLOv7 achieves higher recall through aggressive proposal generation, our method maintains better precision via spatial-aware feature processing. The 1.5% improvement over YOLOv8 with minimal complexity increase demonstrates the effectiveness of our carefully designed modules rather than brute-force scaling.

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

Table IV. *Performance metrics of REFNet compared to leading methods on the DUO dataset.*

Model	Precision(%)	Recall(%)	F1-score(%)	mAP_{50}(%)	$mAP_{50:95}$(%)
Faster-Rcnn(FPN) [35]	77.2	73	75	79.3	54.9
CenterNet [36]	81.8	76	76	82.4	60.2
DETR [37]	76.5	72	74	79.2	54.9
YOLOv5 [38]	84.5	76	**80**	82.4	59.3
YOLOv6 [39]	83.7	**77.4**	**80**	83.8	62.8
YOLOv7 [40]	84	76.2	**80**	83.7	60.6
YOLOv8 [41]	84.8	76.3	**80**	83.7	64.1
YOLOv10 [42]	84.7	75.7	**80**	83.5	63.6
REFNet	**86.9**	76.8	**80**	**84.5**	**65.4**

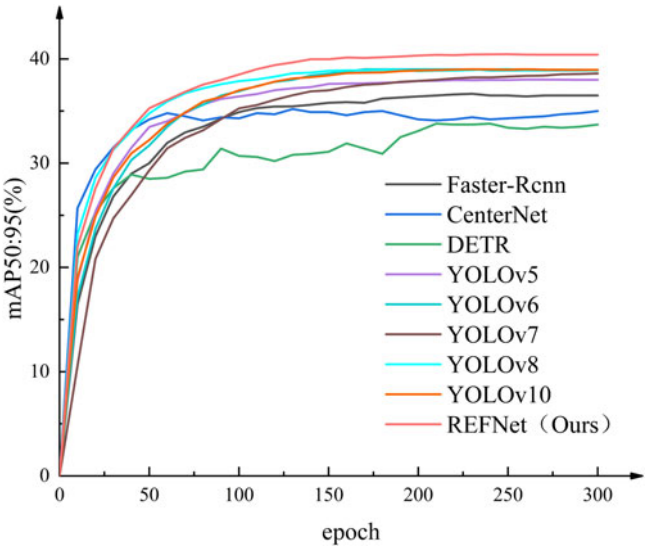


Figure 8. *MAP growth curve for target detection algorithm training on the URPC validation dataset.*

Figure 8 shows the $mAP_{50:95}$ growth curve across 300 epochs for various algorithms. The REFNet algorithm demonstrates a steady performance increase, reaching the highest $mAP_{50:95}$ within the first 100 epochs.

Table IV presents the performance metrics of REFNet compared to leading methods on the DUO dataset. While REFNet trails slightly behind YOLOv7 in recall, it achieves the highest mAP value among all methods and performs strongly across other parameters. As shown in Figure 9, REFNet reaches the highest $mAP_{50:95}$ after 100 epochs, demonstrating its effectiveness in handling underwater target detection.

Figure 10 visually compares the detection results of REFNet with recent algorithms. REFNet significantly outperforms other methods in several key aspects of detection, as shown in the figure. In the first and second rows, our method detects small, blurred distant targets that other methods fail to identify. This demonstrates REFNet's greater sensitivity and accuracy in handling blurred targets, capturing more details. In the third and fourth rows, our model identifies more low-contrast blurred targets compared to other models, which are often overlooked or misclassified. Which demonstrates its superior ability to detect fuzzy targets and handle challenging underwater environments. These visual comparisons reveal two critical advantages: 1) The REG module's multi-scale processing enables detection of distant small targets that conventional NMS tends to suppress. 2) The CIF module's channel-wise attention helps

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

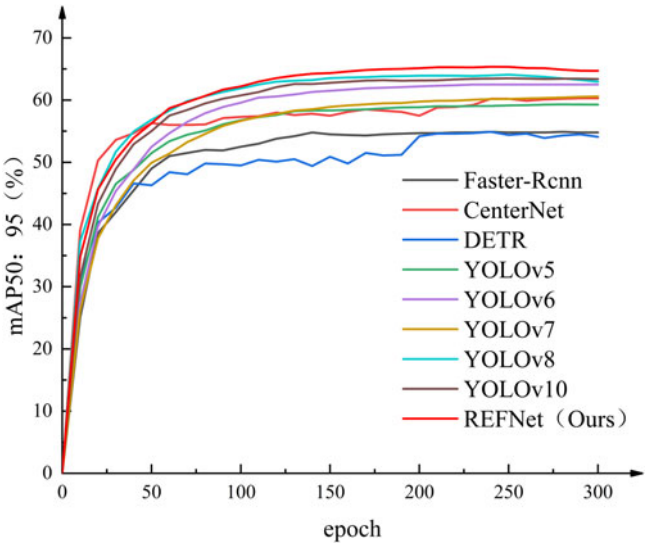


Figure 9. *MAP growth curve for target detection algorithm training on the DUO validation dataset.*

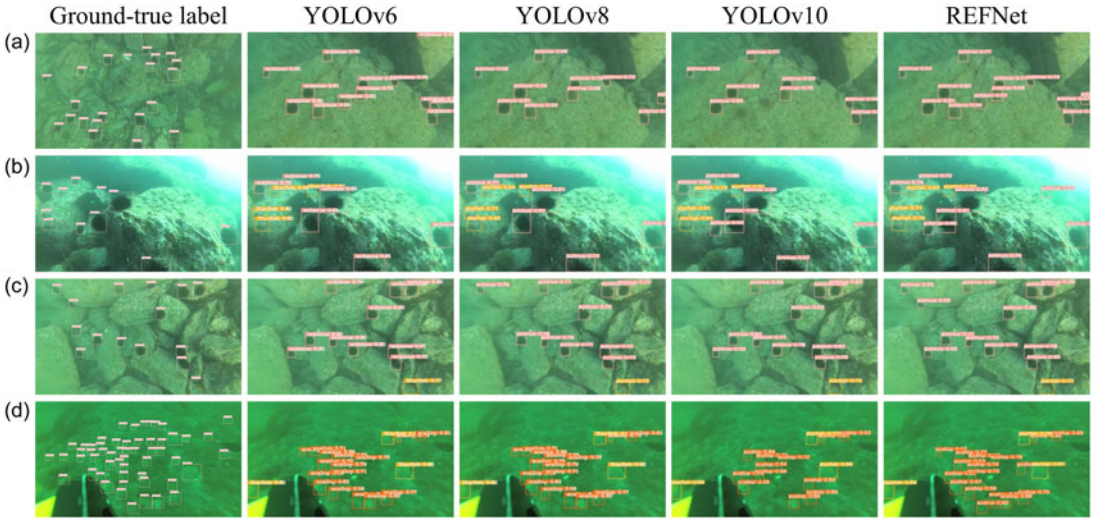


Figure 10. *Comparison of detection results of different algorithms.*

maintain target distinctiveness in low-contrast regions. Such capabilities are particularly valuable for underwater archaeology and marine life monitoring applications.

To further validate its performance, Figure 11 presents heatmaps generated by EigenCAM [43], showing the network's attention to the target area. Comparing the heatmaps in the first and third columns clearly indicates that our network performs better in separating objects from the background. This demonstrates the success of the proposed modules in improving the network's capacity to accurately differentiate between targets. Further analysis of the second and fourth columns reveals that our network focuses more effectively on the target itself. This indicates that the proposed modules allow the network to capture more precisely and detailed features, which is particularly beneficial for detecting blurred targets. Through feature fusion, REFNet progressively enhances detection accuracy and stability by focusing more on the target area. These visualized feature maps further confirm the previous discussion of REFNet's behavior, highlighting our model's clear superiority. The attention localization

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

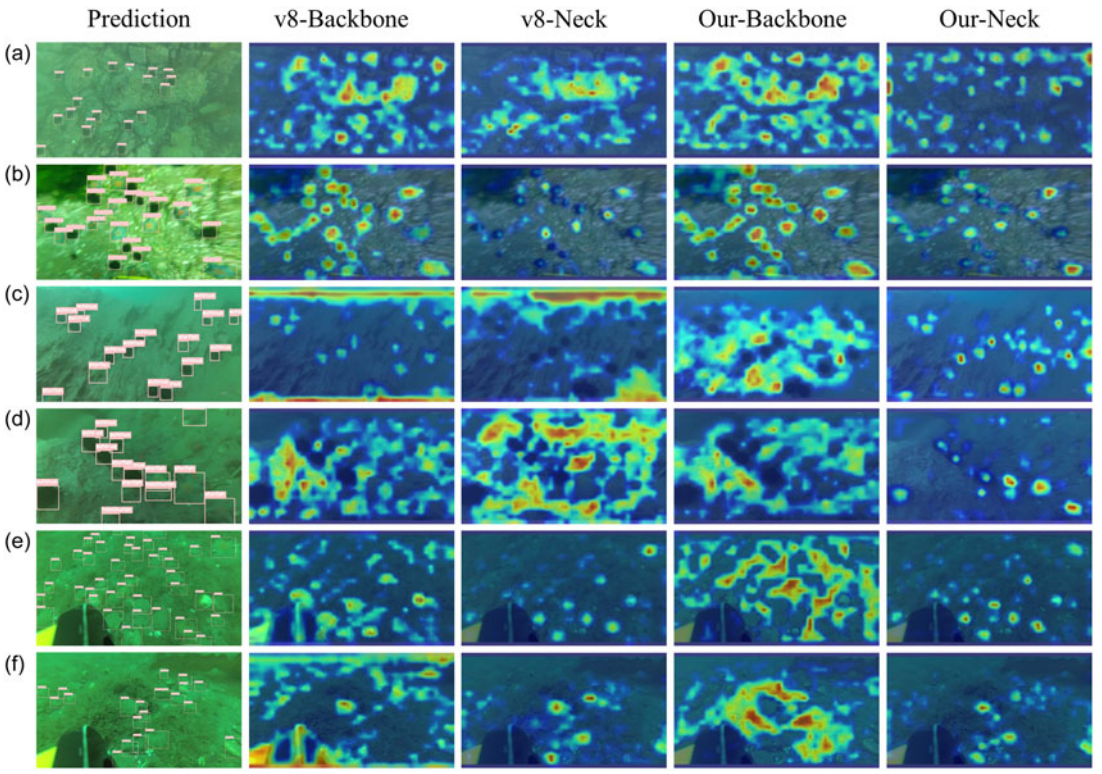


Figure 11. *Visualization of feature maps of REFNet and detectors in common scenarios of the URPC dataset.*

improvement confirms that our architecture successfully addresses the "feature dilution" problem in underwater imaging, where target features often get overwhelmed by scattering artifacts.

In summary, through systematic module design and thorough ablation studies, the REFNet algorithm strikes a well-balanced trade-off between detection accuracy and computational complexity. Our comparative analysis reveals that the performance gains are not merely incremental improvements from model scaling, but rather stem from targeted architectural innovations addressing specific underwater detection challenges. The visual results and comprehensive performance analysis confirm the superiority of REFNet over other leading models, particularly in challenging underwater conditions. Through the visualization of EigenCAM, the paper demonstrates how the proposed improvements enhance the network's performance and validate the design decisions behind REFNet.

While the proposed REFNet demonstrates superior performance in underwater blurred target detection, three key limitations merit discussion. First, the architecture's reliance on single-modality RGB data restricts its ability to leverage complementary information from sonar or LiDAR sensors, particularly in extreme turbidity conditions where visual features become unreliable. Second, the current framework processes individual frames independently, overlooking temporal consistency cues that could enhance detection stability in dynamic underwater scenes. Third, the supervised learning paradigm requires fully annotated datasets, which are scarce and labor-intensive to acquire for specialized underwater applications.

These limitations suggest promising avenues for methodological advancement. Multi-source information fusion strategies [44] could integrate acoustic and spectral data through cross-modal alignment techniques, potentially overcoming visibility constraints in turbid environments. Building on recent advances in multimodal interaction [45], a token division strategy might effectively correlate visual features with auxiliary metadata to resolve low-contrast ambiguities. Furthermore, graph neural networks

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

incorporating latent feature relationships [46] could model spatial dependencies between obscured targets and their environmental context, particularly beneficial for detecting partially occluded marine organisms. For data efficiency, self-paced semi-supervised learning frameworks like S2Match [47] could harness abundant unlabeled underwater imagery while maintaining detection reliability through curriculum-based sample weighting. Future implementations might combine these directions with physics-informed neural networks to explicitly model underwater light propagation characteristics, potentially enabling more robust performance across varying water quality conditions. Such enhancements would particularly benefit real-time applications like autonomous submersible navigation and large-scale marine ecological monitoring systems.

5. Conclusions

This paper presents a network tailored for detecting blurred underwater targets. It improves target detection by enhancing global response and cross-channel information fusion through reparameterization, effectively addressing the challenges of the underwater blur environment. First, this paper introduces the REG module, which enhances the backbone of the detection algorithm by improving its feature extraction ability, enabling it to capture higher-quality image features in challenging underwater conditions. Second, this paper proposes the CIF module, designed to strengthen the neck of the detection network and enhance its ability to perceive fine details in underwater fuzzy targets. Finally, This paper incorporates Shape-IoU as the network's loss function to enhance the precision of bounding box localization. This paper evaluates the REFNet algorithm by comparing it with the latest methods on the URPC and DUO datasets. Experimental results demonstrate that REFNet achieves the highest mAP while maintaining relatively lower computational efficiency. The REFNet achieves a $mAP_{50:95}$ improvement of 1.5% on the URPC dataset and 1.3% on the DUO dataset compared to YOLOv8. Therefore, the proposed REFNet algorithm excels in detecting underwater blurred targets. In future work, we aim to extend the validation of our proposed methods by testing them on a broader range of object detection algorithms, with the goal of enhancing their generalization capabilities.

Authors' contributions. Lei Cai conceived and designed the study. Benshun Li and Lei Cai conducted data gathering. Benshun Li wrote the article.

Funding. This work was supported by Henan Provincial focus on research and development Project (231111220700, 241111110200).

Competing interests. The authors declare that there are no conflicts of interest regarding the publication of this article.

Ethical considerations. None.

References

[1] A. Ismail, M. Mehri and A. Sahbani, "A dual-stage system for real-time license plate detection and recognition on mobile security robots," *Robotica* **1**(1), 1–22 (2025).
[2] S. Xu, M. Zhang, W. Song, H. Mei and Q. He, "A systematic review and analysis of deep learning-based underwater object detection," *Neurocomputing* **527**(12), 204–232 (2023).
[3] Y. Shi, D. Feng, Y. Cheng and S. Biswas, "A natural language-inspired multilabel video streaming source identification method based on deep neural networks," *Signal Image Video Process.* **15**(6), 1161–1168 (2021).
[4] T. Liu, J. Huang and J. Zhao, "Research on obstacle avoidance of underactuated autonomous underwater vehicle based on offline reinforcement learning," *Robotica* **1**(1), 1–25 (2024).
[5] J. Zhang, Y. Liu, S. Zhang, R. Poppe and M. Wang, "Light field saliency detection with deep convolutional networks," *IEEE Trans. Image Process.* **29**(18), 4421–4434 (2020).
[6] H. Liu, X. Sun, J. Gu and L. Deng, "SF-YOLOv5: A lightweight small object detection algorithm based on improved feature fusion mode," *Sensors* **22**(15), 5817–5836 (2022).
[7] S. Chang, F. Gao and Q. Zhang, "A systematic review and analysis of deep learning-based underwater object detection," *Electronics* **12**(13), 2882–2896 (2023).

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

[8] W. Zhang, X. Li, Y. Huang, S. Xu, J. Tang and H. Hu, "Underwater image enhancement via frequency and spatial domains fusion," *Opt. Lasers Eng.* **186**(9), 1008826–1008842 (2025).

[9] X. Jiang, X. Zhuang, J. Chen, J. Zhang and Y. Zhang, "YOLOv8-MU: An improved YOLOv8 underwater detector based on a large kernel block and a multi-branch reparameterization module," *Sensors* **24**(9), 2905–2919 (2024).

[10] Z. Xu, R. Wang and T. Cao, "AquaPile-YOLO: Pioneering underwater pile foundation detection with forward-looking sonar image processing," *Remote Sens.* **17**(3), 360–376 (2025).

[11] L. Cai, B. Zhang, Y. Li and H. Chai, "IFE-net: Improved feature enhancement network for weak feature target recognition in autonomous underwater vehicles," *Robotica* **42**(4), 1231–1245 (2024).

[12] L. Zhou, L. Cai, J. Jia and Y. Gao, "Multi-scale aware turbulence network for underwater object recognition," *Front. Mar. Sci.* **11**(25), 1301072–1301091 (2024).

[13] R. Varghese and M. Sambath. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. **In:** *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India* (2024) pp. 1–6.

[14] R. Qu, C. Cui, J. Duan, Y. Lu and Z. Pang, "Underwater small target detection under YOLOv8-LA model," *Sci. Rep.* **14**(1), 16108–16124 (2024).

[15] Z. Lu, L. Liao, X. Xie and H. Yuan, "SCoralDet: Efficient real-time underwater soft coral detection with YOLO," *Ecol. Inform.* **85**(7), 102937–102956 (2025).

[16] Z. Niu, G. Zhong and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing* **452**(22), 48–62 (2021).

[17] Y. Fei, F. Liu and M. Su, "Real-time detection of small underwater organisms with a novel lightweight SFESI-YOLOv8n model," *J. Real-Time Image Process.* **22**(7), 23–43 (2025).

[18] J. Quan, Z. Zhao, W. Li, Y. Cao and J. Wu, "Enhancing YOLOv8n with multiple attention and MRV module for efficient deep-sea pipeline target detection," *Electronics* **14**(2), 267–288 (2025).

[19] Z. Zhou, Y. Hu, X. Yang and J. Yang, "YOLO-based marine organism detection using two-terminal attention mechanism and difficult-sample resampling," *Appl. Soft. Comput.* **153**(16), 111291–111305 (2024).

[20] J. Max, S. Karen, Z. Andrew and K. Koray, "Spatial transformer networks," *Adv. Neur. Inf. Process. Syst.* **28**(22), 1291–1305 (2015).

[21] W. Li, X. Zhang, Y. Peng and M. Dong. Squeeze-and-Excitation Networks. **In:** *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA* (2018) pp. 7132–7141.

[22] J. Fu, H. T. J.Liu, Y. Li, Y. Bao, Z. Fang and H. Lu. Dual Attention Network for Scene Segmentation. **In:** *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA* (2019) pp. 3141–3149.

[23] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo and Q. Hu. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. **In:** *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA* (2020) pp. 11531–11539.

[24] Z. Li, Y. Dong, L. Shen and YLiu, "Development and challenges of object detection: A survey," *Neurocomputing* **598**(22), 128102–128125 (2024).

[25] Z. Xin, S. Chen, T. Wu and Y. Shao, "Few-shot object detection: Research advances and challenges," *Inf. Fusion* **107**(42), 102307–102332 (2024).

[26] T.-Y. Lin, P. Dollár, R. Girshick, B. Hariharan and S. Belongie. Feature Pyramid Networks for Object Detection. **In:** *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA* (2017) pp. 936–944.

[27] S. Liu, L. Qi, H. Qin, J. Shi and J. Jia. Path Aggregation Network for Instance Segmentation. **In:** *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA* (2018) pp. 8759–8768.

[28] G. Ghiasi, T.-Y Lin and Q. V. Le. NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection. **In:** *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA* (2019) pp. 7029–7038.

[29] L. Chen, Y. Fu, L. Gu, C. Yan, T. Harada and G. Huang, "Frequency-aware feature fusion for dense image prediction," *IEEE Trans. Pattern Anal. Mach. Intell.* **46**(12), 10763–10780 (2024).

[30] C. Sun and F. Zhao, "Multi-level feature fusion network for neuronal morphology classification," *Front. Neurosci.* **18**(16), 1465642–1465663 (2024).

[31] T. Zhang, Y. Liu and Q. Zhao, "Edge-guided multi-scale adaptive feature fusion network for liver tumor segmentation," *Sci. Rep.* **14**(1), 28370–28392 (2024).

[32] X. Ding, Y. Zhang, Y. Ge, S. Zhao, L. Song, X. Yue and Y. Shan. UniRepLKNet: A Universal Perception Large-Kernel ConvNet for Audio Video Point Cloud Time-Series and Image Recognition. **In:** *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)* (2024), pp. 5513–5524.

[33] S. Woo, S. Debnath, R. Hu, X. Chen, Z. Liu, I. S. Kweon and S. Xie. ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders. **In:** *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada* (2023) pp. 16133–16142.

[34] D. Shao, J. Jiang and L. Ma, "Real-time medical lesion screening: Accurate and rapid detectors," *J. Real-Time Image Process.* **21**(4), 134–158 (2024).

[35] S. Ren, K. He, R. Girshick and J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. **In:** *IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA* (2014) pp. 580–587.

[36] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang and Q. Tian. CenterNet: Keypoint Triplets for Object Detection. **In:** *2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South)* (2019) pp. 6568–6577.

[37] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov and S. Zagoruyko, "End-to-end object detection with transformers," *Eur. Conf. Comput. Vision (ECCV), Glasgow, UK* **12346**, 213–229 (2020).

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press

[38] E. Khachatrian, N. Sandalyuk and P. Lozou, "Eddy detection in the marginal ice zone with sentinel-1 data using YOLOv5," *Remote Sens.* **15**(9), 2244–2259 (2023).
[39] Z. Li, P. Xu, X. Chang, L. Yang, Y. Zhang, L. Yao and X. Chen, "When object detection meets knowledge distillation: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(8), 10555–10579 (2023).
[40] C.-Y. Wang, A. Bochkovskiy and H.-Y.M. Liao. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. **In:** *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada* (2023) pp. 7464–7475.
[41] C. H. Kang and S. Y. Kim, "Real-time object detection and segmentation technology: An analysis of the YOLO algorithm," *JMST Adv.* **5**(2), 69–76 (2023).
[42] X. Qiu, Y. Chen, W. Cai, M. Niu and J. Li, "LD-YOLOv10: A lightweight target detection algorithm for drone scenarios based on YOLOv10," *Electronics* **13**(16), 3269–3287 (2024).
[43] M. B. Muhammad and MYeasin, "Eigen-CAM: Visual explanations for deep convolutional neural networks," *SN Comput. Sci.* **2**(6), 47–59 (2021).
[44] L. Xinde, F. Dunkin and J. Dezert, "Multi-source information fusion: Progress and future," *Chin. J. Aeronaut.* **37**(7), 24–58 (2024).
[45] Y. Cai, X. Sui and G. Gu, "Multi-modal interaction with token division strategy for RGB-T tracking," *Pattern Recognit.* **115**(6), 110626–110648 (2024).
[46] X. Shi, Y. Zhang and A. Pujahari, "When latent features meet side information: A preference relation based graph neural network for collaborative filtering,"Expert," *Syst. Appl.* **260**(11), 125423–125451 (2025).
[47] D. Guan, Y. Xing and J. Huang, "S2Match: Self-paced sampling for data-limited semi-supervised learning," *Pattern Recognit.* **159**(26), 111121–111143 (2025).

Cite this article: B. Li and L. Cai (2025). "REFNet: reparameterized feature enhancement and fusion network for underwater blur target recognition", Robotica **43**, 1867–1884. https://doi.org/10.1017/S0263574725000475

https://doi.org/10.1017/S0263574725000475 Published online by Cambridge University Press