

RESEARCH ARTICLE

Shelter Identification for Shelter-Transporting AGV Based on Improved Target Detection Model YOLOv5

DIAN YANG¹, CHEN SU, HANG WU², XINXI XU, AND XIUGUO ZHAO

¹Institute of Medical Support Technology, Academy of System Engineering, Tianjin 300161, China
²Academy of Military Sciences, Tianjin 300161, China

Corresponding author: Xiuguo Zhao (zhaog2011@126.com)

ABSTRACT Shelter identification is the fundamental issue for the shelter-transporting automated guided vehicle to detect and transport shelter effectively. Actively identifying shelter faces the challenge of high accuracy but slow speed using a complex model, and fast speed but low accuracy using a simple model. However, all kinds of target detection algorithms available has difficulty in achieving both high detection accuracy and speed. In this paper, the model YOLOv5n6* is developed based on the modified YOLOv5 model by selecting different model structures, introducing an attention mechanism, and improving loss function and non-maximum suppression. Then, the experiments for shelter recognition were carried out using the model YOLOv5n6*. The experimental results show that the box_loss is reduced by 1.2%, the mAP_0.5:0.95 is improved by 2%, and the detection accuracy is improved by 0.87% for the improved model YOLOv5n6* compared with the YOLOv5n6. However, the YOLOv5n6* size is only 7.2M, and the detection time is increased by 0.2ms. So it is proved that the modified model YOLOv5n6* not only has a significant improvement in the shelter detection ability but also has strong robustness, which meets both the requirements of the recognition accuracy and the detection speed.

INDEX TERMS YOLOv5, automated guided vehicle, target detection, attention mechanism.

I. INTRODUCTION

Field shelter hospitals are of great significance to accomplish medical treatment for the patients in medical support, emergency prevention and control of infectious diseases, and emergency medical rescue of sudden disasters. The field shelter hospital consists of multiple shelter modules, and its deployment involves the cooperation of multiple modules, so the shelter transfer has become an important step during its deployment process. The shelter should be transported to the assignment position to satisfy docking requirements among shelters. Shelter-transporting efficiency seriously affects the deployment of shelter hospitals [1], [2], [3], [4], [5], [6]. Shelter-transporting automated guided vehicle (AGV) as an intelligent transportation device, has higher intelligence and can complete the shelter-transporting at a faster speed and

higher precision. However, the shelter-transporting AGV faces extremely serious problems that shelter-transporting AGV identify shelters have low detection accuracy and slow speed. Shelter-transporting AGV combining computer vision can effectively solve the above problems during the deployment of field shelter hospitals. The target detection algorithm can accomplish the shelter accurate recognition, then improves shelter-transporting AGV efficiency and reduces the deployment time of field shelter hospitals.

Recently, many scholars have proposed different target detection models [6], [7], [8]. For example, the two-stage algorithm model is represented by region-network backbone network series [9], and the two-stage target detection network use region extraction operation. Firstly, the convolution neural network backbone was used to extract image features, and then the possible candidate regions were found from the feature map. Finally, the sliding window operation was carried out on the candidate regions to further judge the

The associate editor coordinating the review of this manuscript and approving it for publication was Jiju Poovancheri¹.

target category and location information. In order to further improve the real-time performance of target detection, some scholars put forward a simplified algorithm model that transforms target detection into a regression problem. With improving detection accuracy and detection speed at the same time, you only look once (YOLO) and single shot detector of single-stage target detection models based on position regression were proposed respectively [10], [11], [12], [13], [14]. A new object detection framework is proposed, which deals with the multi-scale problem of target by adding a multi-angle anchor frame, and a dual-channel feature fusion network was designed to learn local and context attributes along two independent paths [15]. Li et al. [16] presented a method based on the combination of YOLOv3 and kernel correlation filter to track the target vehicle. The YOLOv3 framework was improved in order to solve the problem of super-large scale object recognition in the same scene [17]. However, it has a longer detection time for the model. In the training process, the strategy of dynamic cross-combination was proposed for targets of different scales. The experimental results show that the modified model improves the recognition ability of super-large objects. A mobile detection model based on YOLOv5 was constructed in order to improve the framework. The YOLOv5 was used as the main framework and the backbone was replaced by Mobilenetv2, which compared with the baseline model the number of parameters was reduced by half and the detection speed was increased by 47% [18]. But its improved model reduces the detection accuracy. Actually, the YOLOv5 target detection algorithm combining the global satellite navigation, the visual navigation, and the laser navigation was applied to shelter-transporting AGV navigation to conduct achieve shelter rapid transportation and positioning and also show the good field environment adaptability [19]. However, GPS is prone to failure, and the identification accuracy and efficiency of the shelter still need further improvement.

In summary, this paper proposes the modified YOLOv5 model by selecting suitable model structures, introducing an attention mechanism and improving loss function and non-maximum suppression (NMS), which further improve the recognition accuracy and efficiency of the shelter identification method for the shelter-transporting AGV. Then the experiments for shelter recognition were carried out using the modified YOLOv5 model and the experimental results were discussed. The experimental results show that our improved model solves the problems of long detection time, low accuracy, and GPS failure in the above literature. The work pushes the development of the shelter identification method on improving the accuracy and speed of shelter-transporting AGV's identification shelter.

The main contributions of this paper are summarized as follows:

- We propose an improved model with high accuracy and speed that targets to solve the shelter target detection problem. Specifically, we introduce the attention mechanism, a new loss function and NMS.

- We carried out training and detection experiments using ten pre-trained models of YOLOv5 and selected YOLOv5n6 after carefully considering both detection accuracy and speed.
- We have trained and tested with the proposed model. The results show that our method can effectively improve the performance of the original model and provide reliable target detection capability.

The rest of the paper is organized as follows: Chapter 2 describes the YOLOv5 target detection model and analysis problem. Chapter 3 explains how to improve the YOLOv5n6 model to YOLOv5n6*. Chapter 4 conducts experimental analysis. Finally, Chapter 5 gives the research conclusions.

II. YOLOV5 MODEL AND ANALYSIS

In June 2020, the Ultralytics team proposed the YOLOv5 model based on YOLOv3, which was very fast in recognition, high performance and easy to use [20], [21], [23], [24].

The YOLOv5 algorithm has five network structures, namely YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. The difference between the five network structures is that the width and depth of the cross-stage partial (CSP) structure of the Backbone and Neck are different. The number of convolution and the number of residual blocks are different [25]. YOLOv5 simple network structure is shown in Figure 1 [26], [27], [28], [29], [30], its network model is divided into four parts, which are Input, Backbone, Neck and Prediction.

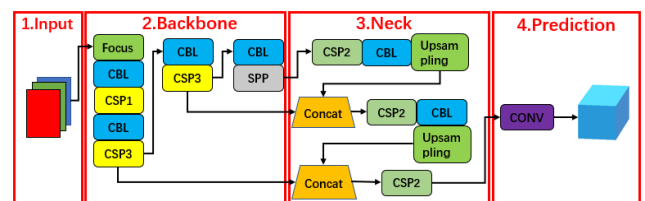


FIGURE 1. YOLOv5 network structure diagram.

(1) Input: The Input includes three parts: Mosaic data enhancement, image size processing and adaptive anchor frame calculation.

(2) Backbone: Backbone is the key part of the network, mainly composed of Focus and CSP. In the Focus structure, the key point is the slicing operation [31]. The CSP structure divides the input into two parts. One part performs a certain operation first, and then carries out convolution operation. And the other part goes straight to the convolution.

(3) Neck: Neck adopts Feature Pyramid Network (FPN) and Pyramid Attention Network (PAN) structure. The FPN structure processes image features from top to bottom. The PAN adopts the bottom-up feature pyramid idea.

(4) Prediction: Prediction includes bounding box loss function and NMS [32], [33], [34], [35].

Through the above discussion, it was can be found that the model with the more complex the structure and the deeper the depth usually has the better the detection effect in the

YOLOv5 model, but has the lower training efficiency and the higher the weights. Complex detection models require more computing resources, which seriously reduces the real-time performance of target detection and is harmful to the deployment of engineering applications. In order to design a model with better performance on the detection speed and accuracy for the shelter recognition in the field environment, the optimal model should be chosen.

YOLOv5 v6 is the latest version of the YOLOv5 series. It integrates many new features and fine-tunes the network structure, proposing the new models YOLOv5n and YOLOv5n6. We train ten pre-trained models of YOLOv5 v6 using the original shelter dataset, and then the trained models are applied to detection experiments. The experimental results show that YOLOv5n6 has high detection accuracy and fast detection speed. So we chose YOLOv5n6 as the original model. The detailed analysis and selection of YOLOv5n6 can be found in Chapter 4 on MODEL SELECTION.

YOLOv5n6 has good detection performance and inference speed compared with the other nine models but has shortages that need to be improved for shelter detection, as the followings:

(1) In the process of shelter-transporting AGV movement, most of the video information is useless. A large number of video information input greatly occupies computing resources and reduces the real-time performance and stability of target detection.

(2) When the prediction box is inside the target box and the size of the prediction box is consistent, the boundary box regression loss function generalized-IOU (GIOU) will degenerate into a simple intersection-over-union (IOU) loss function, which cannot achieve accurate positioning of the prediction box and better optimization of the model. Moreover, the application of GIOU loss function prediction box in horizontal or vertical direction is difficult to optimize, and the convergence is slow, which reduces the training efficiency.

(3) In the process of post-processing, screening of more target frames usually need to be processed by the NMS. The NMS algorithm is sensitive to the setting of the overlap threshold, which set too low will lead to leak and too high will lead to error checking. For the recognition of partially overlapping targets, only unobstructed targets can be detected, and partially obstructed targets have no output detection results.

III. MODEL IMPROVEMENT METHOD

A. ATTENTION MECHANISM APPLICATION

Presently, most researchers introduced attention mechanism to solve the problem of small target detection. However, the attention mechanism can well solve the real-time and stability problems of shelter-transporting AGV shelter detection, so the shelter-transporting AGV will introduce an attention mechanism to accomplish the shelter detection in the process of shelter transportation.

Attention mechanisms can be classified as channel attention, spatial attention, convolutional block attention

module (CBAM), squeeze-and-excitation networks. Among the above four attention mechanisms, CBAM was chosen to applied to shelter detection because of its simplicity and effectivity. It is a lightweight module and can be trained in an end-to-end manner [36], [37], [38], [39], [40]. CBAM combines the attention mechanism of feature channel and feature space. Given a feature map, CBAM will successively infer the attention map along the two independent dimensions of channel and space, and then multiply the attention map with the input feature map to perform adaptive feature refinement. CBAM was integrated into YOLOv5 to help the model quickly find the region of interest in a wide range of images. The structure of CBAM module is shown in Figure 2 [41].

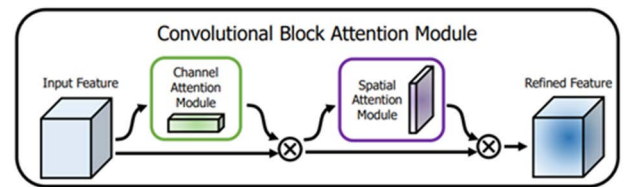


FIGURE 2. CBAM module structure.

During the shelter-transporting AGV movement, CBAM can be used to extract the attention area from the video information obtained by the camera fixed on the shelter-transporting AGV, and focus on the useful information instead of the useless information in shelter detection process.

B. CIOU_LOSS IMPROVEMENT

The original YOLOv5 uses GIOU as the loss function, but its model was low on both optimization degree and convergence speed in the application of shelter detection. However, complete-IOU (CIOU) pays attention to the scale information of the width to height ratio of the boundary frame, and increases the size of the detection frame as well as the loss of length and width, making the prediction frame more consistent with the real frame. Therefore, CIOU was chosen as the loss function. In order to further improve the positioning accuracy of the YOLOv5 model, this study will analyze the influence of the distance between the center point of the detection frame, and the annotation frame and the aspect ratio was added on the basis of considering the overlap area. If the loss function CIOU replacing GIOU was introduced to YOLOv5, the modified YOLOv5 will achieve better results. CIOU_Loss serves as a loss function, as shown in Equation 1.

$$\begin{aligned}
 CIOU_Loss &= 1 - CIOU \\
 &= 1 - \left(IOU - \frac{Distance_{22}}{Distance_{C^2}} - \frac{v^2}{(1 - IOU) + v} \right) \\
 v &= \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^p}{h^p} \right)^2 \quad (1)
 \end{aligned}$$

where, $Distance_{22}$ is the coordinates of the center point of the label box and the prediction box. $Distance_{C^2}$ is the diagonal distance of the smallest bounding rectangle. w^p is

the prediction frame width. h^p is the height of the prediction box. w^{gt} is the width of the label box. h^{gt} is the height of the label box.

C. DIOU_NMS IMPROVEMENT

NMS is mainly used for filtering prediction frames. The highest scoring detection box and other detection boxes are counted for a corresponding IOU value in the classical NMS. All boxes with this value exceeding the NMS threshold are filtered out. IOU is the only factor considered in the classical NMS algorithm. The original YOLOv5 used the classic NMS and had the problem of targets being missed. Therefore, YOLOv5 needs to introduce a better NMS. Distance-IOU_NMS (DIOU_NMS) considers the IOU and the distance between the center points of two enclosures. If both the IOU between two frames is large and the center distance between two frames is large, DIOU_NMS will consider these frames as two objects and will not filter the blocked object.

D. NETWORK STRUCTURE IMPROVEMENT

By introducing the above CBAM, CIOU_Loss, and DIOU_NMS into the original YOLOv5, we obtain the improved network structure, as shown in Figure 3. We add CBAM to the Backbone section in front of the SPP module. It solves the problem of no attention preference in the original network. It allows the network to focus more on detecting the target and improves detection efficiency. Then we inserted CIOU_Loss, and DIOU_NMS into the Prediction section. CIOU_Loss enhances the speed and accuracy of the prediction frame regression. DIOU_NMS solves the problem of missed targets.

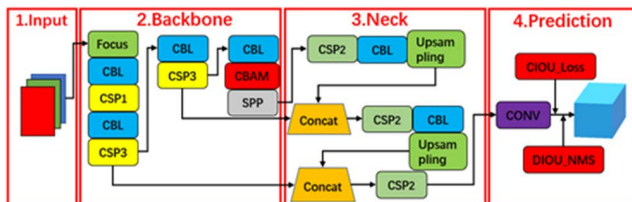


FIGURE 3. Network structure improvement diagram.

IV. EXPERIMENT AND DISCUSSION

A. DATA SET

The original dataset source was divided into two parts, one is the static image data captured by the camera, and the other is the image data obtained from the video screenshots captured by the camera when the shelter-transporting AGV was working. Images were named uniformly, and 1000 images were selected as the total data set of target detection training. The data set was divided into 800 training sets and 200 test sets.

Labeling tool was used to annotate each image in the dataset, as shown in Figure 4. The corresponding relationship between shelter, red cross and label in the data set is shown in Table 1. The anchor frame should completely cover the target in the annotation process, and the annotation object is

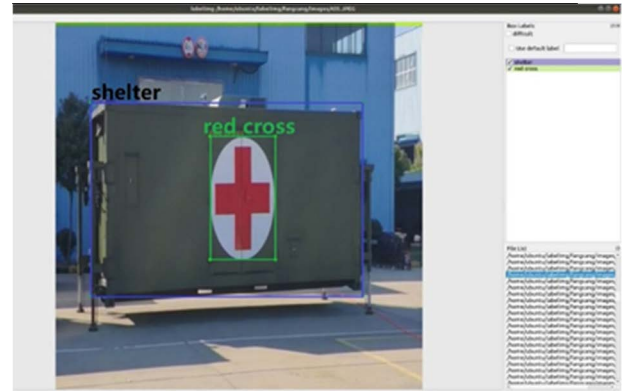


FIGURE 4. Schematic diagram of the data set labeling.

TABLE 1. Correspondence between label.

Category	The shelter	The red cross
Tag name	shelter	red cross
Tag number	0	1

TABLE 2. Index parameters of ten different models.

Model	Depth	Width	Layer	Parameter
YOLOv5n	0.33	0.25	270	1766623
YOLOv5n6	0.33	0.25	355	3096244
YOLOv5s	0.33	0.50	270	7025023
YOLOv5s6	0.33	0.50	355	12326164
YOLOv5m	0.67	0.75	391	21060447
YOLOv5m6	0.67	0.75	481	35281716
YOLOv5l	1.00	1.00	499	46636735
YOLOv5l6	1.00	1.00	607	76170196
YOLOv5x	1.33	1.25	567	86224543
YOLOv5x6	1.33	1.25	733	140045044

the peripheral features of the shelter and the red cross in the middle. YOLO format was selected for annotation. Labeling generate an outer frame in the form of boundary box in the image, and automatically generate a txt file with the same name as the annotation image after the manual annotation result was saved.

B. MODEL SELECTION

Considering comprehensive factors such as training efficiency and detection accuracy, a suitable detection model for shelter-transporting AGV detecting shelter in shelter transport with both better performance and faster speed should be obtained. Therefore, the ten pre-training models of YOLOv5 were chosen to be trained, and the index parameters of the ten pre-training models are shown in Table 2.

The models in Table 2 are listed in ascending order of complexity of network structure. The experimental environment is Ubuntu 18.04 operating system, and based on Pytorch framework. CPU: Intel Core I9-10900K, GPU: NVIDIA RTX

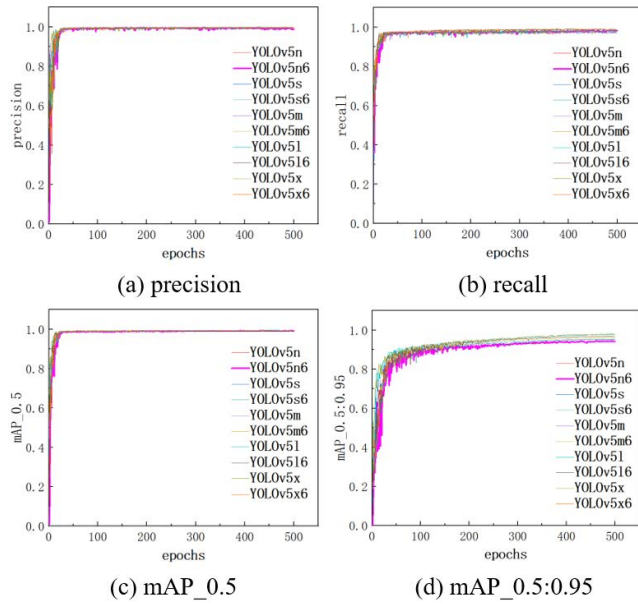


FIGURE 5. Comparison of model training results.

TABLE 3. Experimental training parameters.

Parameter	The numerical
Epochs (Preset training times)	500
Batch-size (Batch)	16
Learning rate	0.01
Momentum term	0.937
Decay (Decay regular term)	0.0005

3090, 24GB. Training parameter settings for ten pre-training models of YOLOv5 are shown in Table 3.

The ten models in Table 2 were trained for a total of 16.059 hours according to the training parameters in Table 3, and the comparison of post-training model parameters are shown.

During the training process, various values change with the number of training steps increasing. The meanings of each value in Figure 5 are as follows:

In Figure.5(a), the precision is equal to the number of correct targets marked divided by the total number of targets marked. The closer to 1, and the higher the accuracy. And in Figure.5(b), the recall rate is equal to the number of correct targets marked divided by the total number of targets that need to be marked. The closer to 1, and the higher the accuracy. In Figure.5(c), the mAP_0.5 (mean Average Precision) represents when IOU is set to 0.5, the AP of all pictures of each category is calculated, and then all categories are averaged. In Figure.5(d), the mAP_0.5:0.95 represents the average mAP at different IOU thresholds (from 0.5 to 0.95 in steps of 0.05).

As can be seen from Figure 5, when the number of training steps reached 200, each value tended to be stable. As the number of training steps reached 500, the curves all achieved a good fitting effect. By comparing the training results, it can be found that both the precision and the recall tend

TABLE 4. Comparison of detection results.

Model	Layer	Parameter size /M	Train time /h	Accuracy / %	Detection time /ms	
YOLOv5n	213	1761871	3.8	0.719	94.24	5.5
YOLOv5n6	280	3089188	6.6	0.857	96.28	6.9
YOLOv5s	213	7015519	14.4	0.813	95.04	5.8
YOLOv5s6	280	12312052	25.1	0.955	96.91	7.6
YOLOv5m	308	21041679	42.6	1.260	95.22	8.3
YOLOv5m6	378	35254692	71.1	1.371	95.88	10.3
YOLOv5l	392	46605951	93.8	1.847	96.46	10.6
YOLOv5l6	476	76126356	153.0	2.002	97.24	12.5
YOLOv5x	444	86180143	173.1	2.929	96.96	14.7
YOLOv5x6	574	139980484	280.9	3.306	96.65	17.6

to reach 1 with the increase of training steps, indicating that all the ten models achieved good training effects, and the mAP_0.5 were also stable around 1 as the number of training steps increased. The mAP_0.5:0.95 increased slowly in the first 100 training sessions, then tended to stabilize and approached 1 slowly, and it was obvious that there is a certain gap in the final stable value for different training models, but the overall results were all greater than 0.9 and the trend was stable.

In order to further analyze the training effect of the model, the ten models after training were tested and compared. The detection results of the trained ten models on the same test set are shown in Table 4.

It can be seen from Table 4 that the more complex structure and more parameters has the longer corresponding training time and the larger weight. From the comparison of performance indexes and detection results of the different trained models, it can be seen that the detection accuracy of relatively complex pre-training models (i.e., YOLOv5m and YOLOv5m6), were not as good as that of relatively simple (i.e., YOLOv5s6 and YOLOv5n6), and the most complex YOLOv5 × 6 was not as good as the simplest YOLOv5s6. Therefore, it can be concluded that detection effect of the complex trained model may not be better in actual applications. YOLOv5n has the minimum depth and width of pre-training model, and the minimum number of model layers, parameters and detection time obtained after training, as well as the minimum weight of model, which is very suitable to deployment on the shelter-transporting AGV. However, it has the lowest detection accuracy of 94.24% compared with other models. The YOLOv5n6 was slightly higher than YOLOv5n on model complexity, but it is nearly 2% higher than YOLOv5n on detection accuracy. Compared to other more complex including YOLOv5m, YOLOv5l and YOLOv5x, the accuracy of the YOLOv5n6 is not low, and detection time of the YOLOv5n6 is only

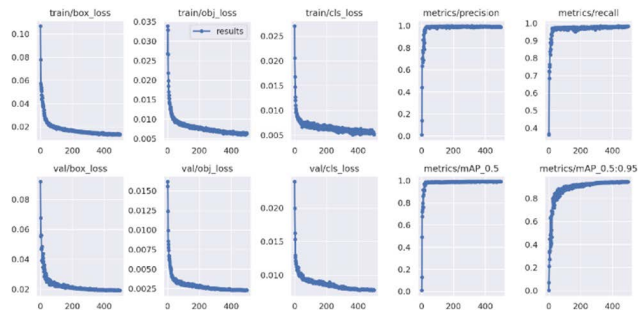


FIGURE 6. YOLOv5n6 training results.

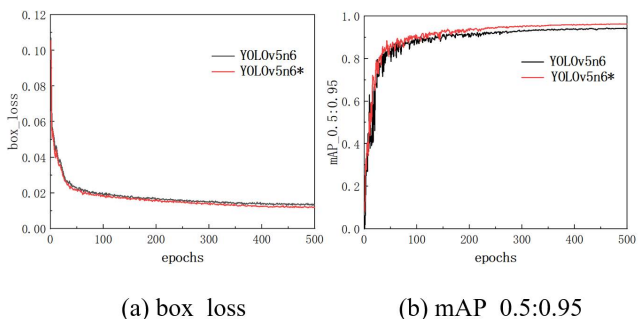


FIGURE 7. Comparison of training results.

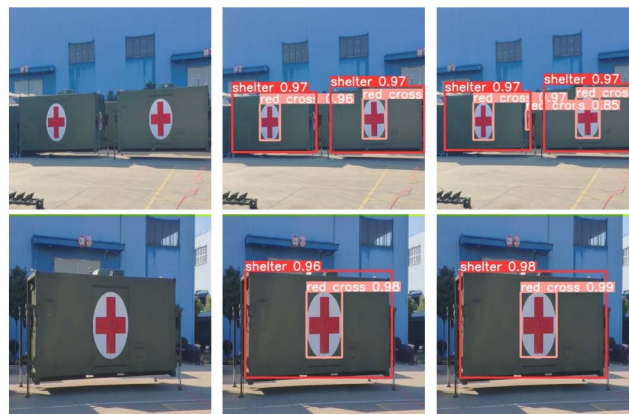
6.9ms, which is far less than the detection time of complex model (i.e., YOLOv5m, YOLOv5l and YOLOv5x) greater than 10ms. The training results of YOLOv5n6 are shown in Figure 6.

According to the YOLOv5n6 training result, the model loss value decreased and tended to be stable with the increase of training steps. The curve fitting state was good, and precision, recall, mAP_{0.5} and mAP_{0.5:0.95} all tended to be stable at 1. Considering the detection accuracy, detection time and model weight, the YOLOv5n6 was selected as the best detection model and applied to shelter detection, which can well achieve the high detection accuracy and speed.

C. DISCUSSION

To solve the problems of identifying shelters with low accuracy and slow speed existing in the process of shelter-transporting AGV shelter detection and further improve the detection performance, the detection model YOLOv5n6* was developed by introducing the CBAM into the model’s main structure, changing the loss function from GIOU_Loss to CIOU_Loss, and selecting a more reasonable DIOU_NMS. The box_loss and mAP_{0.5:0.95} of the YOLOv5n6* and YOLOv5n6 models are shown in Figure 7.

By comparing training results of the two models in Figure 7, it can be seen that box_loss of the improved model YOLOv5n6* and YOLOv5n6 decreased with the increase of training steps and gradually tended to be stable. According to the comparison Figure 7(a), the box_loss of the improved YOLOv5n6* is 1.2% lower than the YOLOv5n6, which meets the requirements of proposed strategy and proves that the improved strategy enables YOLOv5n6* to



(a) original picture (b) YOLOv5n6 (c) YOLOv5n6*

FIGURE 8. Comparison of detection results.

TABLE 5. Comparison of detection results.

Model	Model size /M	Accuracy /%	Detection time /ms
YOLOv5n6	6.6	96.28	6.9
YOLOv5n6*	7.2	97.15	7.1

have higher positioning accuracy. As shown in Figure 7(b), mAP_{0.5:0.95} of YOLOv5n6* and YOLOv5n6 gradually approached 1 during training process, and tended to be stable after 400 training times. The training model reached the fitting. Compared with the YOLOv5n6 before the improvement, the mAP_{0.5:0.95} of the YOLOv5n6* increased by 2%, indicating that the improved model YOLOv5n6* obtained good training results.

In order to further evaluate performance of the improved model YOLOv5n6* and the original model YOLOv5n6 both the YOLOv5n6* and the YOLOv5n6 were tested on the test set. The comparison of detection results of the two models shown in Table 5 and Figure 8.

As can be seen from Table 5, compared with the YOLOv5n6, the detection accuracy of the YOLOv5n6* increased by 0.87%, however, the detection time of the YOLOv5n6* was only increased by 0.2ms. Therefore, the YOLOv5n6* was suitable for the application on a shelter-transporting AGV due to its small size. Figure 8 shows the original picture, and the detection effect pictures by the YOLOv5n6 and the YOLOv5n6*. It can be seen that the detection result will not filter the blocked object from the first row of pictures. However, it can be seen from the second row of pictures that the introduction of CBAM and the change of CIOU_Loss increases the confidence of the detection results. It was proved that the introduction of the attention mechanism and the improvement of the loss function were effective to improve the shelter detection ability and make the YOLOv5n6* model more robustness.

V. CONCLUSION

The shelter detection has the disadvantages of slow recognition speed and low accuracy. Therefore, this paper applies

YOLOv5 to the shelter target detection during the process of the shelter-transporting AGV transshipment shelter. Firstly, a suitable shelter detection model YOLOv5n6 was selected through experiments. And the YOLOv5n6* was proposed by modifying the YOLOv5n6 based on the introduction of the attention mechanism, the application of the loss function CIOU_Loss and the non-maximum suppression function DIOU_NMS. Compared with the YOLOv5n6 model, the box_loss of the improved model YOLOv5n6* is reduced by 1.2%, the mAP_0.5:0.95 is improved by 2%, and the accuracy is improved by 0.87% which shows that YOLOv5n6* was more suitable for shelter-transporting AGV transshipment shelters and has significantly improved detection ability and strong robustness. Under the condition of the model performance improvement, the model size is only 7.2M, and the detection time is only increasing by 0.2ms, indicating that the model YOLOv5n6* can meet the requirements of both identification accuracy and detection speed, and can effectively solve the problems of the low identification accuracy and the detection speed existing in the target detection of shelter-transporting AGV transshipment shelters. In the future, the accuracy and reliability of target detection technology in complex environment need further discussions. An interesting research direction is fusing camera and lidar data to enhance the robustness of current target detection.

REFERENCES

- Z. Yan, I. Lim, and Y. Ge, *Fangcang Shelter Hospitals For COVID-19: Construction and Operation Manual*. Singapore: World Scientific, Jun. 2020.
- Z. Feng, G. Xuan, L. Mengwei, and Z. Ying, "Shelter hospital: Glimmers of hope in treating coronavirus 2019," *Disaster Med. Public Health Preparedness*, vol. 14, no. 5, pp. e3–e4, 2020.
- Z. Pei, Y. Yuan, T. Yu, and N. Li, "Dynamic allocation of medical resources during the outbreak of epidemics," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 2, pp. 663–676, Apr. 2022.
- H. Sheikhbardsiri, G. Khademipour, E. R. Davarani, A. Tavan, H. Amiri, and A. Sahebi, "Response capability of hospitals to an incident caused by mass gatherings in southeast Iran," *Injury*, vol. 53, no. 5, pp. 1722–1726, May 2022.
- R. Neely, K. Haynes, and G. Miller, "Development of a mobile hospital for disaster relief," *JONA, J. Nursing Admin.*, vol. 51, no. 1, pp. 33–37, Jan. 2021.
- J. Wang, T. Xiao, Q. Gu, and Q. Chen, "YOLOv5_CSL_F: YOLOv5's loss improvement and attention mechanism application for remote sensing image object detection," in *Proc. Int. Conf. Wireless Commun. Smart Grid (ICWCSG)*, Aug. 2021, pp. 197–203.
- Q. Fu, J. Chen, W. Yang, and S. Zheng, "Nearshore ship detection on SAR image based on Yolov5," in *Proc. 2nd China Int. SAR Symp. (CISS)*, Nov. 2021, pp. 1–4.
- K. Ding, X. Li, W. Guo, and L. Wu, "Improved object detection algorithm for drone-captured dataset based on yolov5," in *Proc. 2nd Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Jan. 2022, pp. 895–899.
- S. Li, S. Zhang, J. Xue, H. Sun, and R. Ren, "A fast neural network based on attention mechanisms for detecting field flat jujube," *Agriculture*, vol. 12, no. 5, p. 717, May 2022.
- L. Xing, X. Fan, Y. Dong, Z. Xiong, L. Xing, Y. Yang, H. Bai, and C. Zhou, "Multi-UAV cooperative system for search and rescue based on YOLOv5," *Int. J. Disaster Risk Reduction*, vol. 76, Jun. 2022, Art. no. 102972.
- J. Ieamsaard, S. N. Charoensook, and S. Yammen, "Deep learning-based face mask detection using YoloV5," in *Proc. 9th Int. Electr. Eng. Congr. (IEECON)*, Mar. 2021, pp. 428–431.
- H. Gao, W. Wang, C. Yang, W. Jiao, Z. Chen, and T. Zhang, "Traffic signal image detection technology based on Yolo," *J. Phys., Conf. Ser.*, vol. 1961, no. 1, Jul. 2021, Art. no. 012012.
- P. Agrawal, B. Ross Girshick, and J. Malik, "Analyzing the performance of multilayer neural networks for object recognition," 2014, *arXiv:1407.1610*.
- S. Wenxiu and L. Nianqiang, "Application of target detection algorithm based on deep learning in farmland pest recognition," *Int. J. Artif. Intell. Appl.*, vol. 11, no. 3, pp. 1–10, May 2020.
- T. Ray, H. K. Singh, K. H. Rahi, T. Rodemann, and M. Olhofer, "Towards identification of solutions of interest for multi-objective problems considering both objective and variable space information," *Appl. Soft Comput.*, vol. 119, Apr. 2022, Art. no. 108505.
- K. Li, G. Cheng, S. Bu, and X. You, "Rotation-insensitive and context-augmented object detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2337–2348, Apr. 2018.
- S. Chen, Z. Zhang, J. Yang, J. Wang, X. Zhai, T. Bärnighausen, and C. Wang, "Fangcang shelter hospitals: A novel concept for responding to public health emergencies," *Lancet*, vol. 395, no. 10232, pp. 1305–1314, Apr. 2020.
- Z. Ning, X. Wu, J. Yang, and Y. Yang, "MT-YOLOv5: Mobile terminal table detection model based on YOLOv5," *J. Phys., Conf. Ser.*, vol. 1978, no. 1, Jul. 2021, Art. no. 012010.
- P. T. Qu, H. Wu, and X. X. Xu, "Research and design of automatic guided vehicle for field medical shelter transfer," *Mil. Med.*, vol. 46, no. 5, pp. 321–327, 2022.
- Y. Chen, C. Zhang, T. Qiao, J. Xiong, and B. Liu, "Ship detection in optical sensing images based on YOLOv5," in *Proc. 12th Int. Conf. Graph. Image Process. (ICGIP)*, Jan. 2021, Art. no. 11720.
- X. Zhang, H. Fan, H. Zhu, X. Huang, T. Wu, and H. Zhou, "Improvement of YOLOV5 model based on the structure of multiscale domain adaptive network for crowdscape," in *Proc. IEEE 7th Int. Conf. Cloud Comput. Intell. Syst. (CCIS)*, Nov. 2021, pp. 171–175.
- J. Feng, Z. Xu, and Z. Liu, "Improved YOLOv5 with attention mechanism for SAR ship target detection in complex environment," *Int. Core J. Eng.*, vol. 8, no. 6, pp. 290–301, 2022.
- Z. Wei, Z. Qikai, L. Ruizhi, and N. Fu, "Research on camouflage target detection method based on improved YOLOv5," *J. Phys., Conf. Ser.*, vol. 2284, Jun. 2022, Art. no. 012018.
- E. Cengil, A. Cinar, and M. Yildirim, "A case study: Cat-dog face detector based on YOLOv5," in *Proc. Int. Conf. Innov. Intell. Informat., Comput., Technol. (3ICT)*, Sep. 2021, pp. 149–153.
- T. Shilei, Y. Jun, J. Ziqiang, and H. Li, "Approach for improving YOLOv5 network with application to remote sensing target detection," *J. Appl. Remote Sens.*, vol. 15, no. 3, 2021, Art. no. 036512.
- X. Li and J. Li, "Improvement of pedestrian detection algorithm based on Yolo," in *Proc. 2nd Int. Conf. Mech. Eng., Ind. Mater. Ind. Electron. (MEIMIE 2019)*, 2019, pp. 55–60.
- J. Zhou, M. Yan, C. Luo, and X. Xing, "Underwater sonar target detection based on YOLOv5," in *Proc. Int. Conf. Electron. Inf. Eng. Comput. Sci. (EIECS)*, Sep. 2021, pp. 729–732.
- Z.-H. Su and J.-C. Yen, "Cup-to-disk ratio detection of optic disk in fundus images based on YOLOv5," in *Proc. Int. Conf. Electron. Commun., Internet Things Big Data (ICEIB)*, Dec. 2021, pp. 327–329.
- D. P. Carrasco, H. A. Rashwan, M. A. Garcia, and D. Puig, "T-YOLO: Tiny vehicle detection based on Yolo and multi-scale convolutional neural networks," *IEEE Access*, early access, Dec. 22, 2021, doi: 10.1109/ACCESS.2021.3137638.
- L. Kong, J. Wang, and P. Zhao, "YOLO-G: A lightweight network model for improving the performance of military targets detection," *IEEE Access*, vol. 10, pp. 55546–55564, 2022.
- P. Batra, I. Hussain, M. A. Ahad, G. Casalino, M. A. Alam, A. Khaliq, and S. I. Hassan, "A novel memory and time-efficient ALPR system based on YOLOv5," *Sensors*, vol. 22, no. 14, p. 5283, Jul. 2022.
- E. Guney, C. Bayilmis, and B. Cakan, "An implementation of real-time traffic signs and road objects detection based on mobile GPU platforms," *IEEE Access*, vol. 10, pp. 86191–86203, 2022.
- L. Aziz, M. S. B. Haji Salam, U. U. Sheikh, and S. Ayub, "Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review," *IEEE Access*, vol. 8, pp. 170461–170495, 2020.
- S. Verma, S. Tripathi, A. Singh, M. Ojha, and R. R. Saxena, "Insect detection and identification using Yolo algorithms on soybean crop," in *Proc. TENCON IEEE Region 10 Conf. (TENCON)*, Dec. 2021, pp. 272–277.

[35] W. Jia, S. Xu, Z. Liang, Y. Zhao, H. Min, S. Li, and Y. Yu, "Real-time automatic helmet detection of motorcyclists in urban traffic using improved YOLOv5 detector," *IET Image Process.*, vol. 15, no. 14, pp. 3623–3637, Dec. 2021.

[36] L. Zhu, X. Geng, Z. Li, and C. Liu, "Improving YOLOv5 with attention mechanism for detecting boulders from planetary images," *Remote Sens.*, vol. 13, no. 18, p. 3776, Sep. 2021.

[37] J. Wang, T. Xiao, Q. Gu, and Q. Chen, "YOLOv5_CSL_F: YOLOv5's loss improvement and attention mechanism application for remote sensing image object detection," in *Proc. Int. Conf. Wireless Commun. Smart Grid (ICWCSG)*, Aug. 2021, pp. 197–203.

[38] Q. Fu, J. Chen, W. Yang, and S. Zheng, "Nearshore ship detection on SAR image based on Yolov5," in *Proc. 2nd China Int. SAR Symp. (CISS)*, Nov. 2021, pp. 1–4.

[39] K. Ding, X. Li, W. Guo, and L. Wu, "Improved object detection algorithm for drone-captured dataset based on yolov5," in *Proc. 2nd Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Jan. 2022, pp. 895–899.

[40] Z. Wu, D. Zhang, Y. Shao, X. Zhang, X. Zhang, Y. Feng, and P. Cui, "Using YOLOv5 for garbage classification," in *Proc. 4th Int. Conf. Pattern Recognit. Artif. Intell. (PRAI)*, Aug. 2021, pp. 35–38.

[41] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 3–19.



HANG WU was born in 1991. He is an Assistant Engineer with the Institute of Medical Support Technology, Academy of System Engineering, and the Academy of Military Sciences. His research interests include mobile robotics and flexible robotic arms.



XINXI XU is currently a Senior Research Fellow and a Ph.D. Advisor with the Department of Medical Support Technology, Institute of Systems Engineering, Academy of Military Sciences. His research interests include mobile field hospital and intelligent medical platform. He is a Council Member of Chinese Society for Vibration Engineering and Chinese Society of Biomedical Engineering. He received three second prizes of the National Science and Technology Progress Award and four first prizes of the Ministerial and Provincial-Level Science and Technology Awards.



DIAN YANG was born in 1998. He is currently pursuing the master's degree with the Institute of Medical Support Technology, Academy of System Engineering, and the Academy of Military Sciences. His research interests include intelligent mobile medical and health equipment.



CHEN SU was born in 1981. He is a Senior Engineer with the Institute of Medical Support Technology, Academy of System Engineering, and the Academy of Military Sciences. His research interests include health care and security devices.



XIUGUO ZHAO was born in 1979. He is a Senior Engineer with the Institute of Medical Support Technology, Academy of System Engineering, and the Academy of Military Sciences. His research interests include intelligent mobile robot and multi-intelligent body collaboration.

...