

Research Article

Target Detection of Low-Altitude UAV Based on Improved YOLOv3 Network

Haiqing Zhai ^{1,2} and Yang Zhang ^{1,2}

¹School of Computer Science and Technology, Henan Institute of Technology, Xinxiang, Henan 453003, China

²Big Data Engineering Research Center of Henan for Production and Manufacturing IoTs, Xinxiang, Henan 453003, China

Correspondence should be addressed to Haiqing Zhai; zhaihaiqing@hait.edu.cn

Received 7 January 2022; Accepted 18 February 2022; Published 10 March 2022

Academic Editor: Shan Zhong

Copyright © 2022 Haiqing Zhai and Yang Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most existing methods are difficult to detect low-altitude and fast-moving drones. A low-altitude unmanned aerial vehicle (UAV) target detection method based on an improved YOLOv3 network is proposed. While keeping the basic framework of the original model unchanged, the YOLOv3 model is improved. That is, multiscale prediction is added to enhance the detection ability of small-target objects. In addition, the two-axis Pan/Tilt/Zoom (PTZ) camera is controlled based on proportional integral derivative (PID), so that the target tends to the center of the field of view. It is more conducive to accurate detection. Finally, experiments are carried out using real UAV datasets. The results show that the mean average precision (mAP), AP50, and AP75 are 25.12%, 39.75%, and 26.03%, respectively, which are better than other methods. Also, the frame rate is 21 frames·s⁻¹, which meets the performance requirements.

1. Introduction

With the rise and development of UAV technology, it has been widely used in military and civil fields. However, a large number of UAVs pose a certain threat to the flight safety of aircraft and the political sensitivity of images in confidential areas. At the same time, it also brings huge challenges to urban security [1]. For the sake of public safety, local governments prohibit unauthorized UAV flight in airports, meeting places, and other areas [2]. Therefore, monitoring UAVs in specific areas is an urgent need for security. Due to the small volume and low speed of UAV, it is difficult to detect UAV by using traditional radar equipment [3]. In addition, the noise of the city is noisy, so it is difficult to detect UAVs by acoustic sensors [4]. Therefore, how to detect UAV efficiently and accurately from the environment has become an urgent problem to be solved.

So far, the identification methods of UAVs are diverse. The main methods focus on image recognition and radar data analysis [5]. However, traditional analysis methods that rely on spectrum detection and radar data are extremely

susceptible to interference from external environmental factors. The method based on image recognition is limited by the bottleneck of computer technology and communication technology and has not been widely used and developed [6, 7]. At present, there has been a certain amount of research on UAV target recognition in China. For example, Thillainayagi and Senthil Kumar [8] proposed a target detection technology for UAV thermal images based on wavelet transform and singular value decomposition, using discrete wavelet transform and stationary wavelet transform to enhance image texture features and edge features. The experimental results show that the proposed method has smaller errors. However, the detection efficiency needs to be improved. Li et al. [9] proposed a fast and effective moving target detection method, which extracts cross features based on line segments. It has good detection speed and rotation accuracy, but the application scope of the algorithm is small and the practical application has great limitations. Abdulridha et al. [10] proposed a UAV hyperspectral image recognition method. Multilayer perceptual neural network and stepwise discriminant analysis are used to realize the

image detection, which effectively improves the detection accuracy. Yang et al. [11] proposed a UAV object detection model based on rotation and constant depth denoising in view of the difficulties of the multidirectional object, small pixel, and vibration interference of the UAV body in the process of aerial object detection. A selective search method is used to extract the region of interest in the aerial image, and the radial gradient of the region of interest is calculated. Combined with the deep denoising autoencoder, the original data noise is filtered out and the deep features are extracted to realize the detection of aerial image targets. However, deep-level feature extraction will increase the amount of calculation and affect the detection speed of the algorithm.

In addition, image-based target detection technology has gradually been widely used with the significant improvement of computing power and communication technology [12]. Xiaofei [13] proposes a UAV multitarget tracking and path planning method combining basic gray wolf optimizer and Gaussian distribution estimation. It overcomes the problem of real-time optimization of complex projects with traditional models and has good effectiveness and practicability. However, the algorithm focuses on target tracking and path planning, and its performance in target detection needs to be improved. Tao et al. [14] proposed a target search strategy for UAV based on reinforcement learning. Through the reinforcement learning training, the image captured by the drone is analyzed and processed to achieve target detection and tracking, but the detection effect of maneuvering targets is poor. Liu and Zhang [15] proposed an automatic vehicle detection method based on deep learning. Based on the interactive multimodel particle filter algorithm, the performance of the UAV for maneuvering target positioning has been significantly improved. However, the detection effect of small moving targets such as drones has yet to be verified.

Based on the above analysis, a target detection algorithm based on the improved YOLOv3 network is proposed for the problem of low-altitude UAV target detection. The innovations are summarized as follows:

- (1) In order to make up for the inability of the YOLOv3 network to detect small targets, the proposed method adds multitarget prediction to improve the YOLOv3 network. Four different sizes of bounding boxes are provided to match the actual bounding box as much as possible, thereby improving the accuracy of target detection.
- (2) Since drone monitoring is costly and difficult to implement, the proposed method uses a two-axis PTZ camera to track the drone. Among them, the PID algorithm is used to adjust the camera position to achieve efficient detection of low-altitude UAVs.

2. Target Detection Based on Improved YOLOv3

2.1. YOLOv3 Network Structure. The YOLOv3 algorithm uses a fully convolutional network composed of residual blocks as the backbone network. The network depth reaches 53 layers and is called Darknet-53. Its network structure is shown as in Figure 1. YOLOv3 draws on the idea of Feature

	Type	filters	size	output
	Convolutional	32	3*3	256*256
	Convolutional	64	3*3/2	128*128
1×	Convolutional	32	1*1	
	Convolutional	64	3*3	
	Residual			128*128
	Convolutional	128	3*3/2	64*64
2×	Convolutional	64	1*1	
	Convolutional	128	3*3	
	Residual			64*64
	Convolutional	32	3*3	256*256
8×	Convolutional	128	1*1	
	Convolutional	256	3*3	
	Residual			32*32
	Convolutional	512	3*3/2	16*16
8×	Convolutional	256	1*1	
	Convolutional	512	3*3	
	Residual			16*16
	Convolutional	1024	3*3/2	8*8
4×	Convolutional	512	1*1	
	Convolutional	1024	3*3	
	Residual			8*8
	Avgpool		global	
	Connected		1000	
	Softmax			

FIGURE 1: Network structure of YOLOv3 algorithm.

pyramid network (FPN) and uses multiscale features for target detection. On the premise of maintaining the speed advantage, the detection accuracy is further improved, especially the detection ability for small targets is strengthened. Also, the detection effect of high-coverage images is significantly higher than that of YOLOv2 [16].

The YOLO series of algorithms constantly iteratively optimize the detection accuracy on the basis of always maintaining the advantage of high detection speed and gradually improve the detection ability of small targets. At the same time, accurate detection of high-coverage images is realized, the structure is simple, and the background false detection rate is low. It has become the most popular real-time target detection algorithm today.

2.2. Improved YOLOv3. In the detection and recognition process, when two or more types of similar objects appear, the YOLO network will often misidentify them as the same type. This shows that the YOLO network is too poor to distinguish the details. At the same time, the recognition performance of YOLO for objects with large difference in length and width ratio needs to be improved [17]. When the proportion of large and small objects in the image differs greatly, the large objects can be identified in the test results, but the small objects often cannot be detected [18, 19]. The UAV remote sensing image used in the experiment has a lower resolution, a larger image scale, rich information, and more complex details. Direct use of conventional network models for detection and recognition is often not ideal [20, 21]. Therefore, the model algorithm needs to be further improved according to the characteristics of UAV images.

For the above reasons and the analysis of the YOLO model, the YOLOv3 model is improved while maintaining the basic framework of the original model, that is, multiscale prediction is added.

In the YOLOv2 network model, in order to enhance the accuracy of small-target detection, the feature map extracted from the last layer of the network model is connected with the feature map of the previous layer through the pass-through layer. The size of the feature map of the last layer is 13×13 . In YOLOv3, this method is further enhanced. YOLOv3 provides three bounding boxes of different sizes, using similar concepts to extract features of these sizes to form a pyramidal network. YOLOv3 adds several convolutional layers, and the final convolution layer is used to predict the tensor coding including boundary box, target in the box, and classification prediction [22]. YOLOv3 uses the feature fusion of multiple scales, so the number of bounding boxes is much more than before.

Adding a scale prediction to the YOLOv3 network model provides four bounding boxes of different sizes, as shown in Figure 2. The improved YOLOv3 network includes the last layer, and the feature map is 13×13 . There are also 3 upsampled eltwise sums with feature maps of 26×26 , 52×52 , and 104×104 . The largest network model uses the 104 feature map. Also, YOLOv2 takes multiscale into consideration for the training data sampling, and in the end, only the feature map of 13 is used. This should be the place that has the greatest impact on small goals [23]. At the same time, the input image size is changed to 1024×1024 to adapt to large-scale sampling. In the experiment, 12 clusters were selected. Then, divide the dimensional clusters evenly on bounding boxes of different sizes. The 12 clusters are (10×13) , (16×30) , (33×23) , (30×61) , (62×45) , (59×119) , (116×90) , (156×198) , (373×326) , (312×536) , (584×712) , and (869×796) .

2.3. Detection Process. The improved YOLOv3 algorithm does not need to generate a region of interest (ROI) in advance but directly trains the network in a regression way. At the same time, the k-means algorithm is used to cluster the sample bounding boxes, and four groups of bounding box sizes are preset on the four scale sizes respectively, so as to make positioning prediction based on the bounding boxes of 16 sizes. The whole detection process is shown in Figure 3.

First, feature extraction is performed on the original 1024×1024 input image through the feature extraction network. Then, the feature vector is fed into the FPN structure to generate grid areas on 4 scales. They are 13×13 , 26×26 , 52×52 , and 104×104 . Each grid area predicts 4 bounding boxes, resulting in a total of $(104 \times 104 + 52 \times 52 + 26 \times 26 + 13 \times 13) \times 4 = 57460$ bounding boxes. Next, a vector Ω is predicted in each bounding box, which is expressed as follows:

$$\Omega = (t_x + t_y + t_w + t_h) + l_0 + (l_1 + l_2 + \dots + l_n). \quad (1)$$

The first 4 elements t_x , t_y , t_w , and t_h in the vector Ω are the 4 coordinates related to the bounding box, and their relationship is as follows:

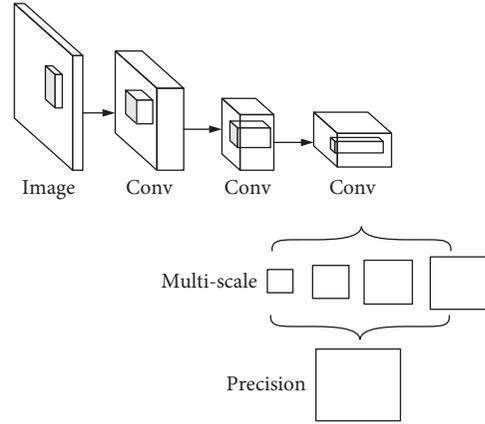


FIGURE 2: Multiscale prediction.

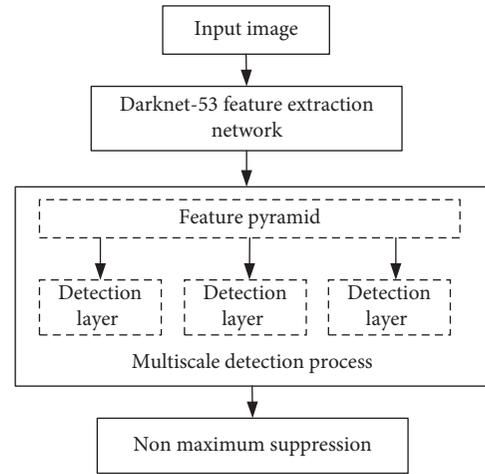


FIGURE 3: Detection process of YOLOv3 algorithm.

$$\begin{aligned} p_x &= \text{sigmoid}(t_x) + b_x, \\ p_y &= \text{sigmoid}(t_y) + b_y, \\ p_w &= g_w \times e^{t_w}, \\ p_h &= g_h \times e^{t_h}, \end{aligned} \quad (2)$$

where b_x and b_y represent the offset of the grid to which the bounding box belongs relative to the upper left corner of the image. g_h and g_w represent the length and width of the predefined bounding box. p_x and p_y indicate the distance from the center of the bounding box of the final prediction result to the upper left corner of the image. p_w and p_h are the length and width of the predicted bounding box.

The l_0 of the vector Ω is expressed as follows:

$$l_0 = \Theta(o) \times \text{IoU}_o^{gt}, \quad (3)$$

where $\Theta(o)$ represents the probability that the object is in the prediction frame. IoU_o^{gt} represents the intersection over union (IoU) of the predicted box and the true bounding box. When using logistic regression to score the prediction box the highest, the probability that the object is in the prediction box is 1, otherwise it is 0. The l_0 in vector Ω represents the

score that the predicted object belongs to one of the classes. When the prediction frame is obtained, the nonmaximum suppression is carried out to obtain the final prediction result.

3. Design of UAV Vision following Control Algorithm

The two-axis PTZ camera is shown in Figure 4. The function of the camera is to control the movement of the camera to keep the target in the center of the video. The control module is a two-axis PTZ with two steering gears. One servo is responsible for controlling the camera to move left and right, and the other controls the camera to move up and down. Each steering gear has an adjustment range of 180° .

The PID control algorithm is expressed as follows:

$$q(t) = \kappa_p \theta(t) + \kappa_i \int_0^t \theta(t') dt' + \kappa_d \frac{d\theta(t)}{dt}, \quad (4)$$

where $q(t)$ is the output of the system, which means the steering gear rotation angle, rad; $\theta(t)$ is the deviation angle between the image center and the UAV center, rad; κ_p , κ_i , and κ_d are all constant coefficients, corresponding to proportional gain, integral gain, and differential gain [24, 25].

Equation (4) is composed of three parts: proportional, integral, and differential. The first part makes the camera rotate with the movement of the drone. The integral part is used to eliminate the stabilization error and prevent the drone from being out of the video center. The derivative part is used to control the rate of change of the deviation [26].

The PID control process is shown in Figure 5. Using OpenCV (Computer Vision Library) to process the camera video stream, improve YOLOv3 detection for each frame of the video stream. After obtaining the position of the drone in the picture, calculate the distance between its center and the center of the picture. The distance parameter is passed to the PID process for calculation, so as to control the rotation of the steering gear.

4. Experiment and Analysis

The experimental platform is a computer with Intel core i7-7700HQ 2.8 GHz CPU and GeForce 1050 ti 2 GB GPU. The minibatch during training is set to 5, and the learning rate is 0.001.

4.1. Dataset. There are few low-altitude UAV detection and recognition methods based on deep learning, and there is no public dataset or standard dataset. Therefore, firstly, the dataset is collected and constructed.

4.1.1. Data Collection. Using visible light detectors and two-axis pan-tilt cameras to take images of 4 types of civilian drones at different times and in different backgrounds, the UAV models are DJI-Elf 3 (DJ-3), DJI-Yu Pro (DJ-Pro), DJI-Yu Mavic 2 zoom version (DJ-M2 Z), and DJI-Yu Air (DJ-Air). In order to ensure the diversity of data, the various flight attitudes of low-altitude drones, including hovering,



FIGURE 4: Two-axis PTZ camera.

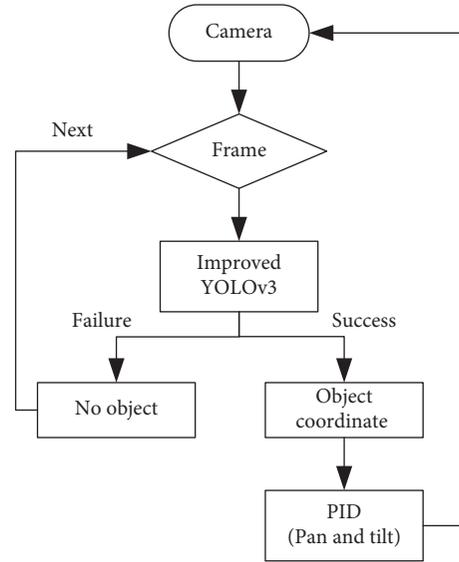


FIGURE 5: PID control process.

rapid ascent and descent, and smooth flight, were fully considered during the shooting process. In the end, 3864 visible light images were obtained.

4.1.2. Data Annotation. In order to ensure the validity of the data, the manual labeling method is adopted to label the samples whose target occluded area is greater than or equal to 50%. Finally, the bounding boxes of the low-altitude drone targets in the 3258 visible light images were labeled, and data with labeled information were obtained. According to the ratio of 5:1, it is divided into the training set and test set.

4.1.3. Image Enhancement. In order to improve the detection and recognition accuracy of the method, the general image enhancement method in the field of target detection is

adopted to enhance the training set, including the operation processing of brightness and contrast.

4.1.4. Data Expansion. Taking into account that the attitude of the UAV during flight is not completely horizontal and inclined, etc., the training set is flipped and rotated at $\pm 10^\circ$ and $\pm 20^\circ$. If the target near the edge in the image is damaged or completely lost after the rotation processing, the sample data are discarded.

4.1.5. Dataset Construction. Through the enhancement and expansion of the images in the training set, three types of datasets are obtained. As shown in Table 1, the training set of UAV-A is composed of low-altitude UAV targets in the original image. The training set of UAV-B is composed of the original image (UAV-A) and the image after image enhancement processing. The training set of UAV-C is composed of UAV-A and UAV-B and their expanded samples (including data sets UAV-A and UAV-B). Among them, the test set uses the same sample for method verification.

4.2. Visual Control Field Test Experiment. The target UAV hovers at the center of the field of view of the cooperative UAV at a relatively long initial distance to make $\tilde{x}(t) = \tilde{x}'$, $\Delta(t) < \Delta'$. In order to quantitatively measure the size error between the actual bounding box and the expected bounding box, the bounding box size error is defined as $\vartheta_\Delta(t) = (\Delta(t) - \bar{\Delta})/\bar{\Delta}$. The experimental result of step response is shown in Figure 6.

It can be seen from Figure 6 that when $t = 2.5$ s, the $v(t)$ step response curve quickly stabilizes. However, there is still a certain steady-state error between the actual bounding box and the expected bounding box.

4.3. Comparison with YOLOv3 Algorithm Classification Effect. In order to describe the classification ability of the improved YOLOv3 algorithm for the dataset, the classification confusion matrix on the data set is calculated. The YOLOv3 classification confusion matrix and the improved YOLOv3 classification confusion matrix are shown in Tables 2 and 3, respectively.

It can be seen from Tables 2 and 3 that DJ-Air has the best detection effect. The accuracy obtained by YOLOv3 and improved YOLOv3 algorithm reached 92.74% and 93.26%, respectively. Compared with other types of UAVs, DJ-Air has obvious characteristics, irregular shape, and easy to distinguish. The detection accuracy of YOLOv3 algorithm for DJ-3, DJ-Pro, and DJ-M2 Z is 88.69%, 92.41%, and 89.91%, respectively. The colors of DJ-3, DJ-Pro, and DJ-M2 Z are not obvious, and the characteristics have certain similarities. In addition, the DJ-Pro and DJ-Air detection results are better in the modified YOLOv3 detection results, with the accuracy reaching 92.51% and 93.26%. Mainly because these categories are quite different from categories other than themselves. Compared with other types of UAVs, the target features are obvious and easier to distinguish. The

detection accuracy of DJ-3 and DJ-M2 Z are 91.74% and 91.98%, respectively. The image performance characteristics of DJ-3 and DJ-M2 Z are similar, so they are easy to confuse.

According to Tables 2 and 3, the classification effect of the YOLOv3 algorithm and the improved YOLOv3 algorithm on the dataset are obtained, as shown in Table 4.

It can be seen from Table 4 that the improved YOLOv3 algorithm has a certain improvement in the recognition and detection effect compared with the classic YOLOv3 algorithm. Compared with the YOLOv3 algorithm, the average detection accuracy rate is increased by about 1.5%. The improved YOLOv3 algorithm adds multiscale prediction, which can strengthen the detection of small targets. Therefore, DJ-Pro, DJ-M2 Z, etc., can be better distinguished.

4.4. Performance Comparison with Comparison Algorithm.

In order to demonstrate the performance of the proposed method, comparison with [8, 11], and [14] was performed. Among them, the comparison experiment stage evaluates the models obtained by different mainstream target detection methods through the same data training. In the experiment, the average accuracy mAP, AP^{50} , and AP^{75} , and frame rate evaluation indicators were used to perform quantitative analysis of detection accuracy and detection speed. The results are shown in Table 5. Among them, AP^{50} is an effective index to evaluate the classification ability of the algorithm. The AP^{75} can reflect the ability of the detection frame to return to the position of the bounding box.

It can be seen from Table 5 that the proposed method has been greatly improved, and its mAP, AP^{50} , and AP^{75} are 25.12%, 39.75%, and 26.03%, respectively. This shows that the proposed improved YOLOv3 target detection framework shows better classification ability and higher frame regression accuracy. The main reason for the obvious improvement of detection accuracy is the use of multilevel feature maps with different scales for target prediction. This greatly improves the detection effect of various targets that change with the drone's viewing angle and flying height. In addition, multiscale target prediction can predict the position and shape of the candidate frame based on image features and generates sparse and arbitrary-shaped candidate frames, which more closely match the real target frame. In addition, the two-axis PTZ camera is controlled based on PID, so that the target tends to the center of the field of view, which is more conducive to target recognition. Thillainayagi and Senthil Kumar [8] realize target detection in UAV thermal image based on wavelet transform and singular value decomposition. The detection model is simple and easy to implement, but the detection accuracy in a complex environment is not high. Its mAP is only 16.14%. Yang et al. [11] proposed a UAV object detection model based on rotation invariant depth denoising. A deep denoising autoencoder is used to filter out the noise of the original data and extract the deep features to realize the target detection of aerial images. Compared with reference [8], its detection accuracy has been improved. However, it is difficult to accurately detect small target objects, and it is easy to cause

TABLE 1: Processing methods and sample numbers of different training sets.

Train dataset	Processing method	Number of samples
UAV-A	Original	3200
UAV-B	Original + image enhancement	6400
UAV-C	Original + image enhancement + data augmentation	65000

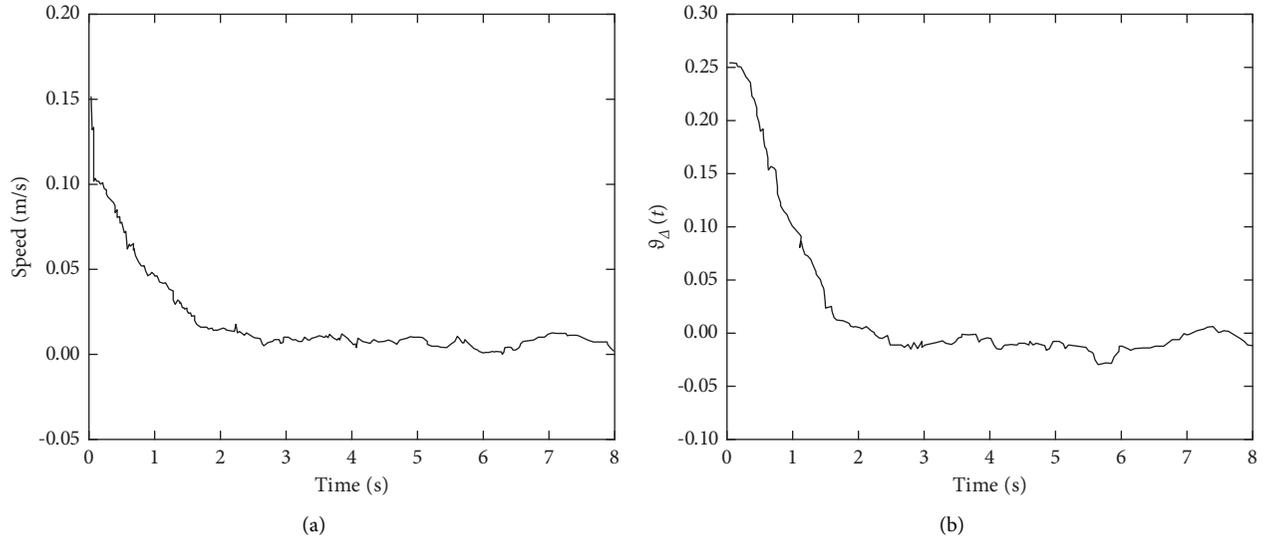


FIGURE 6: Experimental results of step response. (a) Step response curve of flight speed. (b) Boundary frame size error variation curve.

TABLE 2: Confusion matrix of YOLOv3 classification results.

Category	DJ-3	DJ-Pro	DJ-M2 Z	DJ-Air
DJ-3	6382	49	187	21
DJ-pro	274	2336	141	39
DJ-M2 Z	401	108	3625	67
DJ-air	139	35	79	1239
Accuracy rate, %	88.69	92.41	89.91	92.74

TABLE 3: Confusion matrix of improved YOLOv3 classification results.

Category	DJ-3	DJ-Pro	DJ-M2 Z	DJ-Air
DJ-3	6701	43	102	10
DJ-Pro	236	2286	117	26
DJ-M2 Z	295	107	3306	59
DJ-Air	72	35	69	1314
Accuracy rate, %	91.74	92.51	91.98	93.26

TABLE 4: Classification accuracy of two algorithms on datasets.

Category	YOLOv3	Improved YOLOv3
DJ-3	88.69	91.74
DJ-Pro	92.41	92.51
DJ-M2 Z	89.91	91.98
DJ-Air	92.74	93.26
Mean value of accuracy, %	90.96	92.37

TABLE 5: Comparative experimental results of different detection methods on UAV aerial photography data.

Method	Ref. [8]	Ref. [11]	Ref. [14]	Proposed method
mAP, %	16.14	22.37	23.64	25.12
AP ⁵⁰ , %	26.60	34.29	38.08	39.75
AP ⁷⁵ , %	13.62	19.65	24.71	26.03
Frame rate, frames-s ⁻¹	12	19	36	21

confusion. Its AP⁵⁰ is 34.29%. Tao et al. [14] use reinforcement learning to search for drones, and the detection process is easy to implement. Therefore, the detection speed is higher than the 21 frames-s⁻¹ of the proposed method. Its frame rate is as high as 36 frames-s⁻¹. However, due to the lack of PTZ camera assistance, it is not effective in detecting fast-moving targets. Its AP⁷⁵ is 1.32% lower than the proposed method.

5. Conclusion

The rapid development and application of UAV not only brings convenience to the society but also poses a serious threat to public security, personal privacy, and military security. It is becoming more and more important to find unknown UAVs quickly and accurately. Therefore, a target detection method of low-altitude UAV based on improved YOLOv3 network is proposed. The improved YOLOv3 network is used to extract the boundary box of the target UAV. The coordinate error and size error between the boundary box and the desired boundary box are calculated and input into the PID algorithm. Thus, the center moving angle of the two-axis PTZ camera is controlled to realize the accurate detection of the target UAV. The experimental results based on the constructed dataset show that:

- (1) The improved YOLOv3 network has better detection capabilities for UAV targets and is easy to distinguish similar objects. Compared with the YOLOv3 algorithm, the average detection accuracy rate is increased by about 1.5%.
- (2) The proposed method uses the improved YOLOv3 network to obtain the bounding box, and combines the PID algorithm to control the two-axis PTZ camera, which can better obtain the position of the drone. Its mAP, AP⁵⁰, and AP⁷⁵ are 25.12%, 39.75%, and 26.03%, respectively, and the frame rate is 21 frames-s⁻¹. The overall detection performance is the most ideal.

In the next work, we will obtain more drone data sets in scenarios for model training. A model with a faster speed and a smaller footprint will be selected for UAV identification. It will allow the model to get rid of the dependence on the graphics card, making the cost of the system more affordable.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] D. Chen, S. Sun, Z. Lei, H. Shao, and Y. Wang, "Ship target detection algorithm based on improved YOLOv3 for maritime image," *Journal of Advanced Transportation*, vol. 2021, p. 11, Article ID 9440212, 2021.
- [2] J. Abdulridha, Y. Ampatzidis, and S. C. Kakarla, "Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques," *Precision Agriculture*, vol. 21, no. 5, pp. 1–24, 2020.
- [3] Z. Geng, R. Xu, and H. Deng, "LTE-based multistatic passive radar system for UAV detection," *IET Radar, Sonar & Navigation*, vol. 14, no. 7, pp. 1088–1097, 2020.
- [4] J. Yu, Y. Liu, Y. Bai, and F. Liu, "A double-threshold target detection method in detecting low slow small target," *Procedia Computer Science*, vol. 174, no. 5, pp. 616–624, 2020.
- [5] X. Chen, J. Guan, G. Wang, H. Ding, and Y. Huang, "Fast and refined processing of radar maneuvering target based on hierarchical detection via sparse fractional representation," *IEEE Access*, vol. 7, no. 7, pp. 149878–149889, 2019.
- [6] T. Shi, H. Wang, W. Cui, and L. Ren, "Indoor space target searching based on EEG and EOG for UAV," *Soft Computing*, vol. 23, no. 21, pp. 11199–11215, 2019.
- [7] L. Xiaoping, L. Songze, and Z. Boxing, "Fast aerial UAV detection using improved inter-frame difference and SVM," *Journal of Physics: Conference Series*, vol. 1187, no. 2, pp. 032082–032093, 2019.
- [8] R. Thillainayagi and K. Senthil Kumar, "Combination of wavelet transform and singular value decomposition-based contrast enhancement technique for target detection in UAV reconnaissance thermal images," *Journal of Modern Optics*, vol. 66, no. 6, pp. 606–617, 2019.
- [9] Z. Li, C. Meng, F. Zhou et al., "Fast vision-based autonomous detection of moving cooperative target for unmanned aerial vehicle landing," *Journal of Field Robotics*, vol. 36, no. 1, pp. 34–48, 2019.
- [10] J. Abdulridha, Y. Ampatzidis, S. C. Kakarla, and P. Roberts, "Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques," *Precision Agriculture*, vol. 21, no. 5, pp. 955–978, 2020.
- [11] F. Yang, B. Ma, J. Wang, H. Gao, and Z. Liu, "Target detection of UAV aerial image based on rotational invariant depth denoising automatic encoder," *Xibeigongye Daxue Xuebao/ Journal of Northwestern Polytechnical University*, vol. 38, no. 6, pp. 1345–1351, 2020.
- [12] B. Jla, J. A. Rong, and L. A. Wei, "High precision detection algorithm based on improved RetinaNet for defect recognition of transmission lines - ScienceDirect," *Energy Reports*, vol. 6, no. 2, pp. 2430–2440, 2020.

- [13] W. Xiaofei, "A grey wolf optimizer using Gaussian estimation of distribution and its application in the multi-UAV multi-target urban tracking problem - ScienceDirect," *Applied Soft Computing*, vol. 78, no. 8, pp. 240–260, 2019.
- [14] X. Tao, S. Hou, and Y. Wang, "Feasibility of reinforcement learning for UAV-based target searching in a simulated communication denied environment," *Scientia Sinica Informationis*, vol. 50, no. 3, pp. 375–395, 2020.
- [15] X. Liu and Z. Zhang, "A vision-based target detection, tracking, and positioning algorithm for unmanned aerial vehicle," *Wireless Communications and Mobile Computing*, vol. 2021, no. 7, pp. 1–12, 2021.
- [16] O. Doukhi, S. Hossain, and D.-J. Lee, "Real-time deep learning for moving target detection and tracking using unmanned aerial vehicle," *Journal of Institute of Control, Robotics and Systems*, vol. 26, no. 5, pp. 295–301, 2020.
- [17] H.-J. Lim, S.-H. Choi, J. Oh, B. S. Kim, S. Kim, and J. H. Yang, "Adaptive ground control system of multiple-UAV operators in a simulated environment," *Aerospace Medicine and Human Performance*, vol. 90, no. 10, pp. 841–850, 2019.
- [18] S. Pérez-Carabaza and J. Sch Erer, "UAV trajectory optimization for Minimum Time Search with communication constraints and collision avoidance," *Engineering Applications of Artificial Intelligence*, vol. 85, no. 11, pp. 357–371, 2019.
- [19] O. Kechagias-Stamatis and N. Aouf, "A new passive 3-D automatic target recognition architecture for aerial platforms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 1, pp. 406–415, 2019.
- [20] D. Cavaliere, V. Loia, A. Saggese, S. Senatore, and M. Vento, "A human-like description of scene events for a proper UAV-based video content analysis," *Knowledge-Based Systems*, vol. 178, no. 8, pp. 163–175, 2019.
- [21] A. A. d. Santos, J. Marcato Junior, M. S. Araújo et al., "Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs," *Sensors*, vol. 19, no. 16, pp. 3595–3603, 2019.
- [22] M. Rosamilia, A. D. Maio, and A. Aubry, "Single-pulse simultaneous target detection and angle estimation in a multichannel phased array radar," *IEEE Transactions on Signal Processing*, vol. 68, no. 7, pp. 6649–6664, 2020.
- [23] Z. Li, Q. Zhang, and T. Long, "Ship target detection and recognition method on sea surface based on multi-level hybrid network," *Journal of Beijing Institute of Technology (Social Sciences Edition)*, vol. 30, no. 2, pp. 1–10, 2021.
- [24] E. Dong, B. Han, and H. Jian, "Moving target detection based on improved Gaussian mixture model considering camera motion," *Multimedia Tools and Applications*, vol. 79, no. 11, pp. 7005–7020, 2020.
- [25] M. Li, G. Sun, J. Tong, and Z. He, "Training-free moving target detection with uncertain a priori knowledge for airborne radar," *IET Radar, Sonar & Navigation*, vol. 14, no. 3, pp. 372–380, 2020.
- [26] A. Tz, A. Zp, and W. A. Hao, "Infrared small target detection via self-regularized weighted sparse model - ScienceDirect," *Neurocomputing*, vol. 420, no. 3, pp. 124–148, 2021.