International Core Journal of Engineering Volume 7 Issue 9, 2021
ISSN: 2414-1895 DOI: 10.6919/ICJE.202109_7(9).0075

# Research on Small Target Ship Detection in Infrared Image based on YOLOV3

Bai Yu*, Ma Bo

Southwest University for Nationalities, Chengdu 610000, Sichuan, China.

## Abstract

**In target detection, the performance of small objects is far from satisfactory, although impressive results have been achieved for large/medium-sized objects in the large-scale detection benchmark. The reason is that small objects lack detailed enough information about their appearance to distinguish them from the background or similar objects. Remote detection of small target ships is of great significance for shipping safety and military application. This project is a sub-project of the horizontal project of ship identification, for which ships need to be found in a distance. Different from other target detection projects, this project focuses on the processing of data sets. In order to solve the problem of small target detection, the project has carried out high-precision image generation processing on the data set in the preliminary work. Using SRGAN network, the small fuzzy image can be upsampled into the fine super resolution image, and the detailed information can be recovered for more accurate detection. Aiming at ship small targets detection in infrared image recognition rate is low and the accuracy is not high, I have adopted the most mature and best YOLOv3 target detection algorithm, based on the optimized, and the infrared image can effectively avoid the night fog weather influence on testing results, can more accurately determine the ship position, has important application in real life.**

## Keywords

**Small Target Detection; Infrared Image; SRGAN; YOLOv3.**

## 1. Introduction

Maritime safety accidents have been greatly reduced by the navigation of ships equipped with modern technology. However, this is far from enough. According to the investigation of maritime accidents, human factors account for a high proportion. Statistics show that 90% of all collisions are caused by human error. In reducing human error and ensuring collision avoidance and navigation safety, the information and accuracy of data acquisition at the front end of the control system play a decisive role. Infrared target detection and tracking technology is a key technology that infrared imaging system can play an important role in ship navigation and port ship monitoring. This paper studies this key technology. The research of this topic belongs to the relatively new research field both at home and abroad and has important theoretical significance and application value.

## 2. Research status at home and abroad

### 2.1 Research status of ship target detection technology based on infrared image:

#### 2.1.1 Methods based on mathematical morphology.

Dong et al. [4] proposed a method of infrared ship segmentation using mathematical morphology. Jian-Nan et al. [5] further proposed a small ship target detection algorithm based on morphology and image entropy under the background of sea and sky. The author and his project team proposed a target detection algorithm based on morphological reconstruction in infrared image of inland river ships.

### 2.1.2 A method based on wavelet analysis.

Robin et al. [6] proposed a ship target detection method using biorthogonal spline wavelet in simple sea-sky background. Neves et al. [8] proposed a wavelet watershed algorithm for automatic infrared ship target segmentation. Wei et al. [9] proposed using multi-resolution wavelet transform to detect sea antennas in the background of sea and sky, and then using related wavelet energy synthesis to detect ships. Yu-qiu et al. [10] proposed a background suppression technology based on wavelet transform to detect ship targets. Zhou et al. [11] proposed a multi-frame image fusion algorithm combining wavelet transform and morphological model to detect ship targets. King site etc. [12] put forward a kind of based on wavelet and fractal automatic infrared small target detection algorithm, first of all to the original image wavelet decomposition, the calculation of the general picture of low-frequency image fractal parameters, using fractal parameters change in different texture image features, detecting the position of the antenna, water in the water near the antenna using gray characteristic of the infrared target in the search, Finally determine the location of the target.

### 2.1.3 Other methods.

Li et al. [13] proposed a sequential Monte Carlo based infrared ship target detection algorithm 0; Yang et al. [14] proposed an adaptive Butterworth high-pass filter to detect small targets in the sky and sea background. Li Hansong et al. [15], aiming at the detection problem of small targets in infrared ships with low signal-to-noise ratio, proposed the detection method of target segmentation in adjacent rows and columns of the image below the Tianwater line by respectively difference combined with Kapur maximum entropy, and the detection method of target segmentation in time domain by using template smoothing combined with Kapur maximum entropy. On the basis of analyzing the existing segmentation methods, Zhang Fang et al. proposed a segmentation method that uses the salient features of the target to define the segmentation region, which enables the threshold method based on inter-class variance to be applied. At the same time, considering the different gray distribution of the hull and engine in the infrared image, the local segmentation criterion is defined. Liu Yanwu et al. [16] adopted neighborhood average and pseudo median filtering methods for preprocessing and simulation of infrared images of Marine ship targets. He Youjin et al. [17] specifically analyzed the advantages and disadvantages of six edge detection operators, such as Roberts operator, Prewitt operator, Sobel operator, Log operator, Kirsch operator and Canny operator, in detecting infrared images of ships, and concluded that under noise interference, Canny operator is easier to deal with ship infrared image conclusion. Yue-huan wang etc. [27] is proposed based on a multi-resolution pay attention to the mechanism of infrared ship image detection method, this method is to use the note mechanism to reduce the amount of data to be processed, and will pay attention to the process is divided into "advance notice" and "note" two stages, nonlinear sampling model is adopted at the same time, to reduce the advance notice resolution at the same time, make the method can meet the target size changes. During detection, the thermal region of the ship engine or chimney in the infrared ship image is taken as the "pre-attention" feature. The attention is firstly guided to the region of interest, and then the waterline feature is detected within the region of interest. Zhang Tianxu and Feng et al. [18,19] proposed a local recursive segmentation method of two-dimensional maximum inter-class variance based on particle swarm optimization. Kaiyu and Xu Kaiyu et al. [20,21] proposed the detection algorithm of ship targets under the background of sea and sky based on artificial neural network. Zhu Qunying [22] respectively used inter-frame difference and background difference technology to detect ship targets. Miao Dechao [25] proposed the background suppression and small ship target detection algorithm based on multi-direction gradient. Tao Wenbing et al. [26] proposed an infrared target segmentation method for ships on the sea surface based on mean shift filtering and spectral classification.

The above moving ship target detection based on infrared image can only use the existing traditional algorithm when the target is very close and the signal-to-noise ratio is high. However, in the far-sea scene, the existing algorithm cannot detect the ship's target.

 

### 2.2 Research Status of GaN Network in Small Object Detection:

In recent years, GaN has developed rapidly in the field of image generation and data enhancement. In 2018, Bai Y et al. [1] applied GaN network to fuzzy small face detection for the first time, and generated clear high-resolution faces directly from fuzzy low resolution faces by using GaN. Face detection is then carried out. Yancheng Bai et al. [2] proposed multi-task generative adversal networks, which obtained super-resolution images of small objects through the generator of GaN network. During the training process, classification and regression losses in the discriminator of GaN network were propagated back to the generator, enabling the generator to recover more details.

## 3. Research significance

Ship is the important carrier of waterway shipping and the core link of Marine economy. As 90% of China's import trade is completed by water transportation, it is very important to strengthen the supervision of water transportation. There are three key steps in the intelligent ship monitoring system based on video analysis: detecting the ship position, obtaining the clear image of the target ship, and analyzing the image to obtain the ship information (such as ship number, ship type, etc.). In this paper, the research content is the first step to complete the intelligent monitoring system, the realization of ship target detection, to improve the efficiency of cameras to search for the surface of the ship, the focal length of the camera when looking for a ship cannot too big, it will lead to smaller ships in the picture, and small targets detection in traditional target detection algorithm performance is much less than the target object in the large/performance testing. Therefore, it is very important to improve the detection of small target ships for the performance of the whole system. Infrared image can solve the problem of target detection of ships in bad weather such as night, rainy day and fog. In addition, the ship target information obtained through infrared image can be integrated with the information provided by other navigation equipment, so as to provide reliable data for the ship navigation.

## 4. Innovations and features

(1) At present, the target detection of ships is mostly focused on medium and large targets, which is not suitable for target search and detection in large sea area under the premise that the camera is close to the ship or the focal length of the camera is large. In this project, target detection is carried out on small target ships, which can greatly improve the performance efficiency of the system.

(2) At present, many infrared images of ships are based on classical image algorithms, and most of them are large target detection. This project adopts deep learning technology, combined with GaN network and target detection technology to solve the problem of ship target detection under bad light conditions such as bad weather and night.

(3) At present, the algorithm of this project has not been applied too much in the field of small target infrared image ship detection. Moreover, after summarization and exploration by predecessors, it has been proved that our algorithm of this project has achieved very good results in other similar fields of target detection. Therefore, the research of this project is very innovative.

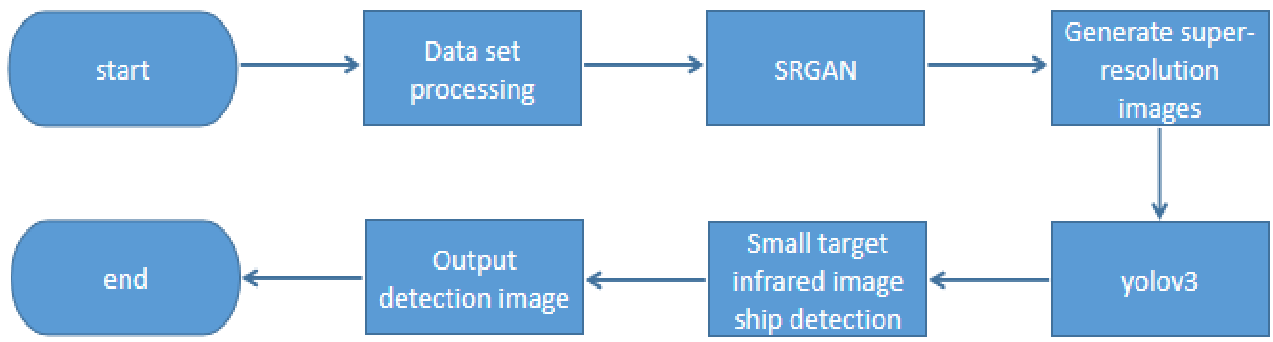


Fig. 1 Project flow chart

 

# 5. Research contents and methods

## 5.1 Project process

The operation process of the whole project is roughly divided into 8 small modules, mainly including data set processing, SRGAN, high-resolution image generation, YOLOV3 and small target infrared image ship detection module. The project flow chart is shown in Figure 1.

## 5.2 Data set processing

The project uses a data set of far-sea (10-12km) ship target detection from the Advanced Research Center for Optics of Shandong University. The data set contains 1044x3 short-wave infrared images, long-wave infrared images and their fusion images, and the size of the data set is continuously updated. We first preprocess the data set, and then use various means of data enlargement to enlarge the capacity of the data set. The larger the capacity of the data set, the more obvious the effect on our training network, and the better the final detection effect.

## 5.3 SRGAN

The structure of GaN network is deeply studied to explore training methods suitable for this project. An excellent GaN application requires good training methods, otherwise the output may be unsatisfactory due to the freedom of the neural network model. Generative adversation network (GADN) is used to generate the appearance information which is lacking in small target image due to its excellent performance to improve the detection accuracy. Realize the target detection of remote ships at night and in bad conditions.

In order to find a suitable network, we chose SRGAN network, which is a generative adversarial network (GAN) for image super resolution (SR). Moreover, it is the first frame to be able to infer photo-realistic natural images with a 4x magnification factor. This network proposes a perceptual loss function, which includes adversarial loss and content loss. Adversarial loss pushes our solution to a natural image manifold using a network of discriminators trained to distinguish between super-resolution images and real images of the original photograph. In addition, we use content loss driven by perceived similarity rather than similarity in pixel space. Our Deep Residual Network is able to recover photo-realistic textures from heavily downsampled images in public benchmarks. Broad mean opinion score (MOS) tests show a significant and significant improvement in perceived quality with SRGAN. The MOS score obtained using SRGaN is closer to the MOS score of the original high-resolution image than that obtained using any of the prior art methods.

## 5.4 Yolov3

### 5.4.1 The advantages and characteristics of YOLOV3 algorithm

You only look once (YOLO) is a state-of-the-art, real-time object detection system. On a Pascal Titan X it processes images at 30 FPS and has a mAP of 57.9% on COCO test-dev.

YOLOv3 is extremely fast and accurate. In mAP measured at .5 IOU YOLOv3 is on par with Focal Loss but about 4x faster. Moreover, you can easily tradeoff between speed and accuracy simply by changing the size of the model, no retraining required!

Prior detection systems repurpose classifiers or localizers to perform detection. They apply the model to an image at multiple locations and scales. High scoring regions of the image are considered detections.

We use a totally different approach. We apply a single neural network to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.

Our model has several advantages over classifier-based systems. It looks at the whole image at test time so its predictions are informed by global context in the image. It also makes predictions with a single network evaluation unlike systems like R-CNN which require thousands for a single image. This makes it extremely fast, more than 1000x faster than R-CNN and 100x faster than Fast R-CNN.

Volume 7 Issue 9, 2021
DOI: 10.6919/ICJE.202109_7(9).0075

YOLOv3 uses a few tricks to improve training and increase performance, including: multi-scale predictions, a better backbone classifier, and more. The full details are in our paper!

YOLOv3 does not have much innovation, but mainly integrates some good solutions into YOLO. However, the effect is still good. Under the premise of maintaining the speed advantage, the prediction accuracy is improved, especially the recognition ability of small objects is strengthened.

### 5.4.2 Yolov3 network architecture

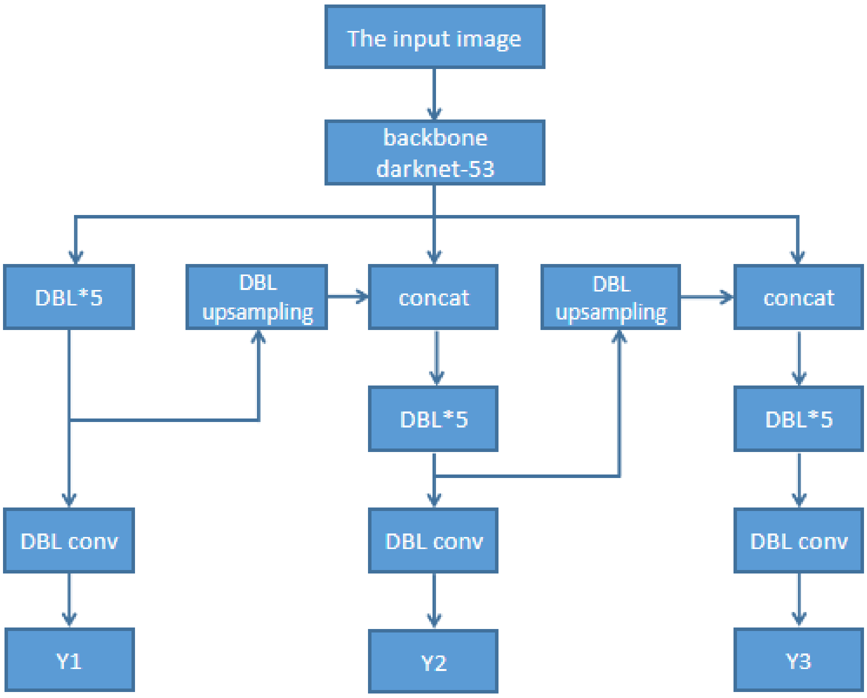


Fig. 2 Yolov3 Network Structure Diagram

YOLOV3 itself uses the full convolutional layer, and even the size modification of the diagram or feature diagram is realized through the convolutional layer. However, the biggest feature of YOLOV3 is Darknet-53 network module. Figure 2 shows the Yolov3 network structure diagram.

In order to achieve better classification effect, the design in Yolov3 trained Darknet 53. Darknet-53 is really strong. Compared to RESNET-152 and RESNET-101, Darknet-53 is not only similar in classification accuracy, but also much faster in computation than RESNET-152 and RESNET-101, and has fewer network layers.

YOLOV3 uses the first 52 layers of Darknet 53 (without full connection layer). YOLOV3 is a full convolutional network, which uses a lot of residual jumping-layer connection. In order to reduce the negative gradient effect caused by pooling, the author directly disuses pooling and uses the stride of CONV to achieve down sampling. In this network structure, convolution with step size of 2 is used for descending sampling.

DBL: DarknetConv2D_BN_Leaky in the actual code, is the basic component of Yolov3. That's convolution plus BN plus Leaky Relu.

Concat: Tensor concatenation. The upper sample of the middle layer of Darknet and a later layer is stitched together. The operation of splicing is different from the operation of residual layer ADD. Splicing will expand the dimension of tensor, while ADD will only directly add without changing the dimension of tensor.

In order to enhance the accuracy of the algorithm in detecting small targets, the UPSAMPLE and fusion method similar to FPN were adopted in YOLOV3 (finally 3 scales were fused, the other two scales were 26×26 and 52×52 respectively), and the detection was performed on the feature maps of multiple scales.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202109_7(9).0075

In the output part, the so-called multi-scale comes from these three predicted paths, Y1, Y2, and Y3 all have a depth of 255, and the rule of side length is 13:26:52. YOLO V3 sets each grid cell to predict 3 boxes, so each box needs to have five basic parameters (X, Y, W, H, confidence), and then there are 80 categories of probability.

The Bounding Box of Yolov3 was further improved by Yolov2. In both YOLOv2 and YOLOV3, K-means clustering is adopted for the object in the image. Each cell in the feature map predicts three bounding boxes, and each bounding box predicts three things: (1) the position of each box (4 values, central coordinates TX and TY, height of the box BH and width BW), (2) an Objectness Prediction, (3) N categories, COCO data set 80 classes, VOC20 classes.

The receptive field corresponding to each of the three tests is different. The receptive field of 32-fold down-sampling is the largest, which is suitable for detecting large targets. Therefore, when the input is 416×416, the three anchor boxes of each cell are (116,90). (156, 198); (373, 326). 16 times is suitable for objects of general size, and the anchor box is (30,61); (62,); (59119). The receptive field of 8 times is the smallest and suitable for detecting small targets, so the anchor box is (10,13); (16, 30); (33, 23). Therefore, when the input is 416×416, there are actually (52×52+26×26+13×13) ×3=10647 proposal boxes. Refer to the test comparison table in Table 1 for details.

Table 1. Comparison table of detection conditions

| feature map | 13*13 | | | 26*26 | | | 52*52 | | |
|---|---|---|---|---|---|---|---|---|---|
| Receptive field | big | | | middle | | | small | | |
| anchor boxes | 116*90 | 156*198 | 373*326 | 30*61 | 62*45 | 59*119 | 10*13 | 16*30 | 33*23 |

Next, I will give you an intuitive sense of the size of the 9 priori boxes. In the figure below, the blue box is the prior box obtained by clustering, the yellow box is the ground truth, and the red box is the grid where the center point of the object is located. The comparison diagram of the detection box is shown in Figure 3.

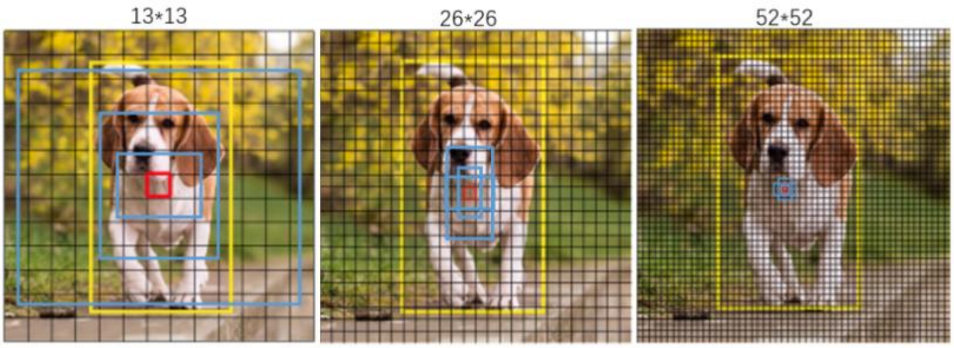


Fig. 3 Comparison diagram of detection box

## 6. Results and Analysis

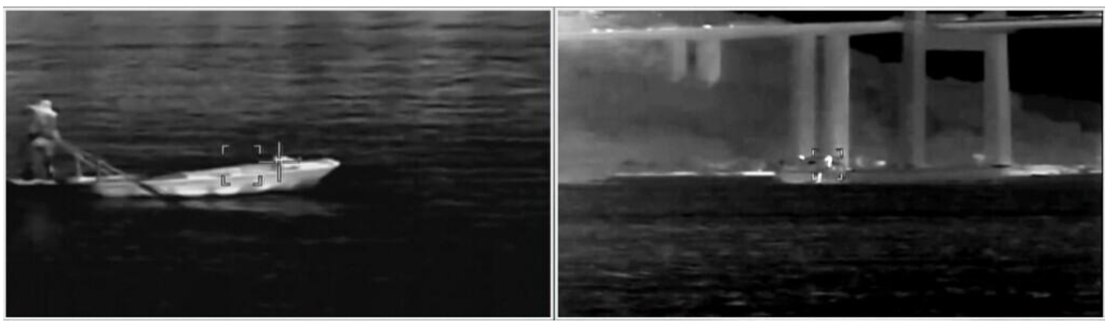

Fig. 4 Infrared image of offshore large target ship

ISSN: 2414-1895
DOI: 10.6919/ICJE.202109_7(9).0075

Fig. 5 Infrared image of a small target ship in the far sea

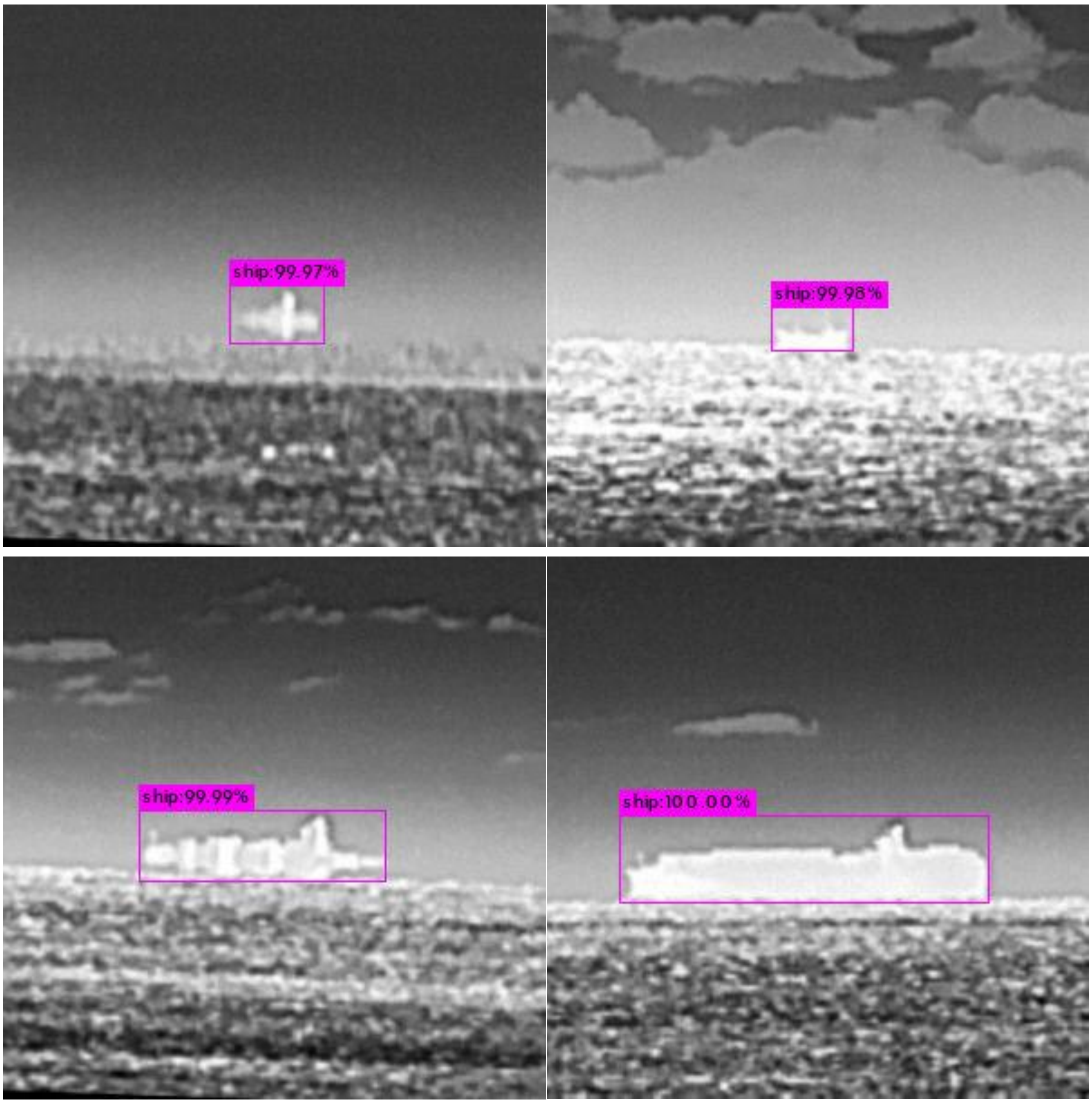


Fig. 6 Infrared image detection effect diagram of a small target ship in the far sea

Fig. 4 is the infrared image of a large target ship taken in close range. In this case, the target is relatively large, the picture is very clear, only with the naked eye can also identify, relatively easy to detect. Figure 5 is the infrared image of a ship with a small target after processing in the long view. In this case, the target is very small and the picture is relatively blurred. At this time, it is impossible to accurately judge the result with the naked eye and it is also difficult to detect. Fig. 6 is the final detection effect. After the training and detection of YOLOV3, significant results have been achieved

ISSN: 2414-1895 DOI: 10.6919/ICJE.202109_7(9).0075

in the detection of ships in infrared images of small targets in the far sea. It can be seen that with the increase of target volume and the improvement of image resolution, the accuracy of target detection is gradually improved. Through the experiment, the project has achieved a very good effect, the overall detection accuracy has reached more than 99%. I have listed four kinds of test renderings from 99.97% to 100.00% for the convenience of showing the results to you. This is shown in Figure 6.

## 7. The application prospect of

The purpose of this task is to realize the detection of small targets of ships through infrared images of far-sea ships. Therefore, this project has broad application prospects both in academic research and in the navigation industry.

### 7.1 Reduce human error in maritime operations

According to the survey, a high proportion of maritime accidents are caused by human factors. Statistics show that 90% of all collisions are caused by human error. At this point, if a set of high-end sophisticated modern technology system equipment assisted, human error will be greatly reduced.

### 7.2 Reduce the incidence of maritime accident

In the complex sea navigation, it is necessary to have a clear grasp of the forward movement in order to carry out safe operations at sea. Through the detection of small target ships in the far sea, the factors affecting navigation safety can be ensured to avoid collision, and the information content and accuracy of the front-end data collection of the control system play a decisive role.

### 7.3 Increase navigational inspection technology means

At present, there is still a lack of systems in China that apply infrared detection technology as a supplementary means of radar to provide auxiliary decision-making for navigation safety and ship collision avoidance. However, this project is focused on the research of infrared images of ship targets, which can be fully applied to future navigation detection technology means.

## 8. Summary and Prospect

The objective of the research results that can be applied in waters land in bad weather on the monitoring and control system for judging the position of the ship in the distance, and the night get ship information lays the foundation for the next step, also can be used for sea ship quickly find attachment shipping position, other navigation information fusion, provide reliable data for navigation of ships at night.

In the future work, I have the following expectations:

(1) We can try to use other network structures and algorithms to detect the target in this data set.

(2) The network can be improved on the basis of YOLOV3 or more complex YOLOV4 and YOLOV5 can be used for detection.

(3) GAN networks with richer functions can be used to generate high resolution images, such as MTGAN networks.

(4) It can process and detect various forms of infrared images in data set, such as mixed wave infrared images, etc.

## References

[1] Bai Y, Zhang Y, Ding M, et al. Finding tiny faces in the wild with generative adversarial network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 21-30.

[2] Y. Bai, Y. Zhang, M. Ding, and B. Ghanem, “Sod-mtgan: Small object detection via multi-task generative adversarial network,” Computer Vision-ECCV, pp. 8–14, 2018.

[3] Zou ZX, Shi ZW Guo YH, and Ye JP. Object Detection in 20 Years: A Survey. Preprint. Available from: arXiv: 1905.05055v1. Cited 13 May 2019.

[4] Dong, Y.-z., et al. Application of soft mathematical morphology in image segmentation of IR ship image [C]. in Signal Processing, 2004. Proceedings. ICSP '04. 2004 7th International Conference on. 2004. 729-732 vol.1.

[5] Jian-Nan, C., et al. A detection method of infrared image small target based on order morphology transformation and image entropy difference [C]. in Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on. 2005. 5111-5116 Vol. 8.

[6] Strickland, R.N. and H. He Il. Wavelet transform methods for object detection and recovery [J]. Image Processing, IEEE Transactions on, 1997, 6(5): 724-735.

[7] LIU Baolong. Research on Character Detection and Recognition of Ship Name Identification Based on Image Analysis and Deep Learning [D]. Hangzhou: Zhejiang University, 2018:1.

[8] Neves, S.R., E.A.B. da Silva, and G.V. Mendonca. Wavelet-watershed automatic infrared image segmentation method [J]. Electronics Letters, 2003, 39(12): 903-904.

[9] Wei, Y., Z. Shi, and H. Yu. Wavelet analysis based detection algorithm for infrared image small target in background of sea and sky [C]. in Image and Signal Processing and Analysis, 2003. ISPA 2003. Proceedings of the 3rd International Symposium on. 2003. 23-28 Vol.1.

[10] Yu-Qiu, S., T. Jin-Wen, and L. Jian. Background suppression based-on wavelet transformation to detect infrared target [C]. in Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on. 2005. 4611-4615 Vol. 8.

[11] Zhou, W., K. Xu, and S. Li. A Method for Ship Target Detection Based on Image Fusion [C]. inElectronic Measurement and Instruments, 2007. ICEMI '07. 8th International Conference on. 2007. 2-923-2-926.

[12] Wang Lidi, Huang Shabai, Shi Zelin. Automatic detection method of small infrared target in sea surface based on wavelet and fractal [J]. Laser & Infrared, 2004, 34(6): 481-483.

[13] Ganhua, L., C. Xuanping, and L. Yunhui. Efficient Target Detection from Infrared Image Sequences Using the Sequential Monte Carlo Method [C]. in Mechatronics and Automation, Proceedings of the 2006 IEEE International Conference on. 2006. 549-554.

[14] Yang, L., J. Yang, and K. Yang. Adaptive detection for infrared small target under sea-sky complex background [J]. Electronics Letters, 2004, 40(17): 1083-1085.

[15] Li Hansong, Lu Chuanyun. Detection of Small Target in Infrared Ship [J]. Infrared and Laser Engineering, 2006, 35(4): 495-498.

[16] Infrared Technology, 2003, 25(3): 9-12. DOI: 10.1016 / j.jcip.2016.01.002 LIU Y W, et al. Preprocessing of ship target infrared image [J]. Infrared Technology, 2003, 25(3): 9-12.

[17] He Youjin, Li Nan. Comparative Study on Edge Detection Methods of Ships Infrared Image [J]. Computer Simulation, 2006, 23(4): 201-203.

[18] ZHANG Tian-xu, et al. A Fast Recursive Infrared Ship Image Segmentation Algorithm [J]. Journal of Infrared and Millimeter Waves, 2006, 25(4): 295-300.

[19] Feng, D. and S. Wen—Kang. Infrared image segmentation with 2-D maximum entropy method based on particle swarm optimizaion [J]. Pattern Recognition Letters, 2005, 26: 597-603.

[20] Kaiyu, X., et al. Target detection based on the artificial neural network technology [C]. in The 9th international conference on Control, Automation, Robotics and Vision. 2006. Singapore.

[21] XU Kaiyu. Research on Detection and Tracking Technology of Moving Ships Based on Infrared Image [D]. Shanghai Maritime University. Doctoral Dissertation

[22] ZHU Qunying. Research on Bridge Pier Collision Prevention Based on Video Image Processing [D]. Wuhan University of Technology. Master Dissertation. 2006.

[23] LIU Jun.Research on Target Detection and Tracking Technology of Moving Ship in Inland River Based on Infrared Image [D]. Chongqing: Chongqing University, 2008:5

[24] CHEN Ren. Research on Ship Target Detection Technology Based on Infrared Image [D]. Shanghai: Shanghai Maritime University, 2005:21

[25] Miao Dechao. Research on Infrared Ship Small Target Detection Technology under Complex Sea and Air Background [D]. Harbin Engineering University. Master Dissertation. 2007.

[26] Tao Wenbing, Jinhai. Journal of Infrared and millimeter Waves, 2007, 26(1): 61-64. (in Chinese)

[27] Zhang Fang, Wang Yuehuan. Infrared Laser Engineering, 2004, 33(6): 603-606. (in Chinese)