Journal of Network Intelligence ©2023 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **8**, Number **2**, May **2023**

A Lightweight Multi-target Detection Method for Infrared Remote Sensing Image Ships

Yong-Mei Zhang*, Rui-Qi Li

School of Information Science and Technology
North China University of Technology
No. 5, Jingyuanzhuang Road, Shijingshan Distric, Beijing 100144, P. R. China
zhangym@ncut.edu.cn
lrq51129@163.com

*Corresponding author: Yong-Mei Zhang

Received October 4, 2022, revised December 5, 2022, accepted March 10, 2023.

Abstract. *Infrared remote sensing ship detection is of great practical significance, whether in military fields of maritime combat dispatching, maritime security, or in the civilian fields of fishery management, import and export cargo transportation, and port management. Aiming at the lower accuracy of existing detection ship methods in multi-category, multi-target and complex backgrounds for infrared remote sensing images, the paper proposes a lightweight multi-target detection method. The CSP Layer of YOLO v5l is replaced by ShuffleNet v2 basic unit, and a coordinated attention mechanism is added to propose a lightweight YOLO v5l improved model, and EIOU is introduced as the loss function of bounding box regression, which combines more bounding box information. The experiment results show the P, R, AP and mAP_0.5 of the presented algorithm are higher than the original model, the parameters are greatly reduced, it also shows better performance for small targets and multi-class detection. In addition, compared with the mainstream target detection models, both are superior to the comparison models. It further illustrates the proposed model has better expression ability for infrared remote sensing ship features. The study is beneficial to providing a reference in subsequent remote sensing target detection.*

Keywords: ship detection, Infrared remote sensing, YOLO v5l, ShuffleNet v2, CA

1. **Introduction.** Compared with traditional detectors, physical field information and other reconnaissance methods, remote sensing technology has unique development advantages because of its observation area, periodic revisit and lack of national border restrictions [1]. With the development of satellite payload and remote sensing imaging technology, the detection and identification of targets in remote sensing images have become possible. As staple transport carriers and vital military targets on the sea, ships play an important role in domestic and international economic development and national security.

Among the existing remote sensing image processing algorithms, most of them are designed for synthetic aperture radar images and visible light remote sensing images, and there are few researches on the detection of infrared remote sensing ship images. Infrared remote sensing can work around the clock, passively detect targets with good concealment, strong anti-interference ability, and long range. It is particularly suitable for detecting infrared dim and small infrared targets in the far infrared band, which makes the research of infrared remote sensing images have irreplaceable theoretical value [2]. Due to the difficulty in obtaining on-board infrared images and the lack of relevant research literature and publicly available datasets, therefore, Infrared remote sensing image target recognition is still in the development stage [3] and has a large research space.

Currently, the detection methods of ships in remote sensing images, mainly based on traditional algorithms and deep learning methods. Traditional ship detection methods, for example method based on gray information [4], visual information extraction method [5-6], machine learning and other methods, rely on artificially designed features and need to design a stable classifier to improve detection accuracy.

Compared with traditional algorithms, deep learning methods have been widely used in human action recognition [7], traffic detection [8], target recognition [9] and other fields. Due to it can automatically learn features from images, its learning and performance capabilities far exceed traditional algorithms. The mainstream target detection algorithms of deep learning include [10]: Two-stage target detection algorithms (RCNN series [11-13]) and One-stage target detection algorithms (YOLO series). The Two-stage detection algorithms consist of generating candidate regions and detailed detection. The One-stage detection algorithms directly output the target category and coordinates while decoding the feature, which are better in reducing time cost and improving computational efficiency. The One-stage target detection algorithms mainly consist of YOLO v3 [14], YOLO v4 [15], YOLO v5, YOLO x [16], SSD [17]. Among them, the YOLO series is constantly updated, and target detection algorithm has a faster recognition speed.

At present, there is relatively little research has been done on infrared remote sensing of ship images. Li et al. [18] firstly performed bicubic interpolation method and labeling of the dataset, and uses the improved YOLO v5s to detect the ship candidate areas, effectively improving the detection accuracy. However, the labeling of the dataset in the method is to manually select ships that meet the characteristics of the dataset and then labeled them. The detection ability of ships and adjacent ships on the coast is weaker. In the field of classic target detection algorithms, Xiang et al. [19] proposed a ship detection method based on salient region extraction and accurate target segmentation, using the saliency of the target, the target search ranges are reduced, accuracy and speed are significantly improved, and the position prediction of the targets in the images are more accurate. However, the method relies too much on the saliency of the target in the global scene, it cannot adapt to the complex and changing background, resulting in the detection method being limited by the scene. Gong et al. [20] proposed a ship target detection method based on guided filtering and adaptive scale local contrast applied to infrared polarization image, first using the Hough transform to detect the candidate sea antenna, segmenting the target region and proposing a sea clutter background suppression algorithm of sea clutter, experiments have been proven to improve ship detection accuracy and check-all rate. The method relies on the target and background local contrast, and the threshold segmentation is more suitable for the case where the contrast between target and background is relatively large, the method has weak generalization ability.

In summary, the existing problems are as follows: 1) In complex marine environments, it is difficult for traditional target detection algorithms to design artificial features with high robustness, which makes the detection method less stable; 2) Although the above methods have achieved good results, most methods are difficult to apply to the detection of complex scenes such as multi-targets and occlusions; 3) The above methods focus on the detection of a single ship category or only detect the ship targets without identifying the targets.

This paper improves the deep learning algorithm to detect and recognize infrared remote sensing ship targets. For complex scenes such as ports, berths, and occlusions, and detect multi-category ships, it can better realize the detection of ship targets in the above scenes and achieve higher accuracy. Specific improvements are as follows: by replacing the Cross Stage Partial Layer (CSP Layer) in the network with the basic unit in ShuffleNetv2, under the premise of ensuring the same feature extraction effects, the number of parameters and calculations are reduced as much as possible. The coordinated attention mechanism is added to make the network pay more attention to the effective features in the process of feature extraction, and retain as much effective information as possible with the network level increases. Efficient Intersection over Union (EIOU) is introduced as the loss function of bounding box regression to fuse more bounding box information. Finally, the improved algorithm can detect multi-class infrared remote sensing ships more effectively.

This paper is organized in the following way: Section 2 presents the related work. Section 3 describes the improved algorithm in detail. The experiment data and evaluation index is introduced in Section 4. The experiment results and analysis are explained in Section 5. The conclusion is drawn in Section 6.

2. Related Work.

2.1. A model based on YOLO V5l. YOLO v5 [21] adds some new improvements to YOLO v4, greatly improving speed and accuracy. YOLO v5 currently has four versions including YOLO V5s, YOLO V5m, YOLO V5l, YOLO V5x. The YOLO V5s network is the YOLO V5 series with the lowest depth and smallest width of the feature maps. However, in detection tasks, YOLO v5s is more suitable for large objectives. The data set in this paper has attributes of smaller targets and poorer images quality. On the premise of ensuring accuracy and mean average accuracy, YOLO v5l is selected as the main network detected in this paper. The network framework diagram is shown in Figure 1.

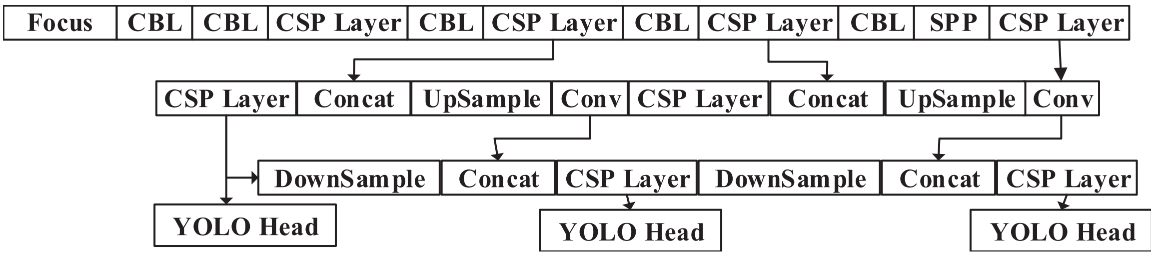


FIGURE 1. The network structures of YOLO v5

2.2. Shufflenet Basic Unit. ShuffleNetV2 [22] network structure is introduced by Megvii Technology after improving ShuffleNetV1 [23]. In the current lightweight neural network, the ShuffleNetV2 neural network uses grouped convolution and channel cleaning to achieve a balance between speed and accuracy. Compared to other CNN, ShuffleNetV2 guarantees precision while reducing computational cost. Figure 2 shows the two commonly used modules of ShuffleNetV2. Figure 2(a) is the module when the stride is equal to 1, which is mainly used to maintain the sizes of the feature map while deepening the number of network layers and increasing the speed. Figure 2(b) is the module when stride is equal to 2, It is mainly used to compress the width and height of the feature layer for under-sampling.

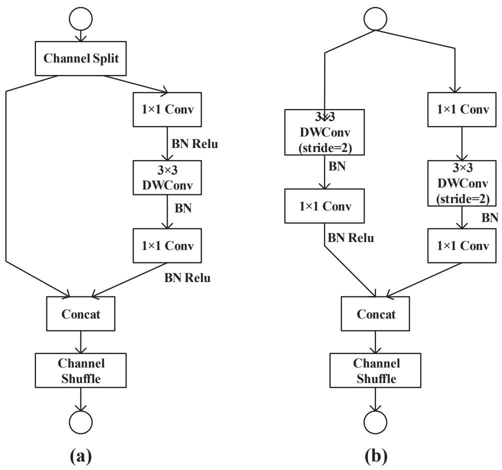


FIGURE 2. ShuffleNetv2 Unit

3. Improved infrared remote sensing ship detection algorithm for YOLO v5l.

3.1. Improve the backbone network and the Neck network. Among the current lightweight neural networks, given the advantages of the ShuffleNetv2 model, this paper uses the basic unit of ShuffleNetv2 to replace the YOLO v5l backbone network. With the deepening of the network hierarchy, the neck network will suffer from network fragmentation and reduce parallelism. In this paper, a module with a stride of 1 in the basic unit of ShuffleNetv2 is employed to improve its neck network. The improved model avoids redundant calculations caused by the CSP Layer in the original structure, and achieves the purpose of a lightweight model. The Shuffle Unit structures in the backbone network are shown in Table 1.

TABLE 1. ShuffleUnit structure introduced in Backbone Network.

Layer	Output size	Stride	Repeat
ShuffUnit1	320×320	2	1
	160×160	1	2
ShuffUnit2	160×160	1	2
	80×80	1	8
ShuffUnit3	80×80	2	1
	40×40	1	8
ShuffUnit3	40×40	2	1
	20×20	1	2

3.2. Add the Coordinate Attention Mechanism. The latest Coordinate Attention (CA) [24] published in CAVR2021 is to embed position information into channel attention. Referring to the literature [25], this paper adds the attention mechanism into the trunk network and the neck network respectively. The improved model structure in this paper is shown in Figure 3.

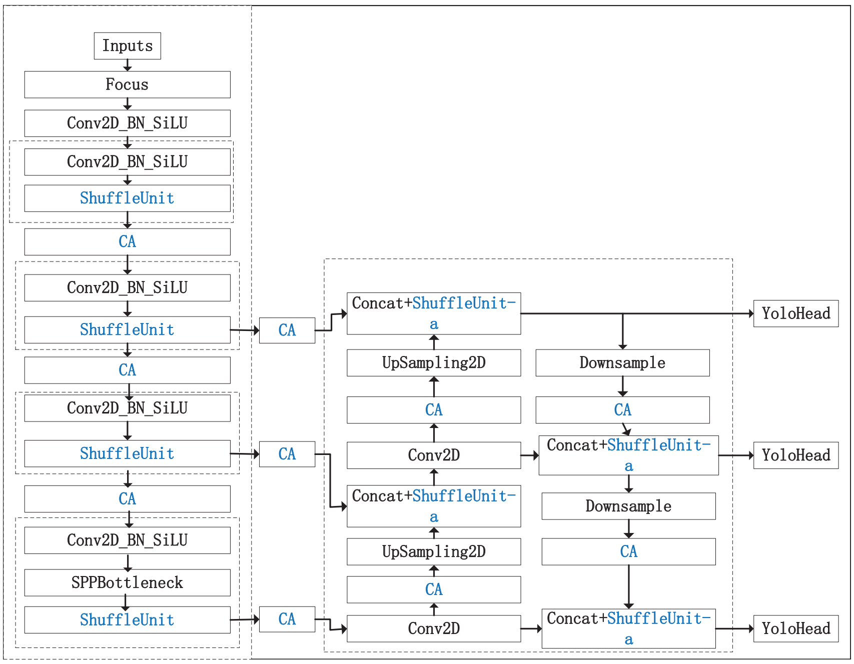


FIGURE 3. Improved overall network structure

3.3. **Loss function of the boundary box.** Although CIOU Loss considers the overlapping area, center point distance, and aspect ratio of bounding box regression, the difference in aspect ratio reflected by v in its formula, rather than the real difference between width and height and its confidence, sometimes hinders the model from effectively optimizing similarity. In order to solve this problem, this paper introduces EIOU loss [26], the formula is as follows.

$$IOU = \frac{A \cap B}{A \cup B} \tag{1}$$

$$L_{EIOU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h 2} \tag{2}$$

b, b_{gt} represents the center points of the predicted frame and the real frame respectively; ρ shows the Euclidean distance between the two center points; c is the diagonal length of the minimum circumscribed rectangle of the real frame and the predicted frame; w, w_{gt} respectively represents the predicted frame and the width of the ground truth box; h, h_{gt} is the width of the predicted box and the ground truth box separately. IOU is the ratio of the intersection and union of the predicted box and the ground-truth box. c_w and c_h are the width and height of the smallest bounding box covering the predicted and ground truth boxes. In frame regression loss, the width and height loss in EIOU Loss make the convergence speed faster and the accuracy higher, so this paper adopts the EIOU loss frame regression loss function with better performance.

4. Experimental data and evaluation index.

4.1. **Experiment dataset.** The dataset adopts the latest infrared ship target recognition database produced by Arrow Technology Co., Ltd. and published in October 2021. The database uses infrared equipment with different resolutions and focal lengths to collect more than 8,000 infrared images in different scenarios. The specific distributions are shown in Table 2. Images are divided into training and test sets according to the ratio of 7:3.

TABLE 2. The dataset.

Classes	Liners	Sailboats	Warships	Canoes	Bulk carriers	Container ships	Finshing boats
Num	1433	5777	2547	4935	1940	695	9118

4.2. **Experiment environment and evaluation indicators.** All the experiments perform on the Ubuntu20.04 operating system, the CPU is 20 Intel(R) Xeon(R) Silver 4210 CPU@ 2.20GHz, the GPU is NVIDIA GeForce RTX 2080Ti with 12G memory, and the Python version is 3.8, the development tool is pycharm2019.3.3, and the framework adopts Pytorch1.11.0. The size of the input is 640×640.

4.3. **Evaluation criteria.** Precision(P), Recall(R), Average Precision(AP), Mean Average Precision (mAP), and parameters are adopted as evaluation criteria. Precision is an index to evaluate the accuracy of the model detection targets, the average precision measures the quality of all categories of the model, which is one of the most important indicators in target detection. The formulae for P, R and mAP list as follows:

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

$$mAP = \frac{1}{classes} \sum_{i=1}^{classes} \int_0^1 P(R)d(R) \tag{5}$$

In the formulas, TP means that the target is a ship, and the detection result is the type of the corresponding ship, FP indicates the target is not a ship, and the detection result is a ship, FN shows the target is a ship, and the detection result is not the type of the corresponding ship.

5. **Experiment results and analysis.**

5.1. **Ablation experiments.** In order to analyze the impact of different improvements on network performance, eight groups of comparative experiments are designed. Each group of experiments apply the same training parameters, where ✓ indicates the corresponding improvement strategy used in the model. Table 3 shows the effects of different methods on the performance of model detection.

TABLE 3. Total category of data from the ablation experiments.

	Backbone	Neck	CA	P(%)	R(%)	mAP_0.5(%)	Parameters(MB)
1				88.18	74.25	84.07	3.03×10^7
2	✓			89.98	68.44	83.13	5.28×10^6
3		✓		93.34	76.58	89.03	3.30×10^7
4			✓	89.83	66.41	79.12	4.68×10^7
5	✓	✓		89.47	74.31	85.09	1.66×10^7
6		✓	✓	91.68	75.67	86.31	3.31×10^7
7	✓		✓	90.37	76.66	86.70	3.03×10^7
8	✓	✓	✓	92.96	89.40	92.03	2.43×10^7

The P, R and mAP_0.5 are all obtained at the IOU threshold of 0.5. Table 3 shows the comparison of ablation experiments in the total categories. In the second experiment, the index is slightly lower than the original model on R, mAP_ 0.5, but the original model is nearly 5.8 times the number of parameters. In the fourth experiment, as far as the original model is concerned, the effect of only adding CA is not better, but the indicators of the improved backbone and the neck network after adding CA are improved compared with the original model. The P, R and mAP_0.5 of the final improved model outperformed the original model, increasing by 4.78%, 15.15% and 7.96%, respectively. As can be seen from the number of parameters, the improved backbone network is minimal, followed by the improved backbone and neck networks. The final improved model parameter count is 80% of the original YOLO v5l parameter count.

Among the improved models, the lightweight improvement of the backbone is not as good as that of the trunk network for the original model. The main reason is that the number of parameters declines and some features cannot be extracted more accurately, which leads to some adverse effects on the extraction of the backbone features. With the addition of the CA module, it focuses more on effective features, and the effect has improved. In general, the presented method has better results. The reason is that the ShuffleUnit module strengthens the compactness of the network, and generally maintains the flexibility of the original YOLO v5l architecture, reduces excessive repeated feature extraction, and lessens the fragmentation of the network.

As can be seen from Table 4, the proposed method in this paper are mostly better than those of the original model and only the accuracy of the Liner is slightly lower than the original model. From the bar chart in Figure 4, it can be seen that the proposed method in this paper has improved mAP values for each type of ship detection results.

5.2. **Performance comparison of classic target detection models.** To verify the detection performance of the improved algorithm for different types of ships in the dataset, the improved algorithm is compared with the classic target detection algorithms. The comparison results are shown in Table 5.

Table 4. Ablation experiment data of single category.

Method / classes	P(%)		R(%)		AP(%)	
	YOLO v5l	The proposed model	YOLO v5l	The propoesd model	YOLO v5l	The proposed model
Sailboats	86.44	90.62	62.43	85.96	80.52	90.39
Warships	94.48	97.47	95.80	97.84	98.21	98.63
Canoes	79.57	87.54	51.19	78.10	65.10	83.15
Bulk carriers	87.46	96.25	88.55	96.42	89.67	97.39
Container ships	93.99	94.29	83.50	96.12	92.39	96.30
Fishing boats	85.89	92.10	65.90	85.02	77.17	89.61
Liners	93.03	92.42	72.41	86.32	85.40	88.74

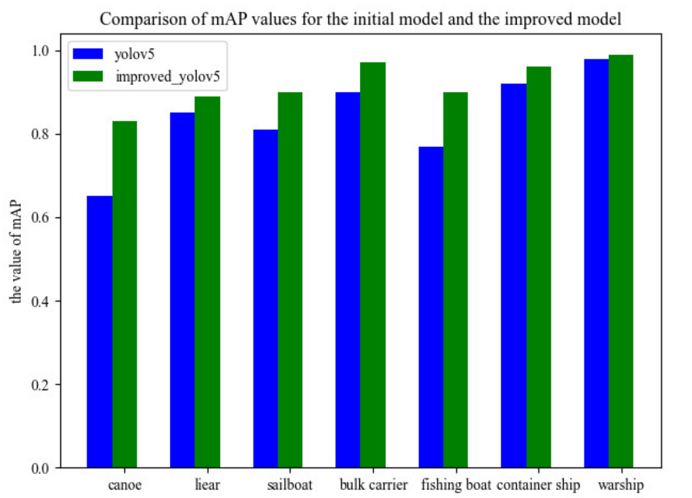


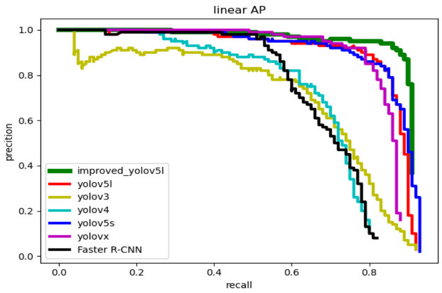
Figure 4. Comparison of mAP values for the yolov5l and this method

Table 5. Classic algorithm comparison results.

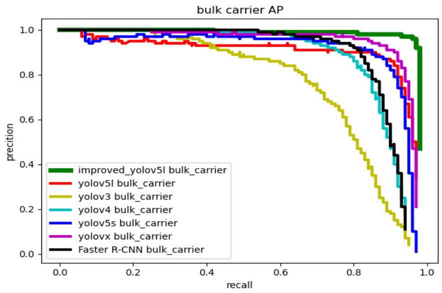
Algorithms	P(%)	R(%)	mAP(%)	Params
YOLO v3	85.42	40.01	67.32	6.15×10^{7}
YOLO v4	86.35	65.58	77.06	6.39×10^{7}
YOLO v5s	90.39	71.24	84.10	7.07×10^{6}
YOLO X-s	91.59	79.72	87.76	5.41×10^{7}
Faster R-CNN	60.02	68.54	64.47	1.36×10^{8}
This algorithm	92.96	89.40	92.03	2.43×10^{7}

For a more comprehensive comparison, this paper selects the classic two-stage target detection algorithm and the one-stage target detection algorithm for experiment comparison. All experiment configurations and parameters remain the same. From Table 5, it can be seen that the P, R and mAP_0.5 values of the presented algorithm are better than the other classical models.

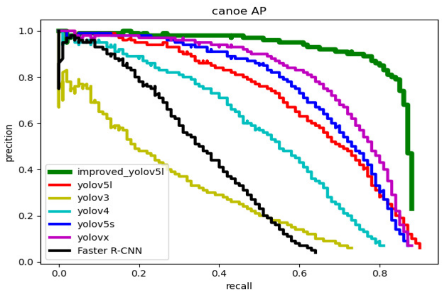
The P-R curve comparison chart of the algorithms in Table 5 after running them on the same test set is further given in Figure 5. The abscissa is the recall rate and the ordinate is the corresponding accuracy rate. It can be seen that the proposed algorithm occupies the largest area in each class of ships, that is to say, under the same recall rate, the ship detection method proposed in this paper has a higher accuracy rate.



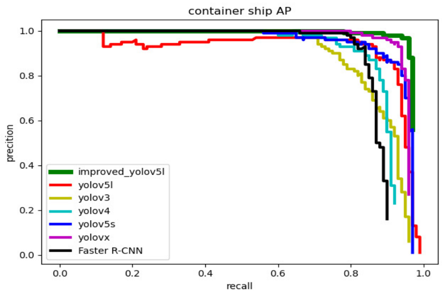
(a) linear AP



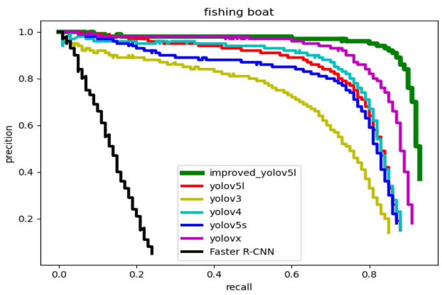
(b) bulk_carrier AP



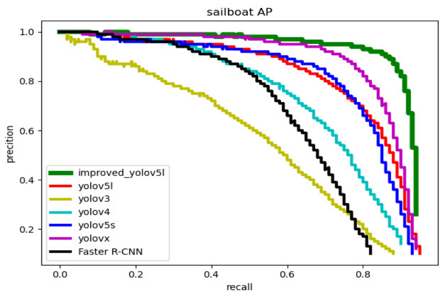
(c) canoe AP



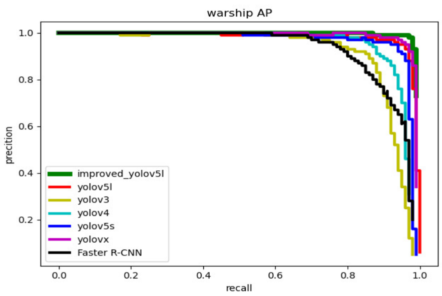
(d) container AP



(e) fishing_boat AP



(f) sailboat AP

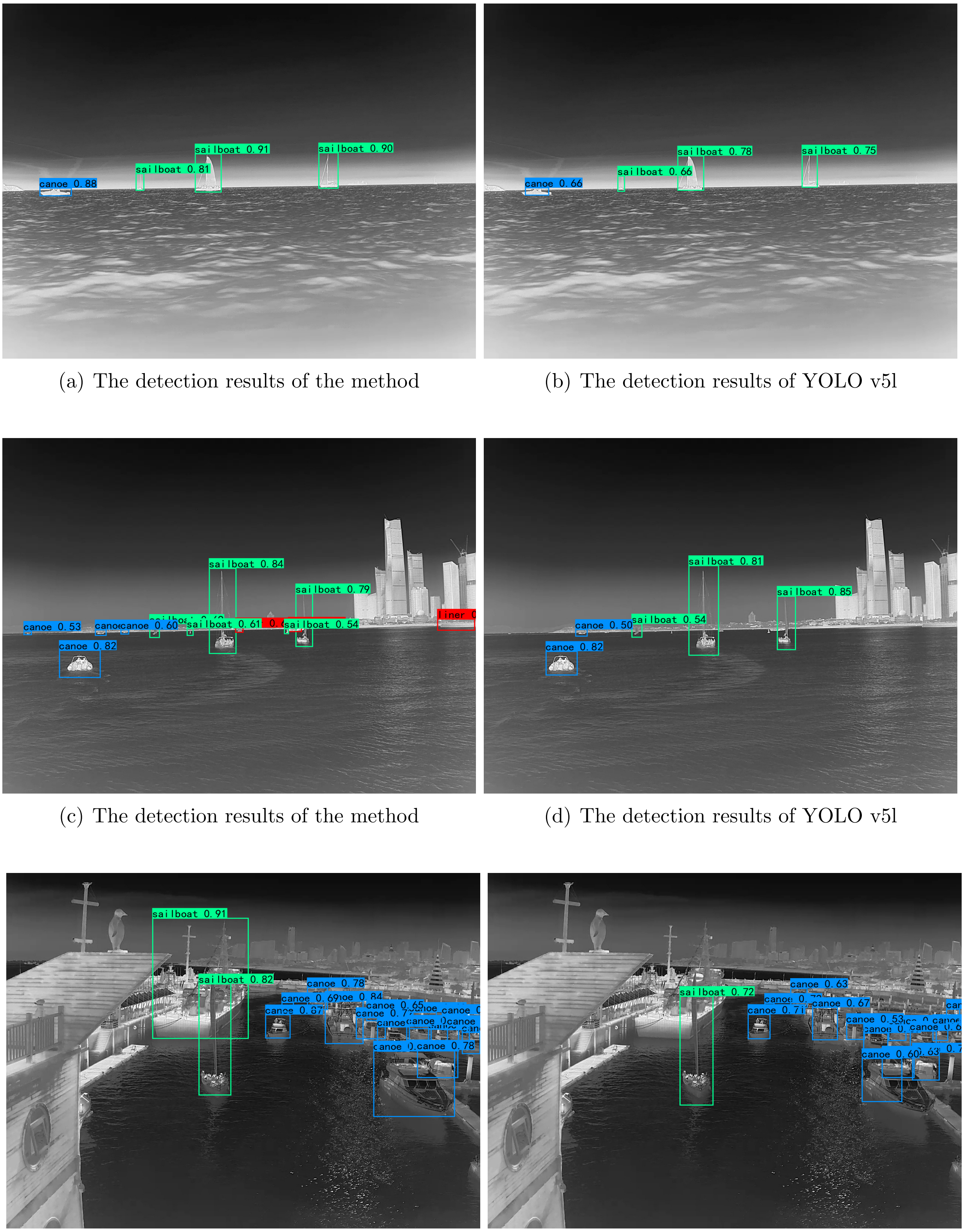


(g) warship AP

FIGURE 5. Comparison of P-R curves of classical algorithms.

Figure 6 selects typical multi-target, multi-class and complex scene detection results for comparison. It can be seen in Figure 6 (a) and (b), both can identify the ships, but the proposed algorithm has higher accuracy. In Figure 6 (c) and (d), the improved algorithm detects small targets missed by the original model, the improved algorithm can better detect small targets in infrared remote sensing images. By observing Figure 6 (e) and (f), it can be seen the improved algorithm detects the large targets missed by the original model and some overlapping targets. The results verified the improved algorithm has stronger robustness in small targets and complex scenes.

The algorithm in this paper improves the YOLO v5l model. On the premise of ensuring that the detection effect is not reduced, the number of parameters is greatly reduced, and



(a) The detection results of the method (b) The detection results of YOLO v5l

(c) The detection results of the method (d) The detection results of YOLO v5l

(e) The detection results of the method (f) The detection results of YOLO v5l

FIGURE 6. The detection results of YOLO v5l and the proposed algorithm.

the coordinate attention mechanism is used and the introduction of EIOU, so that it can pay more attention to the spatial direction and position sensitive information, and make full use of the feature map information, more bounding box information is fused, which is more suitable for detecting small infrared remote sensing targets. The improved algorithm can achieve high detection accuracy without introducing additional computational burden.

6. **Conclusions.** Aiming at the problems of single target category, lower detection accuracy and larger number of model parameters for ship detection in infrared remote sensing images, an improved lightweight model based on YOLO v5l is proposed for detection of multi-target and multi-class infrared remote sensing ships. First, the basic module of ShuffleNet is used alternatively to replace the backbone network of YOLO v5l, in order to reduce the number of parameters and the effectiveness of the feature extraction, the basic module continues to be used to improve the neck network; for the better preserve location and channel information of the feature map, the model can be positioned and identified more accurately, CA modules are added to the backbone and neck networks respectively; the introduction of the EIOU loss function in order to incorporate more edge information. The experiment results show it is 4.78%, 15.15% and 7.96% higher than the original model in terms of P, R, and mAP_0.5, and the parameters are reduced by 80%.

At present, there is a lack of multi-angle and all-round infrared image ship data, as well as infrared remote sensing image data sets. Aiming at the shortage of infrared remote sensing image ship samples, the next problem to be solved is to establish an infrared remote sensing image ship data set to improve image quality. Infrared ship data of simple and complex scenes can be generated by improving deep learning models or algorithms, and can also be combined with other optical remote sensing images to establish infrared remote sensing image ship data sets.

REFERENCES

[1] X.-Q. Meng. "Ship Detection and Motion Monitoring of Remote Sensing Images Based on Deep Learning," M.S. dissertation, Jilin University, 2022.

[2] W.-N. Pan. "Research on infrared imaging ship target detection method based on deep learning," M.S. dissertation, University of Electronic Science and Technology of China, 2018.

[3] R.-L. Ding, J. Li, B.-R. Xie, B.-R. Wang, and Y.-Z. Zhen, "Ship Target Detection Method Based on Infrared Remote Sensing Image," *Optics & Optoelectronic Technology*, vol. 19, no. 6, pp. 24-33, 2021.

[4] J.-Q. Ai, X.-Y. Qi, W.-D. Yu and F. Liu, "A New Ship Wake CFAR Detection Algorithm in SAR Images Based on Image Segmentation and Normalized Hough Transform," *Journal of Electronics & Information Technology*, vol. 32, no. 11, pp. 2668-2673, 2010.

[5] F.-Q. Zhang, T.-Y. Wu and G.-Y. Zheng, "Video salient region detection model based on wavelet transform and feature comparison," *Eurasip Journal on Image and Video Processing*, vol. 58, 2019. [Online]. Available: https://doi.org/10.1186/s13640-019-0455-2

[6] P. Ding, Y. Zhang, P. Jia, X.-L. Chang and R. Liu, "Ship Detection on Sea Surface Based on Visual Saliency," *ACTA Electronic SINICA*, vol. 46, No. 1, pp. 127-134, 2018.

[7] F. Zhang, T.-Y. Wu, J.-S. Pan, G.-Y. Ding and Z.-Y. Li, "Human motion recognition based on VR art media interaction environment," *Human-centric Computing and Information Sciences* vol. 9, no. 1, pp. 1-15, 2019.

[8] K. Sachin, D. Agam, K.Aditya, K. Saru and C.-M. Chen, "LSTM Network for Transportation Mode Detection," *Journal of Internet Technology*, vol. 22, no. 4, pp. 891-902, 2021.

[9] K. Wang, C.-M. Chen, M.-S. Hossain, G. Muhammad, S. Kumar and S. Kumari, "Transfer reinforcement learning-based road object detection in next generation IoT domain," *Computer Networks*, vol. 193, 2021.

[10] Z.-X. Yang, Y.-Q. Tang, J.-J. Zhang, and P.-Z. Geng, "Detection Algorithm of Pedestrian Shoe Area Based on Improved YOLO v4," *Laser & Optoelectronics*, vol. 59, no. 8, pp. 121-130, 2022.

[11] E. -K. Wang, F. Wang, S. Kumari, J.-H. Yeh, C. -M. Chen, "Intelligent monitor for typhoon in IoT system of smart city," *The Journal of Supercomputing*, vol. 77, no. 3, pp. 3024-3043, 2021.

[12] S.-Q. Ren, K.-M. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processiong Systems*, vol. 28, pp. 1-13, 2015.

[13] S.-M. Amira, A.-M. Sayed, A.-E. Reda, M.-A. Hisham, "Object detection using adaptive mask RCNN in optical remote sensing images," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 1, pp. 65-76, 2020.

[14] R.-J. Song, D.-H.Jin, and K. Dovagdorj, "Detection of Insulator Defects Based on Improved YOLOv3," *Journal of Network Intelligence*, vol. 6, no. 4, pp. 859 - 872, 2021.

[15] B. Alexey, C.-Y. Wang, and H.-Y. Mark, "Yolov4: Optimal speed and accuracy of object detection," *arXiv 2020*, 2020. [Online]. Available: https://doi.org/10.48550/arXiv.2004.10934.

[16] G. Zheng, S.-T. Liu, F. Wang, Z.-M. Li, and J. Sun, "Yolox: Exceeding yolo series in 2021," *arXiv 2020*, 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2017.08430.

[17] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A.-C. Berg, "SSD: Single Shot Multi-Box Detector," in *European Conference on Computer Vision (ECCV)*. ECCV, 2016, pp. 21-37.

[18] L.-Y. Li, L.-Y. Jiang, J.-W. Zhang, S.-Q. Wang, and F.-S Chen, "A Complete YOLO-Based Ship Detection Method for Thermal Infrared Remote Sensing Images under Complex Backgrounds," *Remote Sensing*, vol. 14, no. 7, pp. 1534-1547, 2022.

[19] T. Xiang, "Ship Detection in Infrared Image Based on Salient Region Extraction," *Telecommunication Engineering*, vol. 60, no. 7, pp. 785-790, 2020.

[20] J. Gong, J.-W. Lv, L. Liu and R.-C. Chou, "Ship Target Detection in Infrared Polarization Image," *Spectroscopy and Spectral Analysis*, vol. 40, no. 2, pp. 586-594, 2020.

[21] L. Zhu, X. Geng, Z. Li, and C. Liu, "Improving YOLO v5 with Attention Mechanism for Detecting Boulders from Planetary Images," *Remote Sensing*, vol. 18, no. 13, pp. 3776-3784, 2021.

[22] N.-N Ma, X.-Y. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of European Conference on Computer Vision*. ECCV, 2018, pp. 116-131.

[23] X.-Y. Zhang, X.-Y Zhou, M.-X. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile device," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2018, pp. 6848-6856.

[24] Q.-B Hou, D.-Q Zhou, and J.-S Feng, "Coordinate Attention for Efficient Mobile Network Design," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2021, pp. 13713-13722.

[25] F. Xie, B. Lin, and Y. Liu, "Research on the Coordinate Attention Mechanism Fuse in a YOLO v5 Deep Learning Detector for the SAR Ship Detection Task," *Sensors*, vol. 22, no. 9, pp. 3370-3386, 2022.

[26] Y.-F. Zhang, W. Ren, Z. Zhang, Z. Jia, L. Wang and T. Tan, "Focal and efficient IOU loss for accurate bounding box regression," *Neurocomputing*, vol. 506, no. 28, pp. 146-157, 2022.