Marine Target Detection and Ranging Algorithm Based on Monocular Camera and YOLOv5 Algorithm

Guoyu Shao[a,*], Hongyu Fu[b], Lei Zhao[c]

Shipping Research Institute, Bohai University, Jinzhou ,121007, China

[a*]2385460260@qq.com,
[b]fuhongyu@qymail.bhu.edu.cn,
[c]zhaolei1@qymail.bhu.edu.cn

Abstract. In order to improve the detection efficiency and accuracy of target ships during navigation, as well as the distance measurement of target ships. This article uses the YOLOv5 model trained with ship data saturation to detect target ships. Using the Canny algorithm to detect the horizon on the sea surface, and measuring the distance of ships based on the geometric form of the Earth's radius. This article uses non training set ocean images to validate the algorithm as an example. The experimental results show that the algorithm can achieve class detection of ships and distance measurement of distant targets, with a detection mAP (mean average accuracy) of 94.3%. The error value of the distance measurement results for targets compared to GPS data is 8.2% of the actual distance.

Keywords: YOLOv5; Monocular camera; Edge detection; Target Detection

1 Introduction

With the development of vision technology in recent years, it is possible to obtain image data containing huge amount of information for aerial aerial photography at sea level, ship-based equipment, and shore-based equipment. How to efficiently and stably recognize and detect marine targets intelligently is a port and one of the important development directions in the field of shipborne equipment. The current intelligent recognition algorithm gradually improves its recognition accuracy by continuously training the training model through image data.

In 2016, Redmon[1]et al. proposed a new target detection method called YOLOv1 for the first time. The overall detection speed of YOLOvl is very fast. The bottom POLO model processes images in real time at 45 frames per second. A smaller network version of Fast YOLO has a processing speed of 155 frames per second. Subsequently, Redmon et al. proposed the YOLOv2 and YOLOv3 algorithms in 2016 and 2018 respectively. In 2020, the YOLOv4 algorithm was released. The YOLOv4 algorithm accelerates the network for production systems and parallel computing, thereby improving the real-time performance of the detection algorithm. YOLOv5 was released within 40 days, which improved the calculation speed of many detection algorithms compared

G. Zhao et al. (eds.), *Proceedings of the 2024 7th International Symposium on Traffic Transportation and Civil Architecture (ISTTCA 2024)*, Advances in Engineering Research 241, . .
https://doi.org/10.2991/978-94-6463-514-0_55

to YOLOv4.

Based on the monocular camera's maritime target ranging, there are currently some works. GLADSTONE R et al. estimate the distance between the tracked ship and the camera through video[2]; PARK J and CHO Y use the vertical pixel distance between the horizon and the target ship in the image as the measurement distance; multi-target tracking contains two basic problems, namely tracking filtering and Data association issues[3]. Tracking filtering predicts the next moment state of the target, data association solves the identity correspondence between multiple targets, and establishes the target trajectory. Tracking filtering algorithms include optical flow method[4], Kalman filter, particle filter[5], Mean-Shift[6], etc. Data association algorithms include nearest neighbor algorithm, PDAF[7], MHT[8], Hungarian algorithm, KM algorithm and so on.

2 Construction of YOLOv5 Detection Platform

2.1 YOLOv5x Model

The YOLOv5x model has more complex convolution kernels and has a better detection effect. In the Focus structure of the YOLOv5x network, the input image has a size of 608×608 pixels. After the Focus structure, the size of the feature map becomes 304×304×80. In the second convolution operation, the YOLOv5x network uses 160 convolution kernels. Therefore, the feature map obtained is 152×152×160, and the following three convolution and down-sampling operations are also based on the same principle. The final feature map vectors are 19×19×1280, 38×38×1280, 76×76× 1280. After the CSP1 and CSP2 structures, the number of convolution kernels is continuously increasing, and the overall calculation amount of the network also increases. Therefore, as the number of convolution kernels increases, the thickness of the feature maps extracted by the network deepens, and deep learning The ability of the company has also been enhanced.

2.2 Adaptive Anchor Frame Calculation

The target size of the image data is different, and it is difficult to determine the size of the anchor frame. Therefore, an adaptive anchor frame determination method is used to calculate the size of the anchor frame[9]. An anchor frame can be defined by the aspect ratio of the frame and the scale (s) of the frame, which is equivalent to a series of preset frame generation rules[10]. The area of the frame generated by different aspect ratios is the same and has the same size. Border calculation formula:

$$\begin{cases} w \times h = s \\ \dfrac{w}{h} = ratio \end{cases} \Rightarrow \begin{cases} w = ratio \times h \\ ratio \times h^2 = s \end{cases} \tag{1}$$

Where w is the width of the frame and h is the height of the frame. According to the rules of the anchor frame, a series of frames can be generated at any position of the image. The adaptive anchor frame can solve the problem that one window can only detect one target and multi-scale. At the same time, anchor frames of different sizes can mark the predicted target in a targeted manner.

2.3 Loss function

YOLOv5 uses GIoU as the loss function of the frame regression, assuming that the coordinates of the predicted frame are $B^p = (x_1^p, y_1^p, x_2^p, y_2^p)$, The coordinates of the actual label border are $B^g = (x_1^g, y_1^g, x_2^g, y_2^g)$ Specifies the prediction frame $x_2^p > x_1^p, y_2^p > y_1^p$.

Obtain the overlapping frame coordinates through B^p and B^g, the calculation formula is as follows:

$$x_1^I = \max(\hat{x}_1^p, x_1^g),\ \ y_1^I = \max(\hat{y}_1^p, y_1^g),\ \ x_2^I = \min(\hat{x}_1^p, x_1^g),\ \ y_2^I = \min(\hat{y}_2^p, y_2^g) \tag{2}$$

According to the formula, calculate the area SI of the overlapping border. The formula is as follows:

$$S^I = \begin{cases} (x_2^I - x_1^I) \times (y_2^I - y_1^I), x_2^I > x_1^I, y_2^I > y_1^I \\ 0, otherwise \end{cases} \tag{3}$$

Then, find the B^c coordinates of the smallest frame that can contain B^p and B^g, and the calculation formula is as follows:

$$x_1^c = \min(\hat{x}_1^p, x_1^g),\ \ y_1^c = \min(\hat{y}_1^p, y_1^g),\ \ x_2^c = \max(\hat{x}_2^p, x_2^g),\ \ y_2^c = \max(\hat{y}_2^p, y_2^g) \tag{4}$$

Finally, calculate the final loss LGIoU, where U is the area of the set of 2 frames, the formula is as follows:

$$L_{GIoU} = 1 - GIoU = 1 - \left(\frac{S^I}{S^P + S^g - S^I} - \frac{U}{S^c} \right) \tag{5}$$

3 Target Ship Ranging

3.1 CANNY Algorithm Detects the Horizon

The Canny algorithm usually processes grayscale images, so if the camera acquires a color image, it must first be grayscaled. To grayscale a color image is to perform a weighted average according to the sampled values of each channel of the image

As shown in Figure 1, the position of the horizon is calculated and drawn for the CANNY edge algorithm.

(a) Original image (b) Edge detection (c) Draw the horizon

Fig. 1. Horizon detection

3.2 Algorithm for Calculating Ship Distance Based on Horizon

Figure 2 shows the general scheme of problem setting. A camera mounted on a boat at a height of h captures the image of the object. Consider a pixel on the sea surface p represents the object captured by the camera. This pixel corresponds to the distance from the ship d. We should calculate the angle between the line of sight of the small sea area represented by the pixel p and the line connecting the camera and the center of the earth. Therefore, if we know the height of the camera h and the angle described above $\alpha+\varphi$, its field of view (FOV) only starts from a certain angle φ.

assuming that the position of the camera relative to the ship is constant, φ remains constant and can be measured.

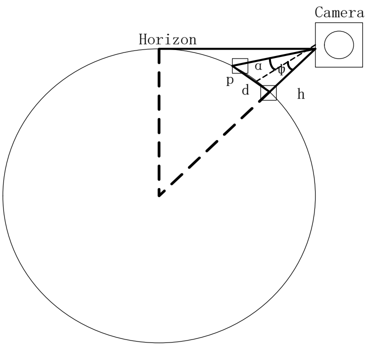

Fig. 2. Distance detection

4 Experimental Results

The experimental platform in this article is the PC-side Windows operating system, running memory 8GB, Intel Core i5-9300H CPU @ 2.40GHz*4 processor, and graphics card NVIDIA GTX1650 (video memory 4GB). The experimental training data set is SEASHIP (7000).

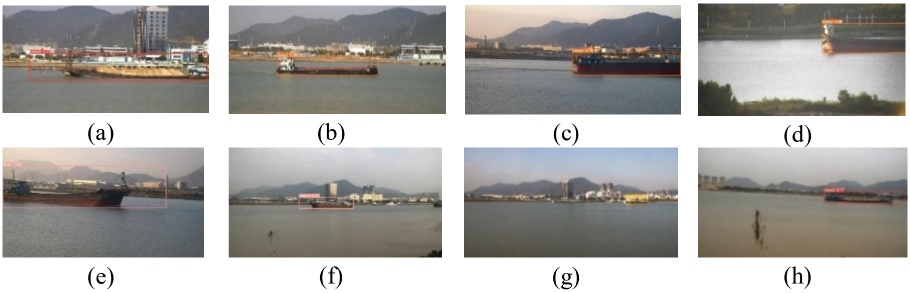

Fig. 3. Ship inspection results

Figure 3 shows the experimental results of ship category detection. During the detection process, the average calculation speed of the trained data model is less than 0.1 seconds, and the mAP can reach 94.3%.

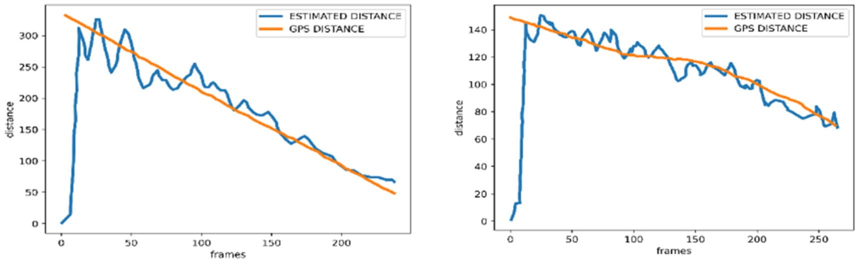

Fig. 4. Comparison Experiment 1 between Test Results and GPS Data

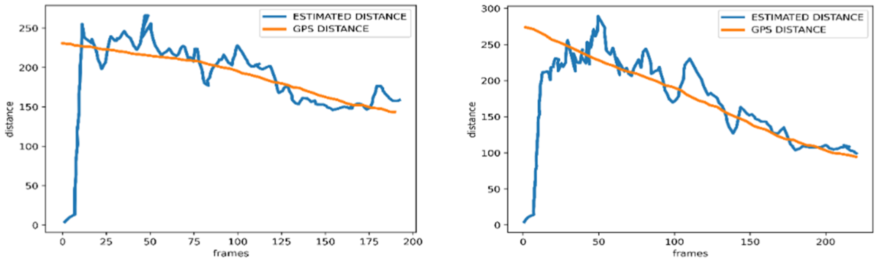

Fig. 5. Comparison Experiment 2 between Test Results and GPS Data

As shown in Figures 4 and 5, a comparison was made between the long-distance distance measurement of ship detection images and GPS positioning data. Among them, there were fluctuations in the overall data of image long-distance distance measurement compared to GPS distance measurement, with an average error of 8.2% of the actual distance, which meets the requirements of general maritime image distance measurement.

5 Conclusion

This article combines YOLOv5 training model with horizon detection ranging algorithm to classify ships sailing on the sea. Experimental results show that using YOLOv5 for ship classification detection can achieve mAP of 94.3%, and the error value of target ranging compared to GPS data is 8.2% of the actual distance, which meets the general requirements of sea image ranging. In the future, algorithm improvements and the introduction of neural networks will be used to reduce data errors and further improve data detection stability.

References

1. Ultralytics, YOLOv5 (2020). [2020－06－10]. https://github.com/ultralytics/yolov5.
2. GLADSTONE R, MSHE Y, BAREL A, et al. Distance estimation for marine vehicles using a monocular video camera2016 24th European Signal Processing Conference (EUSIPCO), IEEE, 2016.
3. PARK J, KIM J, SON N S. Passive target tracking of marine traffic ships using onboard monocular camera for unmanned surface vessel. Electronics Letters, 2015, 51(13):987-989.
4. CHO Y, PARK J, KANG M, et al. Autonomous detection and tracking of a surface ship using onboard monocular vision. 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence(URAI), IEEE, 2015.
5. BREITENSTEIN M D, REICHLIN F, LEIBE B, et al. Robust tracking-by-detection using a detector confidence particle filter. IEEE International Conference on Computer Vision. 2009.
6. COMANICIU D, RAMESH V, MEER P. Real-TimeTracking of Non-Rigid Objects Using Mean Shift. IEEE Conference on Computer Vision ＆ Pattern Recognition Cvpr.2002.
7. BAR－SHALOM Y, DAUM F, HUANG J. The probabilisticdata association filter. IEEE control systems, 2010, 29(6):82－100.
8. REID, D. An algorithm for tracking multiple targets. IEEE Transactions on Automatic Control, 1979, 24(6): 843－854.
9. Zhu Linlin, Geng Xun, Li Zheng, Liu Chun. Improving YOLOv5 with Attention Mechanism for Detecting Boulders from Planetary Images. Remote Sensing, 2021, 13(18):
10. Wang Ying, Hao Zhengyang, Zuo Fang, Pan Shanshan. A Fabric Defect Detection System Based Improved YOLOv5 Detector. Journal of Physics: Conference Series, 2021, 2010(1):