electronics

Article

Maritime Target Recognition and Location System Based on Lightweight Neural Network

Xiao Zhao, Zhenjia Chen *, Min Wang and Jingbo Wang

School of Information and Communication Engineering, Hainan University, Haikou 570228, China; 2021220854000198@hainanu.edu.cn (X.Z.); minwang@hainanu.edu.cn (M.W.); 21220854000176@hainanu.edu.cn (J.W.)
* Correspondence: zjchen@hainanu.edu.cn

Abstract: China's sea surface area is vast, the need to monitor the area is too large, and the traditional human monitoring method consumes a lot of manpower. Additionally, the monitoring period is too long; the monitoring efficiency is too low; and long-term human monitoring can easily cause visual fatigue, as well as missed detection and error detection. At present, the detection of sea surface targets generally includes infrared, visible light and other different means, which can obtain the image information of sea surface targets in different ways. The infrared target detection of the sea surface can be processed in the spatial domain and frequency domain, respectively, but the image resolution is not high in general, and the detection effect is not good because it is easily affected by weather. In this paper, we propose a maritime target detection method based on embedded vision. Based on visible video images, this paper realizes the rapid detection and recognition of sea surface targets. Clouds and waves in ocean images are filtered by adding an image preprocessing module. Compared with the traditional two-frame difference method, this algorithm has better detection capability for sea surface targets. Experiments were carried out in different weather conditions to detect moving ships at sea. By comparing the number of detection boxes and the detection accuracy, the accuracy of this method reaches 90.2 percent. By designing a single camera location algorithm for the marine environment, the world coordinate location of the marine target is realized. On this basis, the communication function is added to realize the intelligent monitoring of the sea surface without human intervention.

Keywords: embedded vision; maritime target recognition; target detection; target location

Citation: Zhao, X.; Chen, Z.; Wang, M.; Wang, J. Maritime Target Recognition and Location System Based on Lightweight Neural Network. *Electronics* **2023**, *12*, 3292. https://doi.org/10.3390/electronics12153292

Academic Editor: Stefanos Kollias

Received: 12 July 2023
Revised: 24 July 2023
Accepted: 26 July 2023
Published: 31 July 2023

1. Introduction

With the continuous progress of society, the sea area has gradually become a prominent area of human activity, and the importance of sea area control is increasing day by day. In recent years, marine exploration equipment has undergone revolutionary changes. With the development of computer vision, its embedded vision has also been well applied. With the development of artificial intelligence, more and more neural network methods have been applied for target recognition.

1.1. Related Works

Image filtering, which is to suppress the noise of the target image while retaining the details of the image as much as possible, is an indispensable operation in image preprocessing, and its processing effect will directly affect the effectiveness and reliability of the subsequent image processing and analysis. Wenkai proposed an image denoising method based on an edge feature fusion network [1]. The edge information of the image is extracted by an edge extraction network based on a Canny operator, adaptively assigning larger weights to relatively important edge information to enhance the edge details of the image to obtain a clear image with more edge information. Yibin Hu proposed an image denoising and fusion method based on the improved wavelet transform [2]. The noise-containing image is denoised using the wavelet thresholding method and the traditional mean value method to obtain two denoised images; then, the image is fused using the wavelet fusion method to obtain the final denoised image.

A convolutional neural network is a kind of feedforward neural network with convolutional computation and a deep structure. It is one of the representative algorithms of deep learning, and its algorithm can be used for target recognition. Yang Zheng et al. proposed edge-computing technology for real-time video stream analysis [3]. Zheng Zongsheng improved the application of YOLO v4 model in fish target detection [4]. Chen Peihao et al. proposed an object recognition system based on the YOLOV3 model [5]. Ma Zhijun proposed a research on target recognition based on multi-feature fusion [6]. The deployment of convolutional neural networks is mainly limited by its large network model size, high temporary result occupancy, and large floating point computation. If it is transplanted directly to embedded devices, it will be difficult to meet the requirements of real-time detection. Due to the insufficient computing power of some devices, its network model can not even run on embedded platforms. Based on these problems, it is particularly important to adopt some lightweight methods to effectively reduce model computation and model volume without changing the performance of existing deep neural network models. Yang Chenyi et al. proposed a lightweight neural network gesture recognition method for embedded platforms [7]. A lightweight mobile netv3-ssd life gesture recognition algorithm is designed by using a deep separable product instead of common convolution. Liang Huagang et al. proposed a lightweight expression recognition method combining improved convolutional neural networks and channel weighting, which simplified the complexity of the network, effectively reduced the network parameters, and increased the detection rate [8]. Xia Guang et al. proposed a small sample iris image segmentation method based on lightweight convolutional neural networks, which not only has significant performance advantages on small sample databases, but also has high segmentation accuracy on large sample databases [9].

Nowadays, the more accurate distance detection methods are binocular vision ranging and LIDAR ranging. Binocular vision ranging is difficult to calibrate, and LIDAR ranging consumes expensive costs. In order to solve these problems, some experts and scholars propose using the monocular vision method for target ranging. The use of the monocular vision method for ranging has the advantages of a low cost and easy calibration. However, the images acquired by monocular cameras lack depth information. Therefore, there is a need to improve the accuracy of monocular ranging. Y. Zha et al. proposed constructing a distance estimation neural network model. A compensation strategy was also proposed to improve the accuracy of long-range ranging [10]. Z. Xu et al. proposed a monocular ranging model without camera internal reference. The mapping model between the longitudinal pixel points of the image and the actual distance is fitted using a function [11]. Improved ArUco markers were proposed by Y. Wang et al. This method is able to achieve a relative error of less than 1 percent [12].

In this paper, we propose a maritime target detection method based on embedded vision. Based on visible video images, this paper realizes the rapid detection and recognition of sea surface targets. Clouds and waves in ocean images are filtered by adding an image preprocessing module. A single optical camera is used to acquire image data for maritime target detection. However, the visible image lacks depth information, resulting in the inability to determine the location of the target. Therefore, when a maritime target is detected, target ranging is performed using the center projection point of the target on the sea surface. The camera head is used to obtain the pitch and roll angles of the camera and establish a monocular vision model. Considering the influence of ship attitude on ranging, the monocular vision ranging model at sea under ship attitude transformation is established by acquiring the ship attitude in real time through the sensor.

1.2. Contributions

Based on the research of marine-information-sensing technology projects of high-precision shipborne intelligent terminals, this paper proposes a maritime target recognition and location system based on a lightweight neural network. By using a lightweight neural network to achieve the rapid identification of sea surface targets in the surrounding environment of ships, the detection and identification efficiency of the whole system is improved by optimizing the recognition function of the lightweight network and the algorithm of the detection module, so that the equipment can adapt to the complex sea surface environment and achieve the purpose of real-time monitoring of the sea surface environment around the ship. At the same time, we use our own algorithm to locate the identified target, and model the current sea environment to detect the sea environment in real time. This provides information support for subsequent ship collision avoidance and path planning, and relevant personnel can use it to monitor the sea surface area, which has very important research value and necessity. This method of surface vessel detection is suitable for shipborne or shore-based embedded terminals in complex marine environments.

2. Maritime Boat Identification

Various neural networks are run in the existing equipment through control variables to identify ship images on the sea surface, and the neural networks and equipment required for target identification are preliminarily screened by comparing detection rates, as shown in Table 1.

Table 1. The recognition rate of different neural networks in different devices.

System	Model	Identify Target	Build Option	Prediction Time (Seconds)
Pi 3B	YOLOv3-Tiny	Boat	None	12.2 (first frame), 9.8 (subsequent frames)
Pi 3B	YOLOv3	Boat	None	20.2 (first frame), 16.6 (subsequent frames)
Pi 3B	Darknet19	Boat	None	22.3 (first frame), 17.5 (subsequent frames)
Pi 3B	YOLOv3-Tiny	Boat	NNPACK = 1, ARM-NEON = 1, NNPACK-FAST = 1	2.2 (first frame), 1.6 (subsequent frames)
Pi 3B	YOLOv3	Boat	NNPACK = 1, ARM-NEON = 1, NNPACK-FAST = 1	3.6 (first frame), 2.3 (subsequent frames)
Pi 3B	Darknet19	Boat	NNPACK = 1, ARM-NEON = 1, NNPACK-FAST = 1	3.9 (first frame), 2.7 (subsequent frames)
Pi 4B	YOLOv3-Tiny	Boat	NNPACK = 1, ARM-NEON = 1, NNPACK-FAST = 1	0.8 (first frame), 0.5 (subsequent frames)
Pi 4B	YOLOv3	Boat	NNPACK = 1, ARM-NEON = 1, NNPACK-FAST = 1	2.5 (first frame), 1.9 (subsequent frames)
Pi 4B	Darknet19	Boat	NNPACK = 1, ARM-NEON = 1, NNPACK-FAST = 1	2.9 (first frame), 2.1 (subsequent frames)

The YOLO-tiny model accelerated by the NNPACK acceleration package has a good recognition rate on the embedded device tested in this paper. In this paper, the YOLO-tiny model is used to identify ships at sea. The YOLO-tiny module is used to detect, identify, classify, and store ships on the sea surface through the ip camera directly connected to the embedded device. When different types of vessels are detected, they are classified according to the type of vessel and stored in a specific folder. In order to avoid the situation of missing detection due to a special type of ship or weather factors, a moving target detection module is added in this paper, and the existing algorithm is improved to make it suitable for the marine environment. Maritime vessel identification models have been deployed to embedded devices, making practical applications of maritime vessel detection possible.

In order to prevent the data storage from being saturated too quickly due to the fast detection rate, this paper adds a storage judgment module, which can not only identify the ship, but also store the identification results in the background for subsequent human verification. Combining edge computing technology and embedded vision technology, ship environment detection and marine target recognition are realized. The specific process of ship identification at sea is shown in Figure 1.

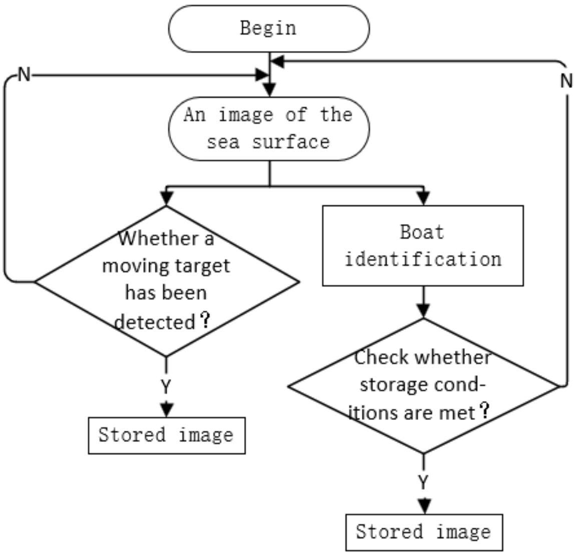

Figure 1. Maritime Boat Recognition Model.

Automated Sampling

The traditional moving target detection mostly adopts a two-frame difference method, which takes the previous frame of the video stream as the background and compares it with the next frame to determine whether there is a moving target in the field of view. The two-frame difference method is a method to obtain the contours of moving objects by differentiating two consecutive frames of video image sequence. The algorithm has the characteristics of a fast, adaptive dynamic environment; is insensitive to scene light; and is suitable for low-illumination scenes.

When there is abnormal target motion in the monitoring scene, there will be an obvious difference between two adjacent images, subtracting two frames, obtaining the absolute value of the pixel difference of the corresponding position of the image and judging whether it is greater than a certain threshold. Then, the motion characteristics of video or image sequences are analyzed. The mathematical formula is as follows:

$$D(x,y) = \begin{cases} 1 & |I(t) - I(t-1)| > T \\ 0 & others \end{cases} \tag{1}$$

D(x,y) is the difference image between two consecutive frames; *I(t)* and *I(t−1)* are the images at *t* and *t*−1 time, respectively; and *T* is the threshold value selected when the difference image is binarized, *D(x,y)* = 1 for the foreground and *D(x,y)* = 0 for the background.

Based on this, this module changes the logic relation and adds the image and pre-processing, making it suitable for complex environments on the sea. Before detection, the contour of the image is enhanced so that the object and background in the image can be easily distinguished. The captured sea surface images are grayed and blurred using Gaussian filters to prevent noise caused by natural vibrations, light changes, or the camera itself. The test results show that the impact of cloud and sea surface ripples on the image cannot be completely eliminated after the median filter processing, and the traditional image processing method directly processing the original image cannot eliminate the impact, as shown in Figure 2.

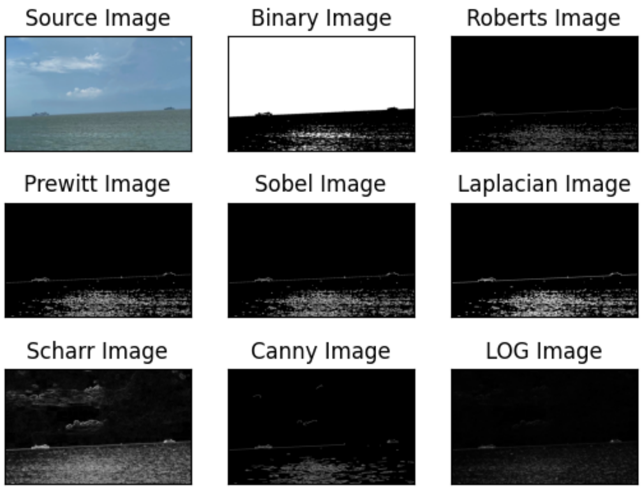

Figure 2. Maritime Moving Target Detection Process.

In order to prevent the influence of sea ripples and cloud changes on the detection results, the image gray processing is a good starting point. In this paper, a grayscale processing algorithm suitable for the sea surface environment is designed, and a series of sea surface samples are tested to adjust parameters and improve the algorithm. The algorithm is as follows:

$$f(x2) = (140/255) * f(x1) + 100 \tag{2}$$

where $f(x1)$ is the gray value of the original image and $f(x2)$ is the gray value after processing. The grayscale histogram of the image processed by the algorithm is shown in Figure 3.

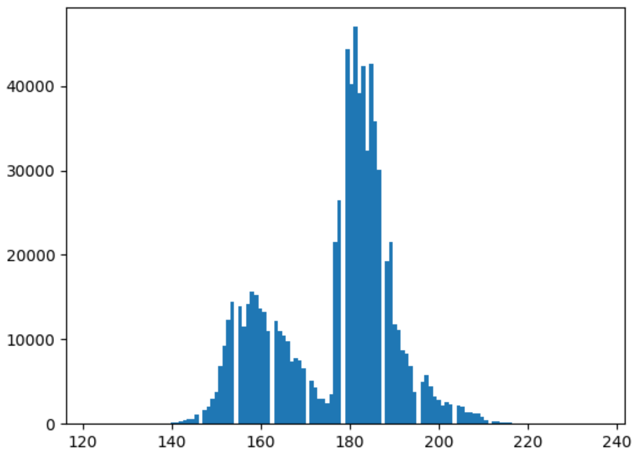

Figure 3. The Gray Histogram of the Image is Processed by the Algorithm.

Through a series of sea surface samples, the algorithm test found that objects without sea ripples and clouds in the image could be extracted by converting the image value to 100–200, and then the median filter, Gaussian filter, and Canny operator's double threshold (50–100) detection are used, respectively, eliminating the influence of the sea surface environment on the operation results. The gray contrast diagram after algorithm processing is shown in Figure 4.

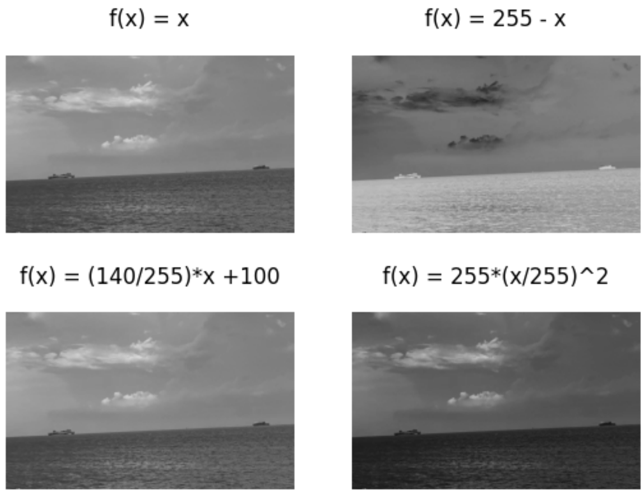

Figure 4. The Gray Contrast Diagram After Algorithm Processing.

The median value of the processed image was binarized to get Figure 5. After the image preprocessing is completed, the difference of two consecutive frames is calculated and then the difference map is processed by expansion corrosion to connect the images, so as to better find the moving objects at sea.

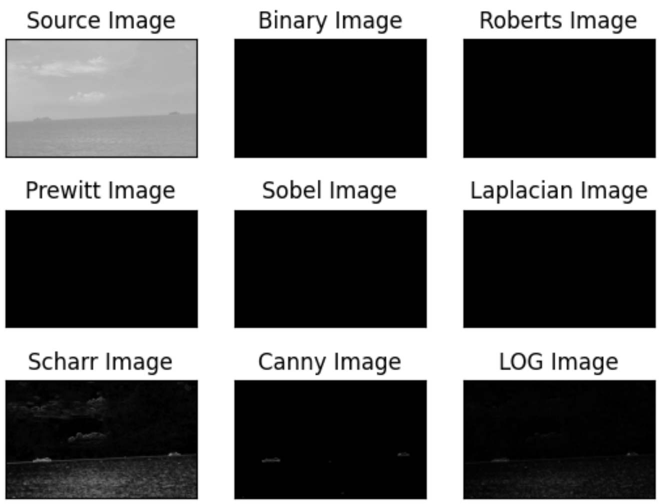

Figure 5. The Image After Binarization.

In order to increase the applicability of the module on the sea surface, a double threshold is added. When the area of the picture frame is less than the minimum threshold or the size of one frame exceeds the maximum threshold, it is judged that no moving target is found to prevent misjudgment caused by camera movement or ship shaking. Finally, the sensitivity and threshold range are set according to the specific location of the embedded device and the residual parameters of the camera, so as to realize the monitoring of moving objects on the sea surface. The flow chart of sea moving object detection is shown in Figure 6.

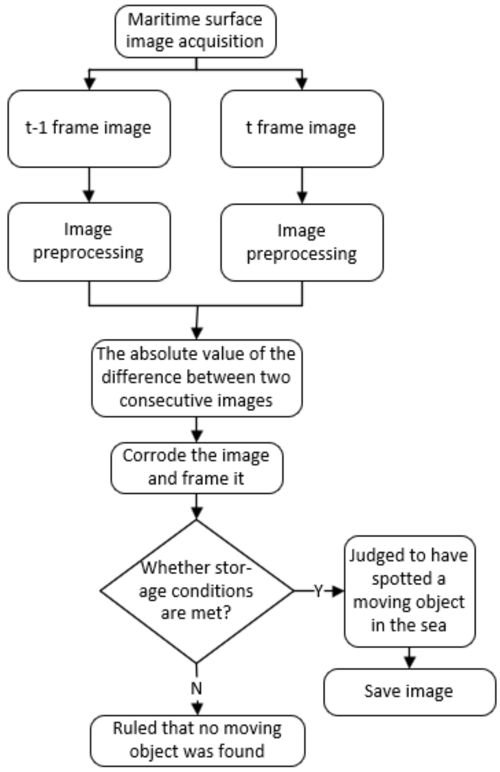

Figure 6. Maritime Moving Target Detection Flowchart.

3. Maritime Boat Identification

The marine boat recognition in this article uses the YOLO algorithm, whose full name is You Only Look Once; it only requires one CNN operation. Compared with the R-CNN algorithm based on the region proposal (R-CNN, Fast R-CNN, Faster R-CNN), the algorithm is faster and has a very high accuracy on medium and small objects.

End-to-end detection is used by yolov3-tiny in predicting pictures, dividing the entire picture into S*S areas, and if the center of an object falls on a certain area, the corresponding network will detect it. For each network, there is a bounding box, which is the prediction area. There are four coordinate parameters for each prediction, xi, yi, width and height tw, th in the upper left corner, and a confidence level. This confidence is the product of logistic regression. The confidence level judges whether the bounding box will be ignored. If it is not ignored, the logistic regression of multi-label classification will be performed to label it. The confidence formula is as formula 3: Pr(Object) = 1 if the target center is in the grid, otherwise Pr(Object) = 0:

$$confidence = \Pr(\text{object}) * \text{IOU}_{\text{pred}}^{\text{truth}} \tag{3}$$

When detecting the target, the class conditional probability predicted by each grid is multiplied by the confidence information predicted by the bounding box to obtain the class-specific confidence score of each bounding box. The formula is as follows:

$$\begin{aligned} &\Pr(\text{Class}_i \mid \text{Object}) * \Pr(\text{Object}) \ * \ \text{IOU}_{\text{pred}}^{\text{truth}} \\ &= \Pr(\text{Class}_i) * \text{IOU}_{\text{pred}}^{\text{truth}} \end{aligned} \tag{4}$$

Through a series of comparative tests, this paper uses yolo-tiny to realize object recognition. Yolo-tiny has only 24 network layers in total, which is much less than yolo3's 107 layers. It has only two yolo layers, yolo16 and yolo23, respectively. The sizes are 13×13 and 26×26, respectively. The network is simple and the amount of computation is small, and it can be run on mobiles or devices.

A large amount of image data collected through self-sampling is used as the dataset for this paper. Deploying the maritime moving target detection module to the embedded device in an offshore environment, a large number of images of maritime moving objects are acquired in rainy day, foggy day, sunny day, and dusk, respectively, and 2000 images, each containing large target vessels and small target vessels, are filtered as the dataset under each condition. After manually labeling the images containing vessels, the images are put into a common workstation to train the model, and the trained weight files are ported to the embedded device to complete the model deployment. The test results show that the model has a good recognition rate for ships at sea, and the recognition rate can reach 2 s per frame required for actual detection.

After identification, it is necessary to conduct subsequent positioning processing on its ship, so there cannot be only one identification result map, and the identified picture also needs to be stored in the device for subsequent human verification. In this paper, by modifying the identified data processing code, the identified ships are cut according to the size of the identification box, and the longitude and latitude of the ships are named and stored in a specific folder according to the identification category, current time, and subsequent modules. After identification, the size of the identification box, as well as the x and y values, are sent to the positioning module in the form of sockets to locate and track the identified ships. The effect diagram after recognition is shown in Figure 7.

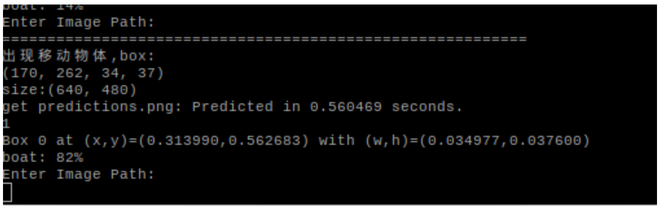

Figure 7. The Recognition Result is Displayed in the Background.

4. Sea Target Ranging Based on Monocular Vision Model

In computer vision, restoring the 3D scene in an image requires the construction of a geometric model of the camera imaging to determine. Using only one camera is a monocular vision model. Compared to binocular vision, monocular vision requires less costs and less computational resources to measure distance. To build a monocular vision model, it is necessary to obtain the internal and external parameters of the camera through camera calibration.

4.1. Obtain Camera Parameters Based on Camera Calibration

The behavioral intent of a target cannot be determined simply by knowing what targets are identified at a given location by the shipboard terminal. By target localization, the location of the target and its travel trajectory can be determined. In this paper, we establish a monocular visual ranging model to perform target ranging.

Camera calibration is a method to obtain the internal and external reference matrices by calculation. The result of camera calibration affects the accuracy of target localization. In order to be able to confirm the internal reference parameters of the camera, this paper uses the Zhang Zhengyou calibration method to calibrate the camera. The Zhang Zhengyou calibration method is one of the most commonly used camera calibration methods in the field of computer vision. This method requires the acquisition of black and white checkerboard grid images in different poses. The internal and external reference matrices of the camera are calculated by the single response matrix. This method is well suited for camera calibration because of its robustness, high accuracy, and simplicity.

The Zhang Zhengyou calibration method requires 10–20 images of the checkerboard grid in different poses. In order to get better calibration results. In this paper, 20 images are selected for calibration. Each image is taken from the calibration board at different rotation angles, as shown in Figure 8. The size of the black and white checkerboard grid is 86, and the size of each grid is 27 (mm). After obtaining the corner point information in the images, the internal and external parameters are calculated by the single response matrix.

Figure 8. Black and White Checkerboard Grid Images at Different Rotation Angles.

The focal length (*fx*,*fy*) and the center point coordinates (*u*,*v*) of the camera's pixels were obtained by calibration, as shown in the Table 2 . The required internal parameters are determined and the monocular visual ranging model can be built.

Table 2. Camera internal reference parameters.

Parameters	*fx*	*fy*	*u*	*v*
pixcel	2181	2177	1265	702

4.2. Target Localization Based on Monocular Vision Ranging Model

After obtaining the internal parameters of the camera and after camera calibration, the monocular visual ranging model can be constructed. The sea target ranging is mainly for targets on the sea surface, such as ships. To construct a monocular vision ranging model at sea, it is also needed to obtain the pitch angle of the camera, as well as the yaw angle. The distance to be measured includes two aspects, the lateral distance Dh and the vertical distance Dv.

The modeling of longitudinal distance is similar to small-aperture imaging, such as the target ranging model shown in Figure 9. Point B and the optical axis projection line are in the same line. At this time, the longitudinal distance is the distance from the target point B to A, where the projection point of point B in the image is b. f is the focal length. The pitch angle of the camera is α. H is the distance from the ship to the sea surface, which is OA, and $\angle\beta = \angle bOc$. The longitudinal distance Dh can be calculated by the above conditions:

$$Dh = H/tan(\alpha + arctan(bc/f)) \tag{5}$$

When the target is not on the optical axis projection line, it is necessary to consider the influence of the lateral distance to the ranging model, where the projection point of the target point T is t. Based on the similar triangle principle, the lateral distance Dv is calculated:

$$Dv = TB = OB * bt/Ob \tag{6}$$

As the camera is deployed to the ship, the attitude of the camera is affected by the sea surface heave. The height H of the camera is affected when the ship has pitch and roll angles. Therefore, the monocular visual ranging in this paper will combine the attitude of the observation vessel and the attitude of the camera. In order to be able to obtain the attitude of the observation vessel, the IMU module is deployed to the observation vessel and the camera. The monocular visual ranging model built on the basis of the ship attitude transformation is shown in Figure 10, where $\angle\gamma$ is the pitch angle and $\angle\eta$ is the cross-roll angle. Since the camera position on the ship is fixed, IG, GK, and $\angle IGK$ are known. When the pitch and cross-roll angles of the observation ship are obtained, the height H of the camera to the sea surface under the ship attitude transformation can be obtained:

$$H = OA \tag{7}$$
$$= OE * cos(\eta) \tag{8}$$
$$= OF * cos(\gamma) * cos(\eta) \tag{9}$$
$$OF = (OK + IK * sin(\gamma))/sin(\angle IFK) \tag{10}$$

At this point, the real-time altitude H can be calculated according to the formula. Based on the height H, the target distance can be calculated using a monocular visual ranging model.

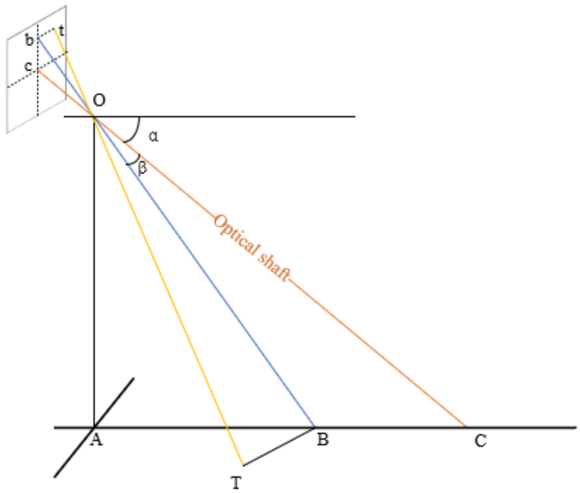

Figure 9. Monocular Visual Ranging Model.

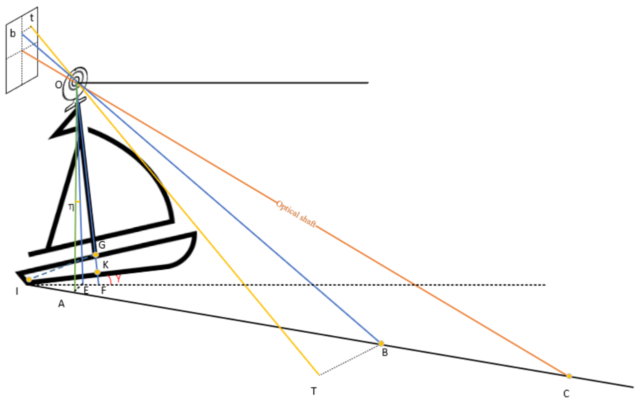

Figure 10. Monocular Visual Ranging Model Based on Ship Attitude.

5. Data Storage and Postback

5.1. Data Storage

For data storage, the same ship may stay in the camera field of view for a long time. In order to prevent data storage from becoming saturated too quickly, a storage judgment module is added in this paper. The identification frame and the objects in the frame are used as the criteria to determine whether to continue to store the next picture. The identification process does not affect the identification, positioning, and data return.

5.2. Data Postback

For data return, when the ship is near the shore docking point, the device is connected to the WiFi at the same time for data transmission backup. The device automatically finds the storage server. Networks are classified into LAN cases and non-LAN cases. By default, the data module is on the LAN. The fixed IP address of the storage server is pinged directly. If the ping succeeds, a connection with the storage server is set up through SFTP.

If the storage server cannot be pinged, the system switches to the non-LAN mode and transmits data through the intranet connected to the file storage server through the domain name. If the connection fails, it is necessary to wait and connect again. The process is repeated until the connection is successful. If the connection is not successful within one minute, the connection is abandoned and the error log is saved to the specified folder. The error log contains the current device IP, connection failure time, number of connection attempts, and error codes returned by the storage server. The data transmission flow chart is shown in Figure 11. In order to make data transmission more stable and efficient, compression is adopted before transmission, and the compressed package of each compression is uniquely identified by the current time name of the storage server.

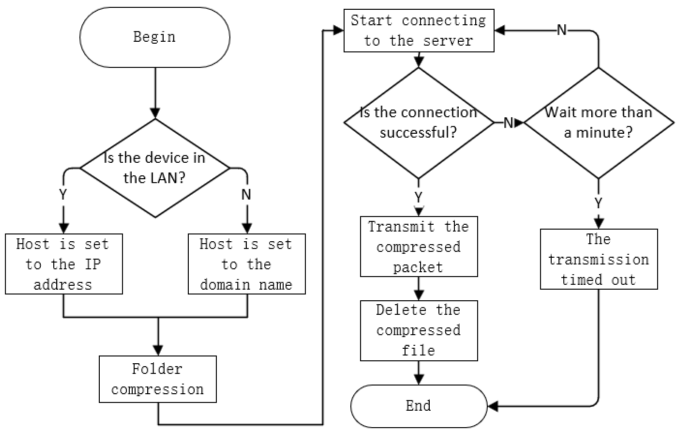

Figure 11. Data Transmission Flowchart.

6. Experimental Result

6.1. Automatic Sampling Module Test Results

In the actual test of the marine environment, two pictures of ships on the sea are taken in succession, and after preprocessing, two frames of difference calculation are carried out to get the difference map. Through frame processing, the specific relative position of the moving target can be found, which also provides conditions for subsequent judgment. The processed picture is shown in Figure 12. By processing the picture frame of the moving area, the specific relative position of the moving target can be found, which also provides conditions for subsequent judgment. The picture after the picture frame processing is shown in Figure 13.

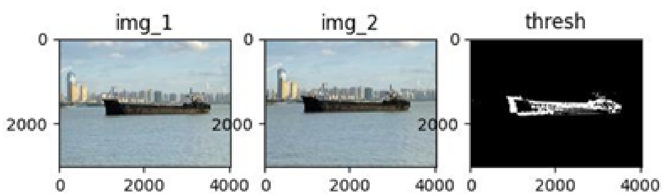

Figure 12. Difference Graph Obtained by Two-Frame Difference Method.

Figure 13. Difference Graph Obtained by Two-Frame Difference Method.

Through image pre-processing, the moving object detection module eliminates the effects of clouds, raindrops, and waves on the experimental results. In extreme weather, this module also has a good detection effect. The module also has a good detection effect on the sea surface on days with rainstorms, as shown in Figure 14.

Figure 14. Sea Surface Moving Target Detection on Stormy Day.

The algorithm provides better detection of sea targets compared to the traditional two-frame differencing method. Detection experiments were conducted on moving vessels on the sea under different weather conditions. Two-thousand consecutive two-frame images containing moving vessels were measured on sunny, rainy, and foggy days, respectively, and the accuracy of the algorithm was derived by comparing the number of detection frames with the actual vessels and the accuracy difference. By testing a large number of consecutive two-frame images containing moving vessels on the sea surface, the accuracy of our method reaches 90.2 percent, which is much higher than that of the traditional method in the sea surface environment. The test results are shown in Figure 15.

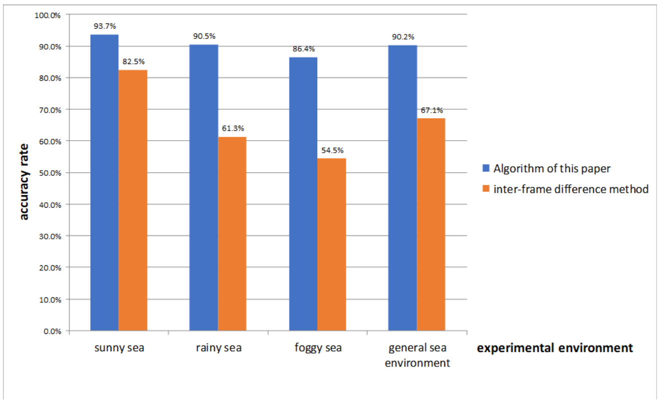

Figure 15. Moving Target Detection Accuracy and Comparison Results.

6.2. Maritime Boat Identification Results

By training the YOLO network with a large amount of sea surface ship image data, the model can well identify some common sea ships, and the recognition results of common ships are shown in Figure 16.

Figure 16. Boat Identification Result.

In order to test the measured effect of the recognition module and the minimum requirement of the equipment, a camera with less accuracy was used to test the equipment during the rainstorm period. The following Figure 17 is the actual measurement of the identification module on the embedded device.

Figure 17. Boat Identification with Bad Equipment Result.

6.3. Boat Positioning Result

The position relationship and distance between the target and the monitoring ship can be obtained by designing an offshore target location algorithm, and the accurate longitude and latitude coordinates of the identified target can be obtained by obtaining the longitude and latitude of the monitoring ship. In order to verify the feasibility of this method, ranging simulation experiments are conducted on the ground. As a test, a person was used as a target sample, positioned, respectively, at 3 m, 4 m, 5 m, 7 m, 8 m, 10 m, 11 m, 12 m, and 15 m. The absolute value of the error between the measured value and the true value was divided by the true value to arrive at the percentage error. The results of the ranging experiment are shown in Figure 18.

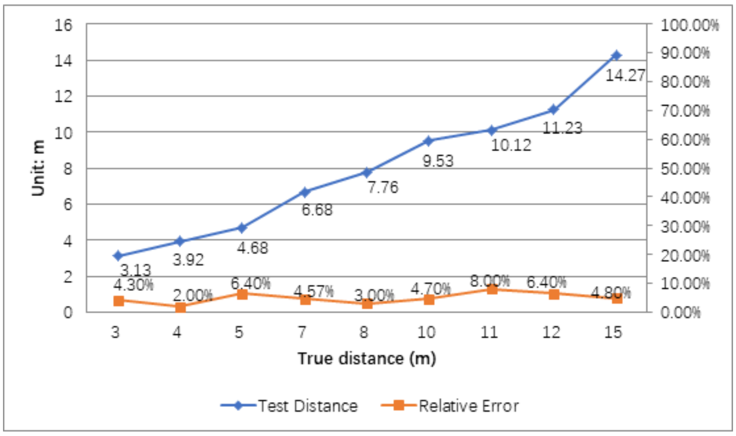

Figure 18. Ship Positioning Result.

The following Figure 19 lists the measured results of the localization algorithm of this paper in a sea surface environment. When the target is farther away from the camera, the distance test error will be larger. For short distances, target ranging can achieve more accurate localization. Therefore, in the future, many experiments will need to be carried out for monocular visual ranging, and a compensation network is proposed to improve the accuracy of this monocular ranging model.

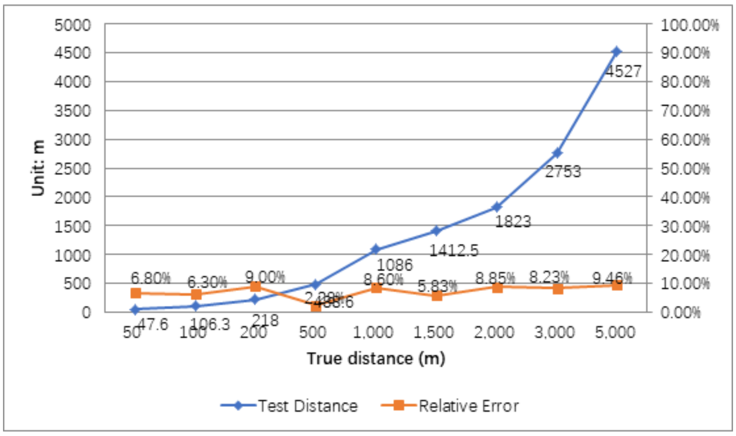

Figure 19. Ship Positioning Result at Different Distances.

The outdoor test target positioning results are shown in Figure 20. By building an embedded device to the bow of the ship, you can get the relative position and distance between the target captured by the camera and the ship after acquiring the attitude of the ship.

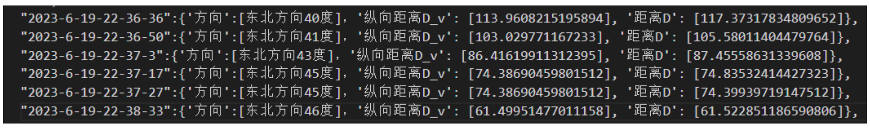
"2023-6-19-22-36-36":{'方向':[东北方向40度], '纵向距离D_v': [113.9608215195894], '距离D': [117.37317834809652]},
"2023-6-19-22-36-50":{'方向':[东北方向41度], '纵向距离D_v': [103.029771167233], '距离D': [105.58011404479764]},
"2023-6-19-22-37-3":{'方向':[东北方向43度], '纵向距离D_v': [86.41619911312395], '距离D': [87.45558631339608]},
"2023-6-19-22-37-17":{'方向':[东北方向45度], '纵向距离D_v': [74.38690459801512], '距离D': [74.83532414427323]},
"2023-6-19-22-37-27":{'方向':[东北方向45度], '纵向距离D_v': [74.38690459801512], '距离D': [74.39939719147512]},
"2023-6-19-22-38-33":{'方向':[东北方向46度], '纵向距离D_v': [61.49951477011158], '距离D': [61.522851186590806]},

Figure 20. Ship Positioning Result.

7. Conclusions

In this paper, lightweight neural networks are used to achieve the rapid identification of sea surface targets in the surrounding environment of ships, and the detection and identification efficiency of the whole system is improved by optimizing the recognition function of lightweight networks and the algorithm of detection modules so that the equipment can adapt to the complex sea surface environment and achieve the purpose of real-time monitoring of the sea surface environment around ships. At the same time, the identified target is located using its own algorithm, and the current sea environment is modeled to detect the sea environment in real time. After determining the longitude and latitude position of the identified target, the identified ship is named according to the identification type, time, and the longitude and latitude coordinates of the target to increase the functionality of the system. This model pays more attention to applications in actual engineering environments. It has good portability and timeliness, and has good practical application effects in the sea environment. In the future, we will continue to optimize the identification network to speed up the identification rate and improve the efficiency of the equipment.

Author Contributions: Formal analysis, Z.C. and J.W.; Data curation, M.W.; Project administration, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Hainan Province Science and Technology Special Fun (ZDYF2022GXJS012).

Data Availability Statement: Not Applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kai, W.E.N.; Juan, J.I.; Xiao, X.U.E.; Rujin, H.E. Image denoising method based on edge feature fusion network. *Microelectron. Comput.* **2023**, *40*, 25–32. [CrossRef]
2. Hu, Y.; Zhou, X.; Shen, A.; Hu, Y. Image denoising and fusion based on improved wavelet transform. *J. Qiqihar Univ. Nat. Sci. Ed.* **2023**, *39*, 33–38.
3. Yang, Z.; He, X.; Wu, J.; Wang, X.; Zhao, Y. Edge computing technology for real-time video streaming analysis. *China Sci. Inf. Sci.* ***2022*** *52*, 1–53.
4. Zheng, Z.; Li, Y.; Lu, P.; Zou, G.; Wang, Z. Application of improved YOLO v4 model in fish target detection. *Mod. Fish.* **2022**, *49*, 82–88.
5. Chen, P.; Han, Y.; Zhong, H.; Yu, Z.; Liu, X. Research on Object Recognition System Based on YOLOV3 Model. *Electron. Test.* **2021**, *21*, 66–67.
6. Ma, Z.; Yang, J.; Sun, Y. Research on object recognition based on multi-feature fusion. *Sci. Technol. Innov.* **2021**, *16* , 33–35.
7. Yang, C.; He, Y.; Zhao, J.; Li, G. A lightweight neural network gesture recognition method for embedded platforms. *Strong Laser Part. Beam* **2022**, *34*, 142–150.
8. Liang, H.; Bo, Y.; Lei, Y.; Yu, Z.; Liu, L. Lightweight Expression Recognition Based on Improved Convolutional Neural Network and channel Weighting. *Chin. J. Image Graph.* **2022**, *27*, 3491–3502.
9. Huo, G.; Lin, D.; Liu, Y.; Zhu, X.; Yuan, M. Lightweight convolution based neural network the small sample of the iris image segmentation. *J. Jilin Univ. Sci. Ed.* **2023**, *21*, 583–591.
10. Zha, Y.; Li, F.; Li, Y.; Hong, L.; Ling, Q. MDE-Net: Monocular Distance measurement based on detected license plates. In Proceedings of the 2022 41st Chinese Control Conference (CCC), Hefei, China, 25–27 July 2022; pp. 5453–5457.
11. Xu, Z.; Zhou, Z.; Wang, L. Research on Monocular Vision Distance Measurement Algorithm Based on Reference Target. In Proceedings of the 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), Wuxi, China, 19–21 July 2019; pp. 192–197. [CrossRef]
12. Wang, Y.; Zheng, Z.; Su, Z.; Yang, G.; Wang, Z.; Luo, Y. An Improved ArUco Marker for Monocular Vision Ranging. In Proceedings of the 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 22–24 August 2020; pp. 2915–2919. [CrossRef]