2025 World Journal of Innovation and Modern Technology, Vol.8, Issue 9 (Sep)
ISSN 2682-5910
DOI: 10.53469/wjimt.2025.08(09).14

Research on Ship Target Detection Based on YOLOv8

Wanqiu Xu

Jiangsu Maritime Institute, Nanjing 211170, Jiangsu, China

Abstract: ***With the rapid development of the marine economy, maritime traffic supervision and safety management have raised higher requirements for ship detection. Traditional methods suffer from low efficiency and high false detection rates, making it difficult to meet the needs of intelligent maritime management. This paper proposes a ship target detection model based on YOLOv8. Experimental results show that the model achieves 99.15% mAP@50 and 85.14% mAP@50:95, effectively handling ship recognition tasks under complex sea conditions, providing reliable technical support for smart ocean construction and maritime supervision.***

Keywords: YOLOv8; Ship Target Detection; Deep Learning; Intelligent Maritime.

1. INTRODUCTION

Against the backdrop of accelerating global trade and ocean development, the density of maritime traffic is continuously increasing, posing higher requirements for real-time detection and recognition of ship targets. Traditional methods, such as radar and manual observation, are inefficient and prone to high false detection rates under complex sea conditions, with limited capability in detecting small vessels, making them inadequate for modern maritime traffic supervision and safety needs. Meanwhile, the increasing number of vessels, the diversity of navigation environments, and the frequent occurrence of extreme weather conditions present further challenges to existing ship detection technologies.

In recent years, the development of computer vision and deep learning technologies has provided new solutions for ship target detection. The introduction of convolutional neural networks (CNNs) has enabled end-to-end automated detection. YOLO series models, with their end-to-end training, efficient inference, and high accuracy, have become one of the mainstream approaches in object detection. From YOLOv3 to YOLOv5, continuous iteration has improved both detection accuracy and inference speed. However, challenges remain in low visibility, dense occlusion, and small target detection tasks.

As the latest generation improved model, YOLOv8 optimizes both architectural design and training mechanisms. It adopts an anchor-free detection mechanism and decoupled detection head, reducing the limitations of prior box designs. At the same time, it incorporates multi-scale feature fusion and dynamic label assignment strategies, enhancing robustness for small target detection and multi-target recognition in complex backgrounds. Furthermore, YOLOv8 maintains excellent inference speed, satisfying real-time detection requirements.

Therefore, research on YOLOv8-based ship detection is of great academic significance and can provide technical support for practical applications in smart ocean development, intelligent shipping management, and port safety supervision. Experimental validation on multi-scenario and multi-modal datasets can further promote the transition of ship target detection from laboratory research to engineering applications, laying a foundation for building intelligent maritime traffic systems.

2. REVIEW OF SHIP TARGET DETECTION TECHNOLOGIES

2.1 Fundamentals of Deep Learning

Deep learning, as an important branch of machine learning, has achieved revolutionary progress in the field of computer vision, providing strong technical support for object detection tasks. Compared with traditional methods based on HOG, SIFT features, and SVM classifiers, deep learning automatically learns hierarchical feature representations from raw data by constructing multi-layer neural networks, avoiding the cumbersome process of manual feature design. In ship target detection, although traditional methods are fast and memory-efficient, they lack robustness under complex sea conditions (e.g., fog, low visibility, strong light interference), often resulting in missed and false detections [1].

Convolutional Neural Networks (CNNs) are the core architecture of deep learning in image processing, constructed by stacking convolutional layers, pooling layers, and fully connected layers [2]. With continuous architectural optimization in recent years, efficient CNN architectures such as ResNet [3] and DenseNet [4] have emerged. Through mechanisms such as residual and dense connections, these networks effectively mitigate gradient vanishing problems and perform well in large-scale image recognition and detection tasks. These advanced architectures provide the foundation for high-accuracy ship detection, particularly in multi-scale target detection under complex sea conditions.

2.2 YOLOv8 Model Theory

As a representative of single-stage object detection, the YOLO series has gained wide attention for balancing real-time performance and accuracy. YOLOv8, released by Ultralytics in 2023, consists of three main components: backbone, neck, and head. The backbone employs a lightweight CNN for feature extraction [5]. The neck adopts a Path Aggregation Network (PANet) structure to achieve multi-scale feature fusion from both top-down and bottom-up pathways, enhancing detection capability for ships of different sizes. The detection head outputs final results, including bounding boxes and classification labels.

2.3 Development and Application of Ship Detection Technologies

Depending on data sources and processing methods, ship detection technologies can be classified into three categories: optical image-based detection, SAR image-based detection, and multi-modal fusion detection. Optical image detection utilizes visible or infrared cameras, offering high resolution and rich detail but is easily affected by lighting and weather. Woo et al. introduced attention mechanisms and feature pyramid networks to improve the robustness of multi-scale ship detection in optical images [6]. SAR image detection, with all-day and all-weather capabilities, is suitable for offshore and extreme environments. As ships in SAR imagery are often small targets, Li et al. introduced an improved feature pyramid structure and designed a noise-resistant loss function to suppress sea clutter interference, significantly improving detection accuracy [7]. Multi-modal fusion integrates optical images, infrared, SAR, and AIS (Automatic Identification System) data, allowing systems to leverage complementary modalities for high-accuracy, robust detection. For example, the DBW-YOLO model combines deformable convolution and a BiFormer attention-based feature enhancement module, effectively improving detection performance for nearshore and small vessels [8].

Ship target detection has evolved from traditional methods to deep learning and multi-modal fusion approaches. The future trend will combine lightweight models with multi-modal data to ensure detection accuracy while enhancing deployment capabilities in edge devices and large-scale applications.

3. EXPERIMENTS AND RESULTS ANALYSIS

3.1 Dataset

This study adopts the public ship detection dataset SeaShips [9-11], which contains 7,000 images in total, covering six categories: ore carrier, fishing boat, bulk cargo ship, general cargo ship, container ship, and passenger ship. The input image resolution is 640×640 pixels. The dataset is randomly split into training, validation, and testing sets at an 8:1:1 ratio.

Table 1: SeaShips Ship Detection Dataset

Ship Type	Ore carrier	fishing boat	bulk cargo carrier	general cargo ship	container ship	passenger ship
Number of Images	2199	2190	1952	1505	901	474

3.2 Results and Analysis

3.2.1 Experimental Results

To comprehensively and objectively evaluate the performance of the model in ship target detection tasks, this paper adopts the following evaluation metrics: precision (P), recall (R), mean Average Precision at IoU threshold 0.50 (mAP@0.50), and mean Average Precision across IoU thresholds from 0.50 to 0.95 (mAP@0.50:0.95). As shown in Table 2, the model achieved 99.15% mAP@0.50 and 85.14% mAP@0.50:0.95. Specifically, mAP@0.50 refers to the mean of the Average Precision (AP) values calculated for all categories when the IoU threshold is set to 0.50, while mAP@0.50 - 0.95 represents the average mAP when the IoU threshold varies from 0.50 to 0.95. The formulas for calculating these performance metrics are as follows:

$$P=\frac{N_{TP}}{N_{TP}+N_{FP}}$$

$$R=\frac{N_{TP}}{N_{TP}+N_{FN}}$$

$$P_A=\int_0^1 P(t)dt$$

$$mAP=\frac{\sum_{n=1}^{N} P_{A,n}}{N}$$

Where: N_{TP} represents the number of correctly recognized positive samples; N_{TN} represents the number of correctly recognized negative samples; N_{FP} represents the number of negative samples misclassified as positive samples; N_{FN} represents the number of positive samples misclassified as negative samples; and N represents the total number of samples across all images.

Table 2: Performance Metrics

Model	P(%)	R(%)	mAP@50(%)	mAP@50: 95(%)
YOLOv8	98.63	98.21	99.15	85.14

3.2.2 Model Training

From the variation trends of various metrics during the model training process (as shown in Figure 1), the box loss, classification loss (cls loss), and distribution focal loss (dfl loss) on both the training and validation sets decreased steadily. This indicates that the model's fitting ability in terms of bounding box regression and category discrimination was continuously enhanced. Meanwhile, the precision and recall metrics increased rapidly in the early stage of training and tended to stabilize at a high level in the later stage, demonstrating that the model possessed excellent detection precision and recall capabilities. In addition, the two metrics of mAP@0.5 and mAP@0.95 also continued to rise, eventually approaching 1.0 and 0.85 respectively. This reflects that the model exhibited strong detection performance across all categories and different IoU thresholds. The overall training process was stable without overfitting, indicating that the model performed excellently on the constructed dataset.

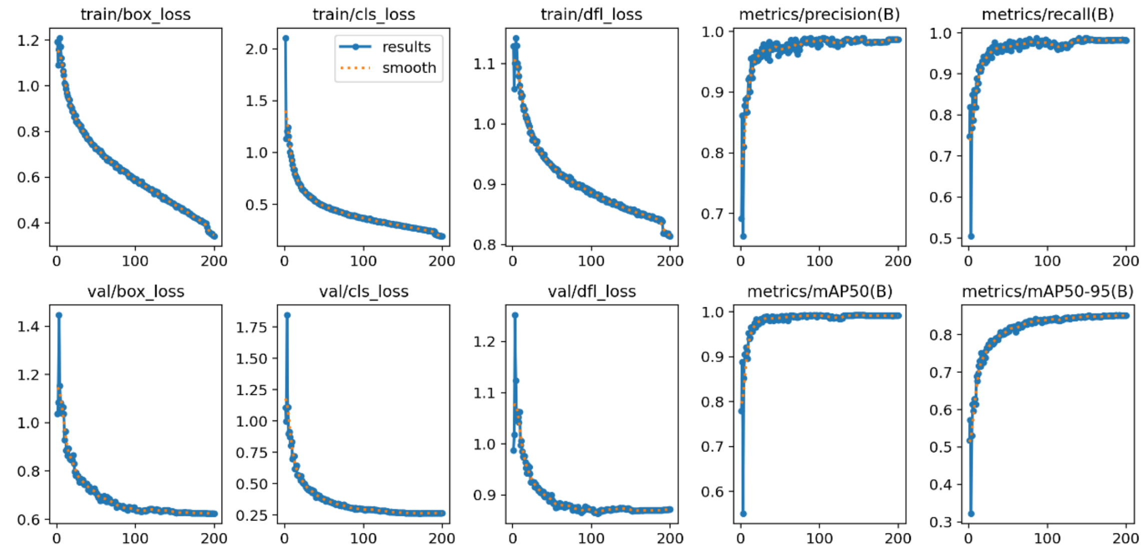

Figure 1: Training Result Graph

3.2.3 Confusion Matrix

The confusion matrix demonstrates the classification performance of the model on seven categories of targets (including six types of ships and the background). As can be seen from Figure 2, the model achieved good recognition results for most categories. For example, "ore carriers" were correctly identified 203 times, with only a small number of them incorrectly classified as "background"; "fishing boats" were correctly identified 175 times, but there were still cases where they were misclassified as "background".

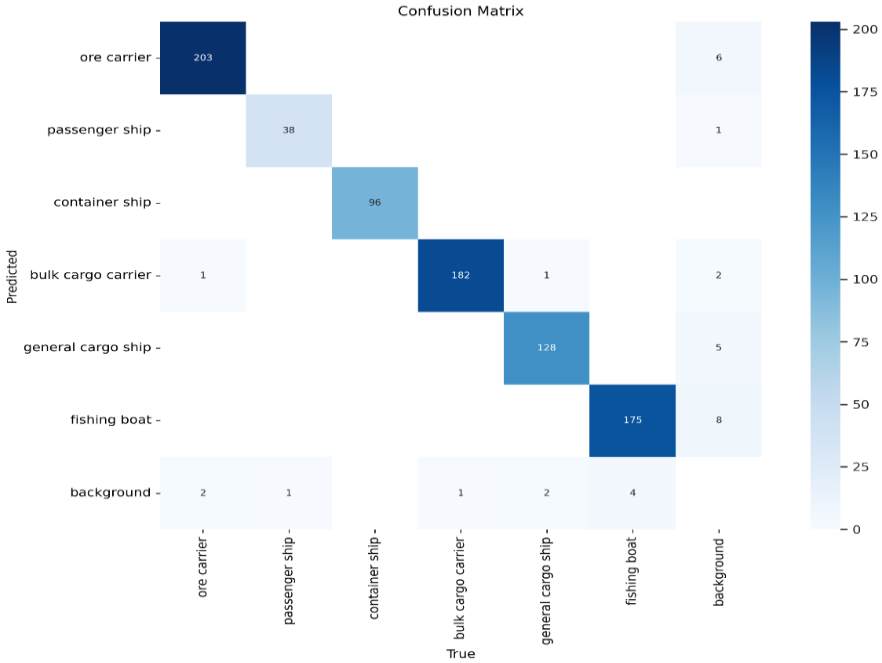

Figure 2: Schematic Diagram of the Confusion Matrix

3.2.4 Visualization Analysis

Figure 3: Ship Detection Visualization Diagram

In ship detection tasks, the application effect of the YOLOv8 model can be intuitively presented through detection result graphs. As shown in Figure 3, it can be observed that the model can detect and identify the types of ships in actual scenarios, and mark them with bounding boxes of different colors.

The visualization results in Figure 3 allow observers to directly judge two core aspects: first, whether the model accurately identifies ship targets; second, whether the positions of the detection boxes generated by the model are accurate. Based on the detection results presented in this figure, researchers can further analyze the reasons for the model's false detection and missed detection issues, and use this as a basis to carry out targeted model improvement work, such as adjusting model parameters and optimizing the network structure. In addition, this figure can also be used to intuitively demonstrate the application effect of the YOLOv8 model in ship detection to non-professional personnel.

4. CONCLUSION

This paper constructed a YOLOv8-based ship target detection model and validated it on the SeaShips dataset. Experimental results showed the model achieved 99.15% mAP@0.50 and 85.14% mAP@0.50:0.95, providing reliable technical support for intelligent maritime, smart shipping, and port supervision applications.

However, limitations remain. The study was based on a single public dataset, lacking validation across datasets and real-world scenarios. Additionally, under extreme weather, small target detection, and dense ship scenes, some false detections and missed detections persist. With the continuous development of deep learning and multi-source data fusion technologies, future ship detection systems are expected to achieve higher accuracy, lower latency, and broader applicability, offering stronger technical guarantees for maritime safety and efficient management.

ACKNOWLDEGEMENT

1) 2024 Jiangsu Provincial College Students' Innovation and Entrepreneurship Training Program (GX-20240158)

2) Scientifc Research Project of Jiangsu Maritime Institute (2024ZKyb11)

REFERENCES

[1] Zhou, Zhenzhen et al. "A Ship Tracking and Speed Extraction Framework in Hazy Weather Based on Deep Learning", Journal of marine science and engineering 11.7 (2023).
[2] Li, Haoxiang et al. "A Convolutional Neural Network Cascade for Face Detection.", Computer Vision and Pattern Recognition (2015): 5325-5334.
[3] He, Kaiming et al. "Deep Residual Learning for Image Recognition", Computing Research Repository (2016): 770-778.
[4] Huang G, Liu Z, Laurens V D M,et al. Densely Connected Convolutional Networks [J]. IEEE Computer Society, 2016.DOI:10.1109/CVPR.2017.243.
[5] Varghese, Rejin, and Sambath M.. "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness", 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS) (2024): 1-6.
[6] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module [J]. Springer, Cham, 2018. DOI:10.1007/978-3-030-01234-2_1.
[7] Li J, Qu C, Peng S, et al. Ship detection in SAR images based on convolutional neural network [J]. Xi Tong Gong Cheng Yu Dian Zi Ji Shu/Systems Engineering and Electronics, 2018, 40(9):1953-1959.
[8] Tang, Xiao et al. "DBW-YOLO: A High-Precision SAR Ship Detection Method for Complex Environments", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 17 (2024): 7029-7039.
[9] Shao Z, Wu W, Wang Z, et al. SeaShips: A Large-Scale Precisely-Annotated Dataset for Ship Detection [J]. IEEE Transactions on Multimedia, 2018, 20(10):1-1.
[10] Zheng, Yanmei, Guanghui Zhou, and Bibo Lu. "Rebar Cross-section Detection Based on Improved YOLOv5s Algorithm." Innovation & Technology Advances 1.1 (2023): 1-6. https://doi.org/10.61187/ita.v1i1.1
[11] Zhao, Xiujuan, Lei Zhang, and Zhonghua Hu. "Smart warehouse track identification based on Res2Net-YOLACT+ HSV." Innovation & Technology Advances 1.1 (2023): 7-11. https://doi.org/10.61187/ita.v1i1.2