International Journal of Science and Engineering Applications
Volume 11-Issue 04, 50 - 55, 2022, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1104.1001

An Improved YOLO v4 Algorithm-based Object Detection Method for Maritime Vessels

Guowen He
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Wenlong Wang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Bowen Shi
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Shijie Liu
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Hui Xiang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Xiaoyuan Wang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Abstract: Ship object detection is the core part of the maritime intelligent ship safety assistance technology, which plays a crucial role in ship safety. The object detection algorithm based on the convolutional neural network has greatly improved the accuracy and speed of object detection, which YOLO algorithm stands out among the object detection algorithms with more excellent robustness, detection accuracy, and real-time performance. Based on the YOLO v4 algorithm, this study uses the k-means algorithm to improve clustering at the input side of image data and introduces relevant berth data in the self-organized dataset to achieve detection of ships and berths for the lack of detection of berths in the existing ship detection algorithm. The experimental results show that the mAP and F1-score of the improved YOLO v4 are increased by 2.79% and 0.80%, respectively. The improved YOLO v4 algorithm effectively improves the accuracy of ship object detection, and the in-port berth also achieves better detection results and improves the ship environment perception, which is important in assisting berthing and unberthing.

Keywords: YOLO v4; deep learning; ship and berth object detection; ship safety; k-means

1. INTRODUCTION

With the increase of maritime cargo transportation, the ship monitoring system is becoming more and more important for the detection of abnormal ship behavior, but the data collected by the monitoring system is usually not complete enough, and even makes the operator make wrong judgments and cause accidents. Therefore, new technologies are needed. To fill in data such as marine obstacles, to ensure the safe navigation of ships at sea, reduce the instability of ship activities, and reduce the occurrence of ship collision accidents [1]. Ship object detection is one of the core functions of the ship monitoring system and an important part of the marine intelligent ship safety auxiliary technology. Therefore, the research on ship object detection has important research significance. A large number of researchers are devoted to the research and development of the safety auxiliary technology for ship object detection, trying to apply image processing, neural network, and other technologies to the field of ship detection, and have achieved certain achievements [2]. However, these methods also have shortcomings such as single applicable scene, poor real-time performance, and low ship detection accuracy. Especially the lack of detection of port berths, the safe berthing and unberthing of ships is the focus of port work, because the captain cannot accurately identify the surrounding of the ship and the port environment will cause many accidents [3]. Therefore, in-depth research on the problem of port berth object detection is required.

Aiming at the complex marine environment, the ship object detection algorithm based on traditional methods such as background subtraction and image segmentation is easily affected by scene changes. The influence of sunlight and waves will greatly reduce the detection accuracy, and it is more sensitive to the selection of object features. The chemical effect is poor and cannot meet the needs of the development of intelligent ships [4]. The object detection algorithm based on deep learning has good real-time performance and average accuracy in various marine environments, can obtain better detection results than traditional methods, and has strong robustness, to a certain extent, it can make Autonomous ships navigate safely and avoid collisions with other ships [5,6].

With the development of artificial intelligence technology, the traditional object detection algorithm is gradually replaced by the detection algorithm based on deep learning, especially the object detection algorithm based on convolutional neural network, because of its higher detection accuracy and faster detection speed, which stands out among many object detection algorithms [7]. The convolutional neural network is mainly composed of a feature extraction layer and feature mapping layer. When an image is input to a convolutional neural network, the network can directly perform computational processing on the image, which greatly improves the performance of the algorithm compared to the process of feature extraction in traditional methods [8]. The object detection algorithm based on a convolutional neural network is divided into two categories: Two-stage and One-stage according to the detection process. The Two-stage detection algorithm first generates the area to be detected, and then detects the object in this area, such as R-CNN [9], Fast R-CNN [10], Faster R-CNN [11], etc., but this type of algorithm detects The speed is slow and cannot meet the needs of real-time detection. The one-stage algorithm directly generates the category probability and position coordinates of the object, and the final detection result can be obtained in a single detection, such as SSD[12], YOLO[13], YOLO v2[14], YOLO v3[15], YOLO v4 [16].

Scholars at home and abroad have done a lot of research on ship object detection using a convolutional neural network algorithm. Zhang [17] et al., based on the Two-stage method, proposed a region-based R-CNN method to solve the problem of poor performance of traditional SAR image-based ship detection methods in small and clustered ships. Detect ships in high-resolution remote sensing images, and achieve higher recall and accuracy on small and clustered ships by improving the Faster-RCNN method. Based on the One-stage method, especially the YOLO algorithm, Chen [18] et al. improved the feature extraction ability of the model in the complex navigation environment by improving the YOLO v3 network and introducing the attention mechanism, and tested it on the SeaShip dataset. The detection accuracy is high. Huang [19] et al. improved the YOLO v3 network and proposed a new method Ship-YOLOv3, which reduced some convolution operations by changing the network structure of YOLO v3, and compared experiments in different environments to detect The time is reduced by an average of 6.06ms, which improves the ship detection accuracy by 12.5% and the recall rate by 11.5% while maintaining high recognition ability.

Aiming at the problems of low detection accuracy and lack of berth detection in the current ship detection model, this paper proposes a ship object detection method based on the improved YOLO v4 algorithm. Based on the YOLO v4 algorithm, the proposed algorithm introduces the K-means algorithm to cluster the a priori boxes in the data set and introduces the berth data in the data set to improve the YOLO v4 algorithm, which improves the detection effect of ships and berths.

2. Method

2.1 YOLO v4 Algorithm

The YOLO (You Only Look Once) algorithm was first proposed by Redmon et al. in 2016 as a regression-based object detection algorithm, which is a typical one-stage algorithm. As the fourth version of the YOLO series, the YOLO v4 algorithm has higher detection accuracy and better real-time performance than the previous work, and the overall structure is quite different from the previous work. The overall structure is shown in Figure 1.

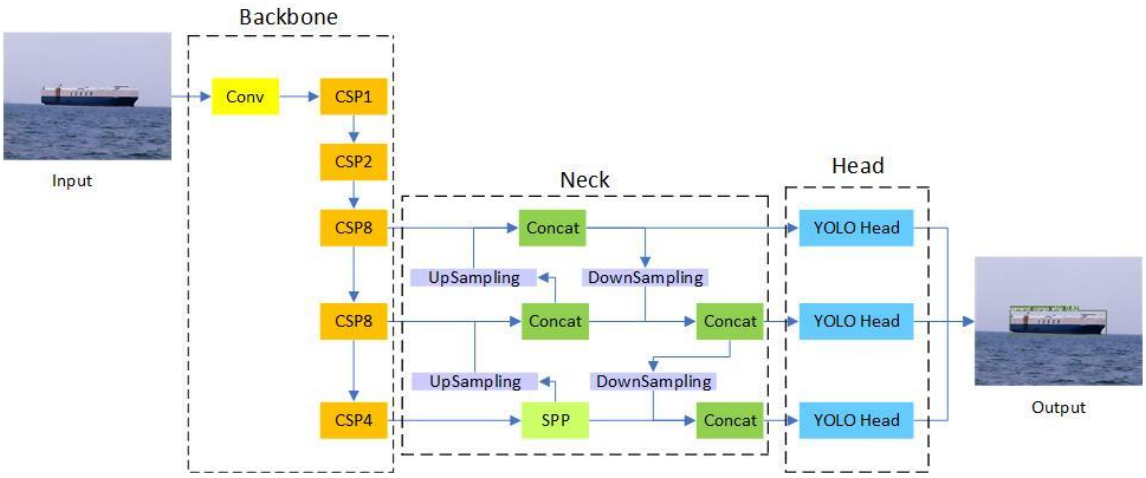

Figure 1. The network structure of YOLOv4

The YOLO v4 algorithm model Backbone backbone network consists of CSPDarknet53, the Neck part consists of a spatial pyramid pooling layer SPP and a feature pyramid PANet, and the Head part consists of the YOLO v3 detection head. The backbone network CSPDarknet53 uses the CSP module composed of ResUnit components as the feature extraction part of the overall structure.

The spatial pyramid pooling layer SPP uses pooling layers of different sizes to process the feature maps obtained through the convolution layer processing and enlarges the receptive field by fusing features of different scales. The SPP structure is shown in Figure 2.

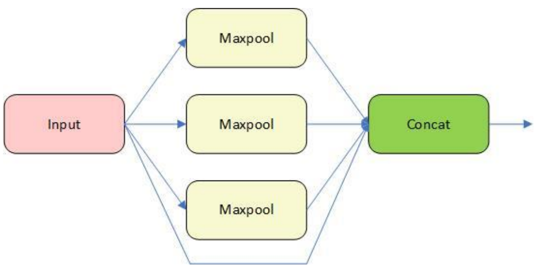

Figure 2. SPP structure diagram

PANet is composed of an FPN+PAN structure, which can perform feature aggregation for different detection layers from different backbone layers. Among them, FPN transfers and fuses high-level feature information through upsampling from top to bottom, and transfers strong semantic features from top to bottom. YOLO v4 is different from YOLO v3's FPN. YOLO v4 adds a bottom-up feature pyramid after FPN, which contains two PAN structures to transfer strong localization features from bottom to top.

2.2 Build ship and berth datasets

This study establishes a ship and berth dataset for evaluating the effectiveness of the improved YOLO v4 detection algorithm. Take advantage of the onboard camera on the smart experimental boat (as shown in Figure 3). In a daytime good weather scenario, real-world images of ships and berths in the port were collected as positive samples, and images without ships were collected as negative samples.

Figure 3 Diagram of the intelligent experimental ship and onboard camera equipment

The intelligent experimental ship is 21.08 meters long, 5.40 meters wide, and 2.90 meters high. It has been converted into an autonomous maritime intelligent ship, equipped with on-board cameras, millimeter-wave radar, a global positioning system (GPS), an inertial navigation system. Due to the insufficient number of images of ships and berths in the port captured by the intelligent experimental ship, this study uses open-source datasets (MS COCO [20] and VOC datasets) to filter out the images of ships that contain them as a supplement to the positive training samples to construct ship and berth images.

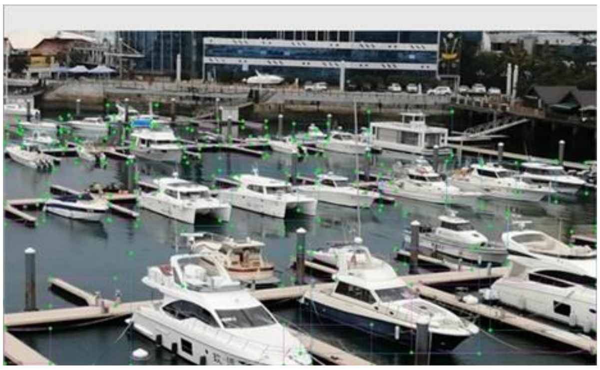

Figure 4 Annotated images

Annotate the ship and berth datasets, and the generated annotation labels include the image name, the object category, and the coordinates and size of the object bounding rectangle. The generated annotated image is shown in Figure 4. A total of 6,634 images containing ships and berths were collected in the ship and berth dataset, and the ratio of 7:3 was randomly selected and divided into a training set and test set.

2.3 Algorithm Improvement Based on K-means Clustering

In object detection tasks, choosing a suitable prior box can significantly improve the speed and accuracy of object detection. The a priori box is a box with a fixed aspect ratio that is preset in the image. In the YOLO v4 algorithm, each grid is generated from three a priori boxes to three bounding boxes. The original YOLO v4 model is trained on the MS COCO dataset and the VOC dataset, and generates 9 sets of default a priori boxes based on 80 categories in its dataset. The results are as follows: (12, 16), (19, 36), (40, 28), (36, 75), (76, 55), (72, 146), (142, 110), (192, 243), (459, 401). However, the default a priori boxes generated on the MS COCO dataset and VOC dataset do not match the object size in the ships and berth datasets used in this paper, which affects the detection accuracy. In this paper, the K-means clustering algorithm is used to cluster the dataset, and 9 sets of adaptive a priori boxes are generated. The generated results are as follows: (22, 14), (77, 19), (41, 41), (152, 42), (80, 85), (226, 89), (129, 201), (334, 161), (359, 349). The clustering algorithm can speed up the convergence of the network and effectively improve the problem of gradient descent in the training process.

$$\mathrm{d(box,\ centroid)} = 1 - \mathrm{IOU(box,\ centroid)} \tag{1}$$

$$IOU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{2}$$

Equation (1) is an indicator for evaluating the bounding box and the clustered prior box, where d is the distance between the object bounding box and the prior box, IOU is the intersection of the bounding box and the prior box, such as Eq. (2) can be used to measure how similar two boxes are. The bounding box with the smallest d can be classified as the prior box. The model can predict the position of the object detection frame from the data of the prior frame and the bounding box, and the formula is as follows.

$$b_x = \sigma(t_x) + c_x \tag{3}$$

$$b_y = \sigma(t_y) + c_y \tag{4}$$

$$b_w = s_w e^{t_w} \tag{5}$$

$$b_h = s_h e^{t_h} \tag{6}$$

Among them, b_x and b_y are the center coordinates of the ship and berth object detection frame predicted by the model, t_x and t_y are the center coordinates of the object bounding box output, c_x and c_y are the coordinates of the upper left corner of the grid corresponding to the entire picture, $\sigma(t_x)$ and $\sigma(t_y)$ are the offsets relative to the c_x and c_y coordinates after being processed by the activation function in the network. b_w and b_h are the width and height of the predicted ship and berth detection boxes, respectively. t_w and t_h are the width and

International Journal of Science and Engineering Applications
Volume 11-Issue 04, 50 - 55, 2022, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1104.1001

height of the bounding box of the object, and s_w and s_h are the width and height of the prior box.

3. Model training and validation

3.1 Model Training

The experimental deep learning framework in this chapter is Tensorflow. The experiments were carried out on a 64-bit computer, the CPU model was Intel Corei5-9400F, the GPU model was NVIDIA GTX1660S, 8GB memory, and the operating system was Windows 10. For the YOLOv4 algorithm, the training parameters are set according to the device hardware, the initial batch size is set to 16, the initial maximum number of training iterations is set to 8000, the initial learning rate is set to 0.001, and the rest of the parameters are set to default values.

During the training process, parameters such as the maximum number of training iterations and the learning rate are continuously adjusted to improve the accuracy of the model and reduce the loss. The learning rate is initially set to a relatively large value, which can speed up the convergence. When the loss curve oscillates or the rate of decline slows, pause training and continue training with a smaller learning rate.

3.2 Evaluation indicators

In this study, mAP (Mean Average Precision) and F1-score are used to evaluate the proposed algorithm.

$$AP = \int_0^1 P(R)dR \tag{7}$$

$$F1\_score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

mAP is the average value of the average precision AP (Average Precision) of the ship and berth detection objects. A curve is drawn according to the accuracy and recall rate of the test results of each category. The area under the curve is the average precision. The calculation formula is shown in formula (7). F1-score is an important indicator to measure the quality of the detection model. It takes into account precision and recall, and is the harmonic mean of precision and recall. The larger the value, the better the model. Its calculation is shown in the formula (8). The accuracy is the detection result index of each category, that is, the percentage of correctly detected ships and berths in all detection areas. The calculation method is shown in formula (9). The recall rate, that is, the percentage of correctly detected ships and berths in all detection results, is calculated as shown in formula (10). Among them, TP is the number of correctly detected ships and berth detection objects, FP is the number of ships and berth detection objects detected in the background, and FN is the number of detected detection objects in the background.

3.3 Model testing and validation

In this study, the model verification of the improved YOLO v4 marine ship object detection algorithm is carried out through the test set part of the ship and berth data set and the real ship test. After the training of the improved model is completed, the object detection test on the test set is shown in Figure 6.

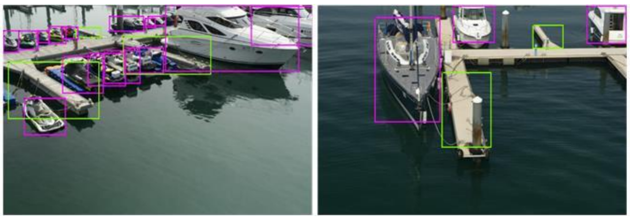

Figure 6 Partial detection results of the test set

The red detection box represents the detected ship object, and the green detection box represents the detected berth object in the port. In comparison, the detection rate of ships is higher than that of berths. On the one hand, because the number of ships in the data set is relatively large, the characteristics of the ships are relatively well fitted, and on the other hand, the characteristics of the ships are relatively compared. Obviously, there are many places where the characteristics of the berths in the port are similar to those of the surrounding environment.

The improved YOLO v4 algorithm is deployed on real ships to detect ships and berths in the port, and the detection effect is shown in Figure 7.

Figure 7 The improved YOLO v4 model real ship detection diagram

The experimental results show that the improved method can achieve accurate detection of ship objects and berths. The improved method proposed in this study is a more effective method for ship detection, and also achieves better detection for berths, which can comprehensively improve the detection ability of YOLO v4 for ships and berth objects at sea.

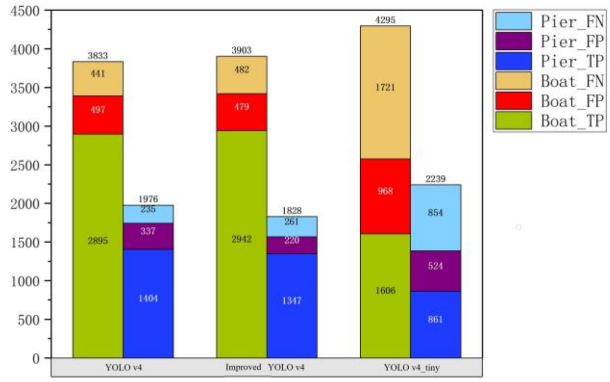

Figure 8 Number of category objects in the YOLO v4 model

This study compares the YOLO v4 model using the default prior frame, the YOLO v4_tiny with smaller model size, and the improved YOLO v4 on the ship and berth datasets. The number of detection objects of ships and berths in different models (as shown in Figure 8), the improved YOLO v4 model has better TP, FP, and FN in the two categories of ships and berths, resulting in better precision and recall The algorithm evaluation index is shown in Table 1.

Table 1 Evaluation indicators of the improved YOLO v4 algorithm

Method	Precision	Recall	F1-score	mAP(%)
YOLO v4	0.8375	0.8526	0.8550	81.19
YOLO v4_tiny	0.6231	0.4893	0.5482	54.45
Improved YOLO v4	0.8599	0.8638	0.8618	83.98

Among them, the mAP of YOLO v4 is 81.19%, and the F1-score is 0.8550. The mAP of the improved YOLO v4 algorithm is 83.98%, and the F1-score is 0.8618. Compared with the unimproved algorithm, the mAP is increased by 2.79%, and the F1-score is increased by 0.80%. The mAP of YOLO v4_tiny is only 54.45%, and the F1-score is 0.5482. Although the YOLO v4_tiny model is small in size and fast in detection, the detection accuracy is not high and cannot meet the needs of ship detection. Compared with the default model, the improved YOLO v4 detection method proposed in this paper further improves the object detection effect.

4. CONCLUSION

Ship object detection technology is of great significance in the fields of sea surface monitoring, ship collision avoidance, and auxiliary berthing and berthing. This paper proposes a method for detecting objects at sea based on the improved YOLO v4 algorithm, focusing on the problem of insufficient detection accuracy of the YOLO v4 algorithm in the ship and berth data sets. On the basis of the YOLO v4 algorithm, the K-means clustering algorithm is introduced to perform Improvements that have further improved the detection accuracy. The experimental results show that the improved YOLO v4 algorithm has improved detection accuracy compared with the original algorithm, mAP and F1-score are increased by 2.79% and 0.80% respectively, and the detection speed meets the real-time requirements, which verifies the effectiveness of the improved algorithm. This research provides a theoretical reference for the further practical application of the maritime intelligent ship monitoring system.

5. REFERENCES

[1] Abreu F H, Soares A, Paulovich F V, et al.A trajectory scoring tool for local anomaly detection in maritime traffic using visual analytics[J].ISPRS International Journal of Geo-Information,2021, 10 (6): 412.

[2] Qian Jiang, Zhang Guirong, He Ping, et al. Ship parameter estimation method based on background modeling [J]. Ship Electronics Engineering, 2019, 39 (05): 78-83.

[3] Kim H, Kim D, Park B, et al.Artificial Intelligence Vision-Based Monitoring System for Ship Berthing[J].Ieee Access,2020, 8: 227014-227023.

[4] Zhang L, Zhang Y, Zhang Z, et al.Real-time water surface object detection based on improved faster R-CNN[J].Sensors,2019, 19 (16): 3523.

[5] Fu H, Song G, Wang Y.Improved YOLOv4 marine target detection combined with CBAM[J].Symmetry,2021, 13 (4): 623.

[6] Kim K, Hong S, Choi B, et al.Probabilistic ship detection and classification using deep learning[J].Applied Sciences,2018, 8 (6): 936.

[7] Wang Can, Bu Leping. A review of target detection algorithms based on convolutional neural networks [J]. Ship Electronics Engineering, 2021, 41 (09): 161-169.

[8] Wang Xinli, Jiang Fucai, Ning Fangxin, et al. Ship target detection based on improved convolutional neural network [J]. China Navigation, 2018, 41 (02): 41-45+51.

[9] Uijlings J R R, Van De Sande K E A, Gevers T, et al.Selective Search for Object Recognition[J].International Journal of Computer Vision,2013, 104 (2): 154-171.

[10] Girshick R. Fast r-cnn[C].Proceedings of the IEEE international conference on computer vision,2015: 1440-1448.

[11] Ren S Q, He K M, Girshick R, et al.Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J].Ieee Transactions on Pattern Analysis and Machine Intelligence,2017, 39 (6): 1137-1149.

[12] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C].European conference on computer vision,2016: 21-37.

[13] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C].Proceedings of the IEEE conference on computer vision and pattern recognition,2016: 779-788.

[14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C].Proceedings of the IEEE conference on computer vision and pattern recognition,2017: 7263-7271.

[15] Redmon J, Farhadi A.Yolov3: An incremental improvement[J].arXiv preprint arXiv:1804.02767,2018.

[16] Bochkovskiy A, Wang C-Y, Liao H-Y M.Yolov4: Optimal speed and accuracy of object detection[J].arXiv preprint arXiv:2004.10934,2020.

[17] Zhang S, Wu R, Xu K, et al.R-CNN-based ship detection from high resolution remote sensing imagery[J].Remote Sensing,2019, 11 (6): 631.

[18] Chen D, Sun S, Lei Z, et al.Ship Target Detection Algorithm Based on Improved YOLOv3 for Maritime Image[J].Journal of Advanced Transportation,2021, 2021.

[19] Huang H, Sun D, Wang R, et al.Ship target detection based on improved YOLO network[J].Mathematical Problems in Engineering,2020, 2020.

[20] Lin T-Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context[C].European conference on computer vision,2014: 740-755.