

Research Article

Open Access



Towards intelligent shipping: image-enhanced ship detection and situation analysis in low-light scenes

Xinqiang Chen^{1,2}, Rui Yang¹, Yuzhen Wu^{1,3}, Han Zhang⁴, Prakash Ranjitkar⁵, Octavian Postolache^{6,7}, Yiwen Zheng⁴, Zichuang Wang⁸

¹Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China.

²Chongqing Key Laboratory of Green Logistics Intelligent Technology, Chongqing Jiaotong University, Chongqing 400074, China.

³SPG Qingdao Port Group Co., Ltd., Qingdao 266011, Shandong, China.

⁴College of Ocean Science and Engineering, Shanghai Maritime University, Shanghai 201306, China.

⁵Department of Civil and Environmental Engineering, The University of Auckland, Auckland 1010, New Zealand.

⁶Instituto de Telecomunicações (IT), Lisbon 1049-001, Portugal.

⁷Department of Information Science and Technology, Iscte - Instituto Universitário de Lisboa, Lisbon 1649-026, Portugal.

⁸Shanghai Ship and Shipping Research Institute Co., Ltd. Shanghai 201306, China.

Correspondence to: Yuzhen Wu, Institute of Logistics Science and Engineering, Shanghai Maritime University, 1550 Haigang Av., Shanghai 201306, China. E-mail: wuyz@qdport.com

How to cite this article: Chen, X.; Yang, R.; Wu, Y.; Zhang, H.; Ranjitkar, P.; Postolache, O.; Zheng, Y.; Wang, Z. Towards intelligent shipping: image-enhanced ship detection and situation analysis in low-light scenes. *Intell. Robot.* **2025**, *5*(3), 662-78. <https://dx.doi.org/10.20517/ir.2025.34>

Received: 17 May 2025 **First Decision:** 1 Jul 2025 **Revised:** 23 Jul 2025 **Accepted:** 24 Jul 2025 **Published:** 7 Aug 2025

Academic Editor: Xin Jin **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

Aiming to address the problems of insufficient ship detection accuracy and high miss rate for small targets in water transport traffic situational awareness under low-light conditions, this paper proposes an EG-YOLO+ framework that integrates ship image enhancement and ship detection. The method achieves adaptive enhancement of low-light images through the unsupervised enhancement-low-light-image-network generative adversarial network (EnlightenGAN) model, effectively solving the problem of detail loss in traditional methods under extreme lighting conditions; subsequently, based on the latest You Only Look Once version 11 (YOLOv11) architecture, it innovatively introduces the squeeze-and-excitation channel attention mechanism, significantly improving the detection accuracy of small-target ships through dynamic feature channel reweighting. The experimental results on the self-constructed maritime dataset show that the proposed method can effectively identify image targets in low-light environments, even small targets, with a 6-percentage point improvement in mAP over the baseline YOLOv11.

Keywords: Ship image enhancement, ship detection, YOLOv11, EnlightenGAN, ship trajectory analysis



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



1. INTRODUCTION

Water transport traffic is an important part of the national economy and social life, and promoting intelligent and information-based development in the field of water transport has become a general trend. To address the shortcomings of the current water transport traffic system under low-light conditions, this paper combines ship image enhancement technology, ship detection, and situational awareness technology into the water transport traffic situational awareness system. Enhancement-low-light-image-network generative adversarial network (EnlightenGAN) and You Only Look Once version 11 (YOLOv11) are used to enhance low-light images and classify ship positions for situational analysis. Empirical results demonstrate the effectiveness of this approach in promoting intelligent water transport.

Ship image enhancement technology has made significant progress in recent years under the impetus of deep learning, with a wide range of applications such as medical imaging, intelligent monitoring, agricultural detection, automatic driving and other fields. Scholars at home and abroad have carried out in-depth research around traditional enhancement methods, deep learning enhancement techniques, and specific scenes, and scholars have focused on the optimization of algorithms based on signal processing. Sun *et al.* proposed an ship image enhancement method based on multi-weight fusion for the special needs of the intelligent monitoring scene of a mine, the method innovatively combines the Gamma algorithm with hue, saturation, and value (HSV) spatial transformation, dynamically adjusts the light distribution through the improved two-dimensional Gamma function, and ultimately adopts Gauss-Laplace pyramid to achieve image fusion^[1]. Shorten and Khoshgoftaar systematically sorted out the technical system of image data enhancement, and classified the existing methods into the traditional techniques, such as geometric transformation and color space enhancement, and the emerging methods, such as adversarial training, generative adversarial networks (GANs), and other emerging methods^[2]. Xu *et al.* further proposed an information-rich three-dimensional taxonomy, subdividing ship image enhancement algorithms into three categories: model-free, model-based, and optimization strategy-based, which provided a theoretical framework for technology selection^[3]. Nagaraju *et al.* provided a theoretical framework for the identification of crop diseases in a crop disease identification task by designing systematic image acquisition algorithms and enhancement processes. A model performance improvement of 5.14% was achieved, and the enhancement was significantly better than six existing benchmark methods^[4]. Hyper-Laplacian Reflection Priors proposed by Zhuang *et al.* significantly improve the detail recovery ability of underwater images by modeling the sparsity of reflection components^[5].

The application of GANs in ship image enhancement has become a research hotspot in recent years. A review study by Eldeen *et al.* points out that GANs show unique advantages in solving the problem of scarcity of medical image data^[6]. In special scenarios, GANs-based data enhancement can effectively expand the training samples and improve the generalization ability of the model. The study by Gupta *et al.* in the field of agriculture further confirms that GAN-generated real samples can improve the accuracy of crop recognition by 15%-20%, opening up a new way to solve the problem of insufficient samples in agricultural images^[7].

The You Only Look Once (YOLO) series of algorithms in the ship recognition part has been the focus of attention. Wang *et al.* proposed an improved YOLO version 5 (YOLOv5) algorithm for the special environment of coal mine underground, which improved the ship recognition accuracy by 9.1% and the mean average precision (mAP) by 3.6% by introducing the efficient channel attention (ECA)_s channel attention mechanism and the efficient intersection over union (EIOU) loss function^[8]. The YOLO network for Sweeping Robots version 2 (YOLO_SRv2) algorithm developed by Lv *et al.* improves ship recognition accuracy in the backbone network by 15%-20%. This algorithm integrates a maximum pooling structure and

a convolutional block attention module (CBAM) into the backbone, achieving 86.1% mAP on the object detection for sweeping robots in home scenarios (ODSR-HIS) dataset while maintaining a high frame rate of 180 frames per second (FPS)^[9]. Dai *et al.* proposed the YOLO-Former model, which fuses the Vision Transformer with a convolutional attention mechanism, achieving an 11.3% improvement in mAP for foreign ship detection and establishing a new paradigm for integrating Transformer and YOLO architectures^[10].

Zhao and Ding proposed an asynchronous deep reinforcement learning framework, which effectively improves the ship detection rate under low-light conditions through multimodal image edge contour detection and adaptive light intensity fusion^[11]. Cao *et al.* developed the YOLO-SF algorithm, which innovatively combines instance segmentation with the Mobile Vision Transformer version 2 (MobileViTv2) backbone and the Varifocal Loss function to achieve a balance between accuracy and speed in fire detection tasks^[12]. Gomaa *et al.* developed a vehicle counting system based on YOLOv2 and the Kanade–Lucas–Tomasi (KLT) tracker, improving the processing speed to 18.7 FPS, which is 93.4% higher than that of traditional methods^[13].

Both ship image enhancement and ship detection have seen significant progress; however, existing research on ship detection in low-light environments remains insufficient, with clear gaps both domestically and internationally. Ship detection research often directly deals with the original low-quality images or uses simple off-the-shelf enhancement methods for preprocessing, lacking synergy between the two^[14]. For example, the multimodal recognition study of Zhao and Ding involves low-light conditions, but its enhancement module and recognition network are independent of each other^[11]. The mine monitoring study of Sun *et al.* improves Gamma correction, but is not trained jointly with the ship detection model^[1]. This study builds on this foundation.

Although existing studies have made progress in low-light image enhancement and ship detection, two critical gaps remain: (1) Most methods process enhancement and detection independently, leading to suboptimal synergy; (2) Small-target detection under extreme lighting suffers from feature flooding due to redundant channel responses. To address these issues, EG-YOLO+ proposes a cascaded framework that integrates EnlightenGAN's unsupervised enhancement with YOLOv11's dynamic channel reweighting via squeeze-and-excitation (SE) modules. This design ensures joint optimization of illumination correction and feature prioritization, particularly for small targets.

2. METHODS

The new model EG-YOLO+ proposed in this paper is divided into two parts: ship image enhancement and ship detection. Ship image enhancement is based on EnlightenGAN model with unsupervised GAN using generator-discriminator interaction to achieve the desired effect^[15]. The YOLOv11 algorithm from the YOLO series is chosen as the backbone for ship detection, and the SE module is added to improve the performance of small ship detection by learning the weights of each channel through the global average pooling and the fully-connected layer, and automatically reinforcing the channels that are critical for the corresponding detection tasks^[16].

2.1. EnlightenGAN

GAN is an unsupervised learning model, which mainly consists of a generator and a discriminator. Adopting EnlightenGAN not only solves the problem of low light enhancement, but also proposes a novel unsupervised framework that can be trained without paired data, using a U-shaped network (U-Net) as the backbone network of the generator, and fusing low-level features with high-level features through pixel-

level attention and multi-scale feature extraction operations in order to obtain more feature information. Ship image enhancement is always desired to enhance the dark areas so that the global image brightness is balanced and not overexposed or too dark^[17].

The discriminator integrates the sub-feature retention loss of the local and global discriminators and the adversarial loss of the global and local generators, randomly crops 5 paths from the two images before and after enhancement, and uses the original least squares GAN as the adversarial loss. The structure of the discriminator is shown in [Figure 1](#).

2.2. Ship detection

In this paper, YOLO is used as the backbone network for ship detection, which can treat ship detection as a single regression problem, and is able to predict the coordinates of target species and bounding boxes in real time^[18]. Compared with the previous generations of YOLO series, the new version of YOLOv11 has higher computational efficiency while maintaining higher accuracy. The YOLOv11 model consists of four parts: Input, Backbone, Neck, and Head^[19]. Input performs standardized preprocessing and data enhancement on the image to provide the network with inputs in a uniform format. The Backbone provides the network with inputs in a uniform format through the Backbone extracts multi-level features through the improved Cross Stage Partial Darknet53 (CSPDarknet53) structure, forming feature pyramids at different scales^[20]; Neck fuses and optimizes the multi-scale features to enhance the detail information of small targets and the semantic information transfer of large targets; Head directly predicts the targets and outputs the detection results^[21].

To address the problem that detection accuracy for medium and large targets is high but remains low for small targets^[22], YOLOv11 optimizes small ship detection performance by introducing a high-resolution detection layer. This enhancement significantly improves the recognition of small target objects, making YOLOv11 a suitable choice for the backbone model.

The convolution operation of traditional YOLOv11 treats all feature channels equally; although it has been more accurate for small ship detection, the channel feature redundancy in complex scenes significantly reduces the performance of small ship detection. The introduction of the SE module in YOLOv11 can effectively improve the small ship detection accuracy, while maintaining the detection performance for medium and large targets. The SE module solves the problem of small target feature flooding under low illumination through dynamic channel reweighting. Squeeze stage global average pooling generates channel statistics, compresses spatial dimensions. Excitation stage two-layer full connection layer learns nonlinear relationship between channels, suppresses redundant background channels such as water surface reflection, and improves response intensity of high-frequency detail channels of small targets. The network structure of the SE module is shown in [Figure 2](#). The channel reweighting mechanism enables the network to automatically differentiate the feature requirements of targets of different sizes, enhancing high-frequency detail channels for small targets while retaining semantic-dominant channels for large targets, thus significantly improving small ship detection performance without compromising the accuracy for medium and large targets.

The SE module is added at the cross-scale connection of the feature pyramid network (FPN) layer to optimize the channel weight allocation of feature maps with different resolutions, and to enhance the visibility of small targets in the deep features without destroying the integrity of the backbone features. The workflow is divided into Squeeze, Excitation, and Reweighting, based on a given input feature map, as given in

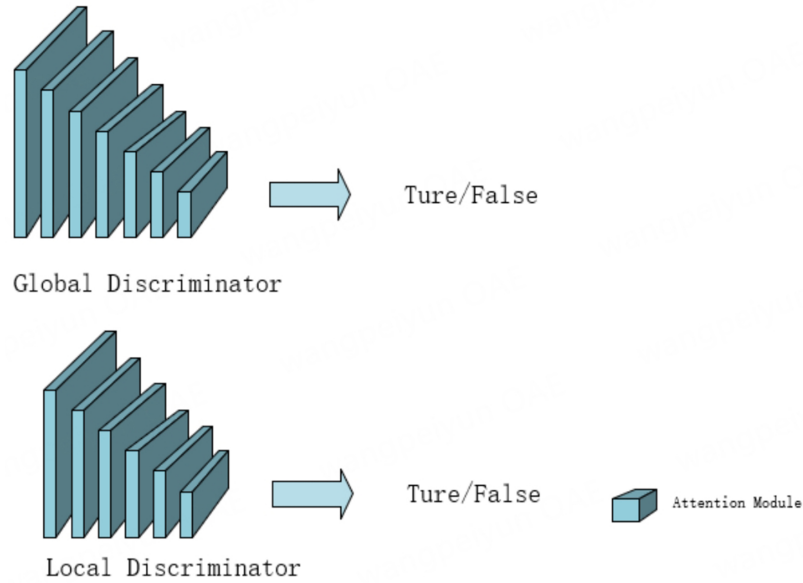


Figure 1. Structure of GAN network discriminator. GAN: Generative adversarial network.

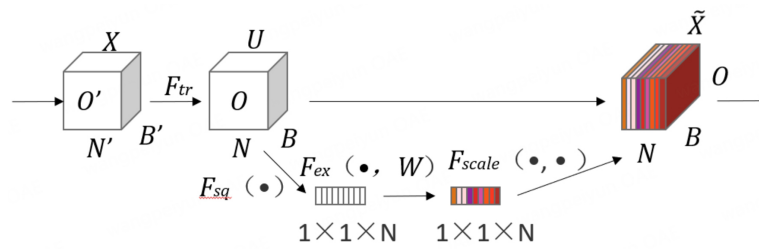


Figure 2. SE module structure. SE: Squeeze-and-excitation.

$$X \in R^{O \times B \times N} \quad (1)$$

The SE module first generates channel statistics, as given in

$$z \in R^N \quad (2)$$

through global average pooling, and then learns activation weights through two FC layers, and finally outputs a reweighted feature, as determined by^[23].

$$\tilde{X} = s \cdot X \quad (3)$$

This mechanism allows the network to automatically enhance high-frequency detail weights when detecting a small target. targets by automatically enhancing the response strength of high-frequency detail channels, such as edges, corner points, and other places.

2.3. Overall framework of EG-YOLO+

This study employs a cascaded design with staged processing to reduce system coupling, consisting of three key components: the front-end, back-end, and joint optimization module. The front-end utilizes EnlightenGAN to perform illumination correction on input images and outputs enhanced results, while the back-end incorporates an improved YOLO network for object detection. The detection network implements multi-scale feature extraction and dynamically adjusts channel-wise feature weights in the FPN layers through SE modules. The training process follows a two-stage strategy: the enhancement module is first trained independently, followed by end-to-end fine-tuning of the entire detection network. The workflow is illustrated in Figure 3. The design innovatively proposes a two-stage framework called EG-YOLO+, which revolves around an “enhancement-detection” collaborative framework. It is the first to deeply integrate unsupervised image enhancement with an improved YOLOv11 object detection model to solve the problem of enhancement and detection tasks in low-light scenarios.

The pseudocode for EG-YOLO+ is given in Algorithm 1. It summarizes EG-YOLO+’s two-stage workflow. Stage 1 enhances illumination via EnlightenGAN without paired training data. Stage 2 applies SE-based channel reweighting to suppress redundant features (e.g., water reflections) before detection. \odot denotes element-wise multiplication.

3. EXPERIMENTS

In this section, the experimental details will be introduced. We will demonstrate the accuracy of the EG-YOLO+ in analyzing the ship’s operational situation under low-light conditions through extensive experiments. This study adopts the recommended configuration of the YOLO framework, utilizing the AdamW optimizer with an initial learning rate (lr0) of 0.01, along with learning rate scheduling and data augmentation strategies. For the ship detection task, we adjusted the input resolution to 640×640 to balance small target detection accuracy and computational efficiency, and set the batch size to 8 to accommodate GPU memory constraints. The theoretical computational complexity of the model is approximately 10.8 giga floating point operations (GFLOPs), with the YOLOv11-SE component contributing about 7.7 GFLOPs and the image enhancement module accounting for approximately 3.1 GFLOPs. The total parameter count is 7.8 M.

3.1. Datasets

The current publicly available maritime ship detection dataset generally has some limitations; maritime technology is developing rapidly but the old datasets are difficult to reflect the latest ship characteristics. In order to break through the bottleneck of existing data, this study independently constructs a maritime target detection dataset, which covers more than 1,800 images in daytime, night, foggy day and strong reflection water surface. The image scene is changeable. This method enhances the diversity of the dataset. Some of the pictures are shown in Figure 4; this method enhances the diversity of the dataset. Some of the images have problems such as blurring and color deviation, which can effectively verify the accuracy of the experiment and greatly enhance its persuasiveness.

3.2. Experimental condition

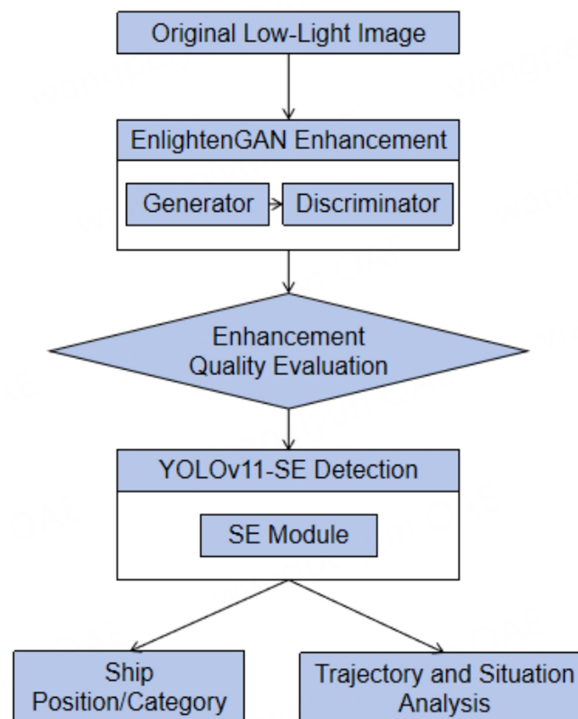
The hardware and software required for the experiment are shown in Table 1.

3.3. Ship image processing

Firstly, we preprocess the dataset by deleting excessively blurred images or those without visible ships, leaving 1,698 images, which are distinguished into training set, validation set, and test set according to the ratio of 6:2:2, and the images are experimented with. The ship images are labelled using LabelMe, and each frame of the images in the training set corresponds to a manually labelled file in JSON format, which

Table 1. Parameters of hardware and software equipment

Device	Parameter
Memory	64GB
CPU	Intel Xeon Gold 6230R @ 2.10GHz
Graphics card	NVIDIA Quadro RTX5000,16GB
Operating system	Linux (x64)
Programming language	Python 3.8.20
Learning framework	Pytorch 2.4.1+CUDA 12.1
cuDNN Version	9.1.0

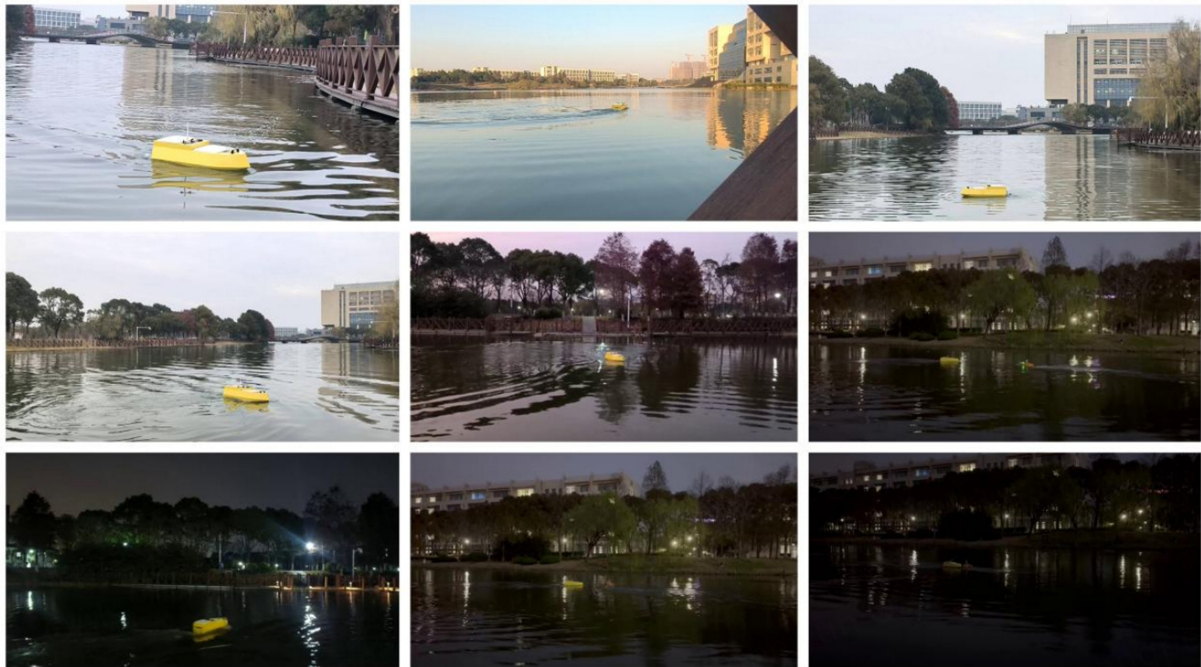
**Figure 3.** Flowchart.

contains information such as target ID, target category, and four-dimensional coordinates of the frame.

This study addresses ship images in low-light situations, so ship image enhancement of low-light images is required, using an unsupervised GAN model, consisting of a generator and a discriminator. Dark light enhancement can be achieved by training the GAN model without pairing images using the proposed global-local discriminator.

Algorithm 1 EG-YOLO+**Input:** Low-light image I_{low} **Output:** Detection results B , Confidence scores C **Stage 1: Image Enhancement**1: $I_{\text{enhanced}} = \text{EnlightenGAN}(I_{\text{low}})$ **Stage 2: Ship Detection**2: $F = \text{CSPDarknet53}(I_{\text{enhanced}})$ 3: for each feature map f in $\text{FPN}(F)$:4: $z = \text{GlobalAvgPool}(f)$ 5: $s = \sigma(W2 \cdot \text{ReLU}(W1 \cdot z))$ 6: $f' = s \odot f$ 7: $B, C = \text{YOLOv11_Head}(f')$ **return** B, C

Algorithm 1.

**Figure 4.** Some images of boat 1 and boat 2 taken under different water conditions.**3.4. Ship detection**

Ship detection has always been one of the important research directions in the field of computer vision, and its importance in the domain of shipping is also self-evident. In this study, YOLOv11 is used as the base model for the ship detection part; as the latest model of YOLO, it is famous for its excellent detection speed and end-to-end training process, and further improves the detection accuracy and robustness by optimizing

the network architecture and training algorithm^[24].

Recognition speed is an important indicator in ship inspection tasks, which usually require ship detection in a short period of time, and the efficiency of the YOLOv11 makes it particularly suitable for this task; however, it was found during the experiments that its recognition accuracy is slightly worse under low light conditions, especially when the hull of the ship is partially covered or the environment is very dark^[25]. The YOLOv11 may have difficulty in effectively detecting small target sizes, resulting in missed detections or false alarms. To address this problem, the study focuses on low-light images to enhance the image quality for more accurate recognition. Adding the SE module pays better attention to the features of small targets, excels in small ship detection accuracy, improves detection capability, and is able to adaptively recalibrate the channel feature response by learning the channel attention weights; comparison pictures are shown in Figures 5 and 6.

In the low-light scene as in Figure 5, the ship detection model exhibits a missed detection, and boat 2 is not detected. However, after optimization by ship image enhancement, boat 2 is successfully identified and the detection confidence of both boats is significantly improved.

Ship detection will also have obvious misdetection situations, misidentifying the brighter lights on the shore as ships, and the confidence value of misidentification reaches 0.52, which is a major recognition error. After the image quality optimization process, this kind of misidentification problem is effectively solved, while the recognition accuracy of the original ship target is also improved.

The experiments use the following indicators to evaluate the effectiveness of EG-YOLO+ ship detection, respectively.

(1) Precision: Precision is used to reflect the model detection accuracy, as given in

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

where TP refers to the number of targets that the model strives to detect, and FP refers to the number that the model misdetects; i.e., actually does not exist but is incorrectly detected as having a target.

(2) Recall: Recall reflects the model's ability to avoid missed detections, as given in

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

where FN is the number of targets missed by the model, those that are real but not detected.

(3) mAP50 (mean average precision at IoU = 50%) and mAP50-95

mAP50 and mAP50-95 are two core evaluation metrics in the ship detection task, which are used to measure the combined detection performance of the model at different levels of stringency. mAP50 is the mean average precision computed at the intersection and concurrency ratio threshold of 0.5, which is then averaged across all the categories of APs, and focuses on whether or not the target is detected, whereas the mAP50-95 has a threshold of 0.5-0.95 and focuses on whether the detection location was accurate, as calculated by



Figure 5. Example of missed ship detection in a low-light scene.



Figure 6. Comparison of ship misdetection.

$$mAP = \frac{1}{n} \sum_{i=1}^N AP_i = \frac{1}{n} \sum_{i=1}^N \int_0^1 P(R) dR \quad (6)$$

where AP is the average accuracy and N is the number of target detection classes.

This study comprehensively evaluates the performance of a ship detection model trained using augmented images under realistic low-light water conditions. In order to validate the effectiveness of the improved model, mainstream ship detection algorithms are selected for evaluation and comparative validation. We perform ship detection on both original and enhanced images, and the comparison results are shown in [Table 2](#).

This model achieves a precision of 0.852, which is significantly higher than YOLOv11 (0.821) and SSD (0.755), with a lower false detection rate. The recall of 0.804 outperforms all the compared models, with a lower miss detection rate for real targets. mAP50 outperforms YOLOv11's 78.4% with 80.1%, and the detection is more accurate especially at high Intersection over Union (IoU) thresholds. The variant YOLOv11-self-calibrated attention (SCA) with the addition of the attention mechanism is lower than the original YOLOv11 in terms of precision and mAP, while the model proposed in this study successfully overcame this performance degradation, indicating that the optimization strategy is indeed effective.

In practical applications, the variation of light will affect the performance of the ship detection model, and the robustness of the model can be evaluated by calculating the change rate of mAP of the model under

Table 2. Comparison of metrics of different algorithms

Model	Precision	Recall	mAP50 (%)	mAP50-95 (%)
SSD	0.755	0.709	72.9	45.8
YOLOv11	0.821	0.785	78.4	55.7
YOLOv11-SCA	0.789	0.727	75.7	48.6
Ours	0.852	0.804	80.1	58.3

The bold formatting highlights superior performance metrics. YOLOv11: You Only Look Once version 11; SCA: self-calibrated attention.

Table 3. Comparison of the parameters of the algorithms under different conditions of illumination

Mould	Normal-light mAP (%)	Low-light mAP (%)	Rate of change in mAP (%)
SSD	72.4	55.9	23.6
YOLOv11	78.2	62.3	20.5
YOLOv11-SCA	75.9	58.7	22.7
Ours	82.6	68.3	17.1

The bold formatting highlights superior performance metrics. mAP: Mean average precision; YOLOv11: You Only Look Once version 11.

different light conditions. The results are shown in [Table 3](#).

YOLOv11-SE has the smallest change in mAP under different light conditions, which is only 17.1%, indicating that it is the most adaptable to light changes and has the best robustness, and the present model has superior mAP values under both normal and low light conditions. Compared with the original YOLOv11 model, the mAP of this model is improved by 4.4% under normal light and even more by 6% under low light conditions, which indicates the superiority of this model.

[Figure 7](#) demonstrates the comparison of the effectiveness of different methods in the low-light image ship detection task. From the figure, it can be seen that compared with other ship detection methods without ship image enhancement, the effect of this study is remarkable. After ship image enhancement, the target can be more effectively extracted from the image. Even small targets can be recognized with high confidence, with no omissions or false detections, resulting in superior recognition performance. The confidence scores for the two ships (0.83 and 0.91) are higher than those of the original model (0.80 and 0.87), and the SCA module is not as good as the original model on the basis of YOLOv11, and the SSD effect is in the middle [[Figure 7A–D](#)]. This indicates that the performance gain of EG-YOLO+ is robust across different scene types and not only dependent on increased brightness.

3.5. Vessel posture analysis

Firstly, we preprocess the unlabeled night ship images, including low light enhancement, ship detection and then we depict the ship trajectory by frames for situational analysis. In this section, single-vessel and two-vessel scenarios are designed.

The change in X-coordinate over time for a single ship trajectory is shown in [Figure 8](#).

The horizontal coordinate represents time in frames, which can be interpreted as the change of ship trajectory over time, ranging from 80 to 560, and the vertical coordinate represents the position of the ship's X-coordinate at a certain moment in time, ranging from 0 to 850. The overall shape of the curve resembles that of a sinusoidal wave, which indicates that there is a “round-trip” motion of the ship in the X-direction.



Figure 7. Visualization comparison results. (A) SSD; (B) YOLOv11; (C) YOLOv11-SCA; (D) The proposed model in this work. YOLOv11: You Only Look Once version 11.

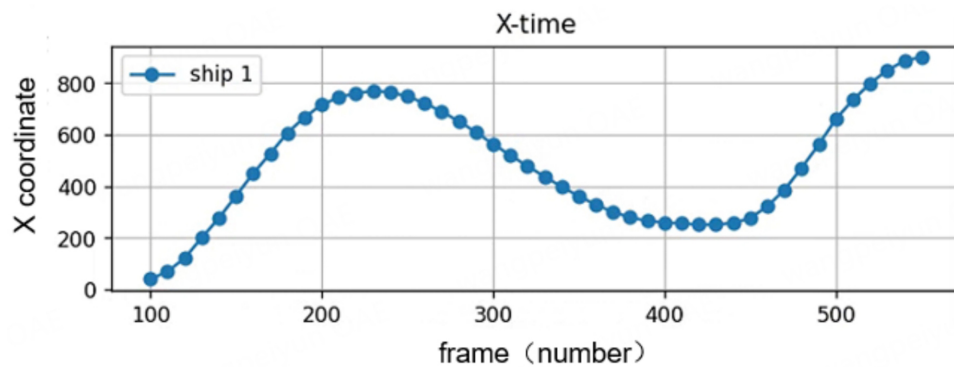


Figure 8. Relationship between the x-axis motion of the single vessel and time.

The overall shape of the curve is similar to a sine wave, indicating that the ship has a “round-trip” motion in the X direction. The analysis of key time periods is shown in [Table 4](#).

In [Figure 9](#), the vertical coordinate indicates the position of the ship’s Y coordinate in each frame, with a range of 530-670, and the horizontal coordinate is the same as that of the X chart, with a range of 80-560, also with periodic motion, indicating that the ship is not only moving back and forth in the X direction, but also doing a “back and forth” motion in the Y direction.

The trend characteristics are shown in [Table 5](#).

Table 4. Motion analysis table for single vessel in x-axis direction

Time interval (frames)	X coordinate change	Motion description
80-220	Upward trend	Ship accelerates to the right
220-330	Downward trend	The ship is gradually slowing down and turning left
330-420	Approaching stability	Ship's X-coordinate is generally stable, may be slowing down or in steering mode
420-560	Upward trend	Ship moves to the right again (acceleration)

Table 5. Motion analysis for single vessel in y-axis direction

Time interval (frames)	Y coordinate change	Motion description
80-300	Downward trend	Ship moving south (Y coordinate decreasing)
300-500	Upward trend	Ship moving north (Y coordinate increasing)
500-560	Slightly down	Vessel trimming to the south or slowing down to turn around

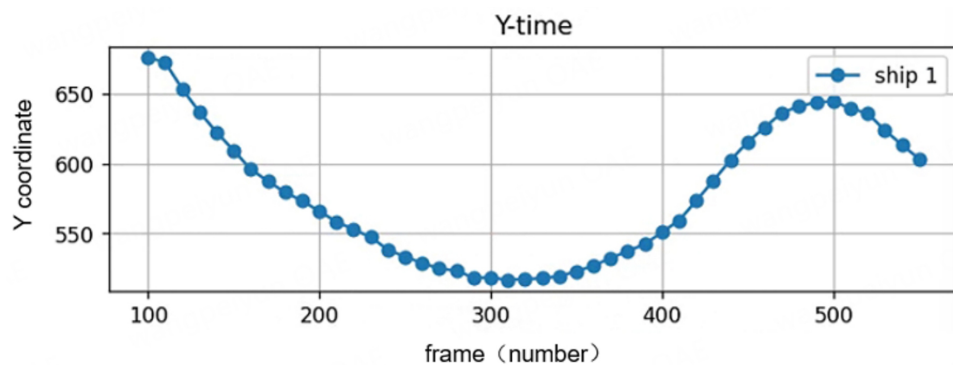
**Figure 9.** Relationship between the y-axis motion of the single vessel and time.

Figure 10 shows two ships travelling in opposite directions, with the X values approaching or even overlapping near frame 180-200 indicating that the two had a head-on rendezvous or brush at this point in time, with a great potential risk of collision. However, after that moment, they both turned back, and this behavior is mostly “avoidance” to prevent collision.

Ship 1’s initial position is at about 1,200 on the right, and as time passes, the ship’s X value gradually decreases, indicating that ship 1 is moving from right to left, and at about frame 180-200, the X value reaches a minimum, and then begins to rise, implying that the ship is experiencing a turn-around or turn-back behavior, driving backward for some distance.

Ship 2 is initially positioned at an X value of about 200 on the left, which gradually increases over time, indicating movement from left to right toward Ship 1. Around frames 180-200, the X value reaches a maximum of approximately 600, after which it begins to decrease, suggesting that a turn-back maneuver is occurring.

The overall trend of the two ships on the Y-coordinate chart is highly similar, indicating that they are mostly sailing on the same or parallel courses. The two ships remained relatively stable at first, then the y-value increased, reaching a peak of about 440-450 at around frame 150, after which it began to decline, reaching a low point of about 330 at around frame 300, and then increased again. As shown in Figure 11, the overall lateral movement of the hulls indicates that the two ships are performing some form of “lateral

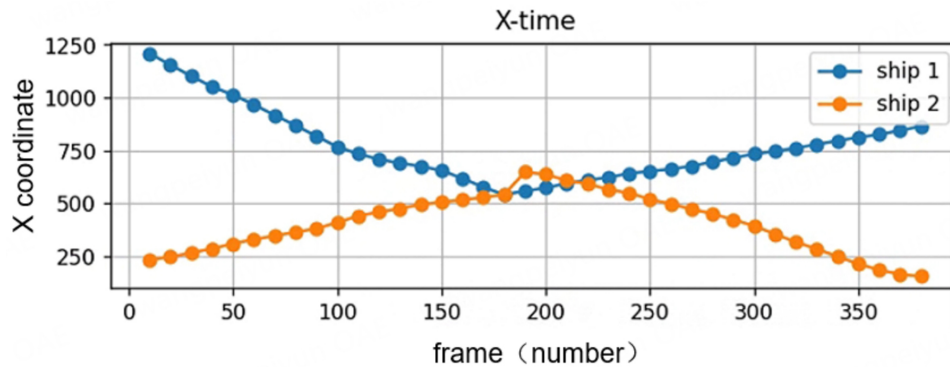


Figure 10. The motion of the x-axis of the double vessel with respect to time.

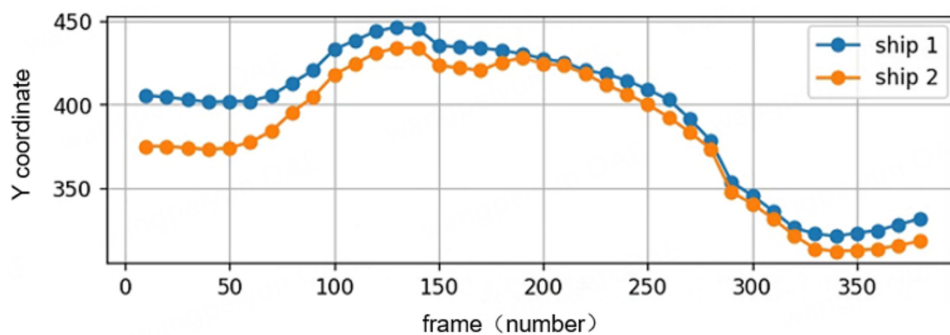


Figure 11. The motion of the y-axis of the double vessel with respect to time.

adjustment” or “obstacle avoidance operation”. In particular, between the 150th and 300th frames, the two ships decline almost synchronously, further suggesting that they may be avoiding each other to prevent a collision. This is also consistent with the state of the X-coordinate.

The trajectories of the two ships in the two-dimensional plane are shown in Figure 12. At frame 180, the X values of the two ships are the same and the Y values are almost coincident, implying that the positions of the two ships in the two-dimensional plane are almost identical, which is consistent with a collision or brushing situation. The two ships come from opposite directions and reach the midpoint at almost the same time, which has the typical characteristics of a rendezvous and creates the risk of collision. After the meeting, both vessels show signs of turning around, and the Y-direction trajectories are highly similar, which indicates that the two vessels avoid collision by spontaneous avoidance behavior, and the two vessels basically maintain a stable heading except for the intermediate part of the avoidance adjustment.

4. CONCLUSION

Aiming to address the current key technical difficulty of small target object recognition for water transport traffic situational awareness under low-light conditions, the EG-YOLO+ framework proposed in this study effectively solves this difficulty by deeply fusing EnlightenGAN ship image enhancement with the improved YOLOv11 ship detection algorithm. In terms of ship detection, the shortcomings in the representation of small target information are effectively addressed by fusing the SE module and YOLOv11. The method achieves 80.1% mAP₅₀ detection accuracy on the self-built dataset, which represents a 6-percentage point improvement over the original YOLOv11 model (74.1% mAP₅₀) under identical experimental conditions. Especially in complex scenarios such as ship encounters, the system can accurately identify ship trajectories

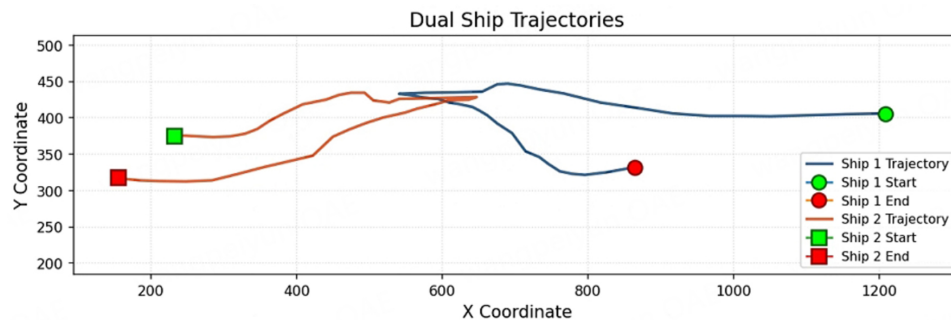


Figure 12. Two-dimensional planar motion trajectory of the double vessel.

and analyze collision avoidance behaviors, which verifies its utility value in intelligent shipping systems. This study not only provides a new technical idea for ship detection in low-light environments but also offers reliable technical support for the intelligent development of water transport. Future research will further optimize the model efficiency and expand the multimodal sensing capability to improve the generalization performance of the method in different application environments.

DECLARATIONS

Authors' contributions

Writing - review and editing, supervision, formal analysis, conceptualization: Chen, X.

Writing - original draft, visualization, validation, methodology, investigation: Yang, R.

Data curation, validation: Wu, Y.

Formal analysis, conceptualization: Zhang, H.

Writing - review and editing: Ranjitkar, P.

Writing - review, data curation, conceptualization: Postolache, O.

Writing - review and editing, performed data acquisition: Zheng, Y.

Supervision, methodology, conceptualization: Wang, Z.

Availability of data and materials

The dataset of the study used is self-constructed. Other raw data that support the findings of this study are available from the corresponding author upon reasonable request. All figures were programmatically generated using Python libraries, ensuring compliance with copyright regulations of the respective platform.

Financial support and sponsorship

This work was jointly supported by the National Natural Science Foundation of China (Nos. 52472347, 52331012), Open Fund of Chongqing Key Laboratory of Green Logistics Intelligent Technology (Chongqing Jiaotong University) (No. KLGLIT2024ZD001), and Open Fund of Jiangxi Key Laboratory of Intelligent Robot (No. JXINTROB-2024-201).

Conflicts of interest

Chen, X. is an editorial board member of *Intelligence & Robotics* and a guest editor of the special issue “*Intelligent, Safe, and Green Shipping-oriented Maritime Data Exploitation and Knowledge Discovery*”. Chen, X. was not involved in any aspect of the editorial process for this manuscript, including reviewer selection, manuscript handling, or final decision. The other authors declare that there are no conflicts of interest. The author Wu, Y. is affiliated with SPG Qingdao Port Group Co., Ltd. This affiliation does not involve any conflicts of interest with the implementation of this study or the publication of its results. The author Wang, Z. is affiliated with Shanghai Ship and Shipping Research Institute Co., Ltd. He contributed to the real-

world validation and computational optimization of the proposed framework, ensuring its practical applicability in maritime environments. All conclusions remain academically objective and independently verified. All other authors declare no competing interests.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2025.

REFERENCES

1. Sun, L.; Chen, S.; Yao, X.; Zhang, Y.; Tao, Z.; Liang, P. Image enhancement methods and applications for target recognition in intelligent mine monitoring. *J. China. Coal. Soc.* **2024**, *49*, 495-504. DOI
2. Shorten, C.; Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. *J. Big. Data.* **2019**, *6*, 197. DOI
3. Xu, M.; Yoon, S.; Fuentes, A.; Park, D. S. A comprehensive survey of image augmentation techniques for deep learning. *Pattern. Recognit.* **2023**, *137*, 109347. DOI
4. Nagaraju, M.; Chawla, P.; Kumar, N. Performance improvement of deep learning models using image augmentation techniques. *Multimed. Tools. Appl.* **2022**, *81*, 9177-200. DOI
5. Zhuang, P.; Wu, J.; Porikli, F.; Li, C. Underwater image enhancement with hyper-laplacian reflectance priors. *IEEE. Trans. Image. Process.* **2022**, *31*, 5442-55. DOI PubMed
6. Khalifa, N. E.; Loey, M.; Mirjalili, S. A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artif. Intell. Rev.* **2022**, *55*, 2351-77. DOI PubMed PMC
7. Gupta, H.; Kotlyar, O.; Andreasson, H.; Lilienthal, A. J. Robust object detection in challenging weather conditions. In *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, USA. Jan 03-08, 2024. IEEE; 2024. pp. 7508-17. DOI
8. Wang, M.; Bao, J.; Zhang, Q.; et al. Research on target recognition algorithm and track joint detection method for unmanned monorail crane working in underground coal mine. *J. China. Coal. Soc.* **2024**, *49*, 457-71. DOI
9. Lv, Y.; Zhou, Y.; Chen, Q.; Chi, W.; Sun, L.; Yu, L. YOLO_Srv2: an evolved version of YOLO_SR. *Eng. Appl. Artif. Intell.* **2024**, *130*, 107657. DOI
10. Dai, Y.; Liu, W.; Wang, H.; Xie, W.; Long, K. YOLO-former: marrying YOLO and Transformer for foreign object detection. *IEEE. Trans. Instrum. Meas.* **2022**, *71*, 1-14. DOI
11. Zhao, G.; Ding, J. Research on multi-modal image target recognition based on asynchronous depth reinforcement learning. *Aut. Control. Comp. Sci.* **2022**, *56*, 253-60. DOI
12. Cao, X.; Su, Y.; Geng, X.; Wang, Y. YOLO-SF: YOLO for fire segmentation detection. *IEEE. Access.* **2023**, *11*, 111079-92. DOI
13. Gomaa, A.; Minematsu, T.; Abdelwahab, M. M.; Abo-Zahhad, M.; Taniguchi, R. Faster CNN-based vehicle detection and counting strategy for fixed camera scenes. *Multimed. Tools. Appl.* **2022**, *81*, 25443-71. DOI
14. Tran, D. Q.; Aboah, A.; Jeon, Y.; Shoman, M.; Park, M.; Park, S. Low-light image enhancement framework for improved object detection in fisheye lens datasets. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, USA. Jun 17-18, 2024. IEEE; 2024. pp. 7056-65. DOI
15. Jiang, Y.; Gong, X.; Liu, D.; et al. EnlightenGAN: deep light enhancement without paired supervision. *IEEE. Trans. Image. Process.* **2021**, *30*, 2340-9. DOI
16. Rasheed, A. F.; Zarkoosh, M. YOLOv11 optimization for efficient resource utilization. *J. Supercomput.* **2025**, *81*, 7520. DOI
17. Gao, H.; Ji, S. Graph U-Nets. *IEEE. Trans. Pattern. Anal. Mach. Intell.* **2022**, *44*, 4948-60. DOI
18. Wang, J.; Yang, P.; Liu, Y.; et al. Research on improved YOLOv5 for low-light environment object detection. *Electronics* **2023**, *12*, 3089. DOI
19. Alif, M. A. R. YOLOv11 for vehicle detection: advancements, performance, and applications in intelligent transportation systems. *arXiv* **2024**, arXiv:2410.22898. <https://doi.org/10.48550/arXiv.2410.22898>. (accessed 29 Jul 2025)
20. Cai, Y.; Luan, T.; Gao, H.; et al. YOLOv4-5D: an effective and efficient object detector for autonomous driving. *IEEE. Trans. Instrum. Meas.* **2021**, *70*, 1-13. DOI
21. Wu, Z.; Guo, K.; Wang, L.; Hu, M.; Ren, S. A collaborative learning-based urban low-light small-target face image enhancement method. *ACM. Trans. Sen. Netw.* **2023**,. DOI
22. Hao, S.; Wang, Z.; Sun, F. LEDet: a single-shot real-time object detector based on low-light image enhancement. *Comput. J.* **2021**, *64*, 1028-38. DOI
23. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern*

- Recognition*, Salt Lake City, USA. Jun 18–23, 2018. IEEE; 2018. pp. 7132–41. [DOI](#)
24. Zhang, M.; Ye, S.; Zhao, S.; Wang, W.; Xie, C. Pear object detection in complex orchard environment based on improved YOLO11. *Symmetry* **2025**, *17*, 255. [DOI](#)
 25. Su, M.; Pu, R.; Wang, Y.; Yu, M. A collaborative siege method of multiple unmanned vehicles based on reinforcement learning. *Intell. Robot.* **2024**, *4*, 39–60. [DOI](#)