FARMLAND OBSTACLE RECOGNITION BASED ON IMPROVED FASTER R-CNN
/
基于改进 FASTER R-CNN 的农田障碍物识别

Xiangyu BAI[1,2], Kai ZHANG[2], Ranbing YANG[1,3], Zhiguo PAN[1], Huan ZHANG[1], Jian ZHANG[*1,3],
Xidong JING[1], Shiteng GUO[1], Sen DUAN[1]
[1] College of Electrical and Mechanical Engineering, Qingdao Agricultural University, Qingdao/ China
[2] National Key Laboratory of Intelligent Agricultural Power Equipment, Henan/ China
[3] College of Mechanical and Electrical Engineering, Hainan University, Haikou/ China
Tel: +8618561879013; E-mail: zhangjian_qau@163.com
Corresponding author`: Jian Zhang
DOI: https://doi.org/10.35633/inmateh-75-29

ABSTRACT

For the accurate detection of obstacles in complex farmland environments, ResNet50 is adopted as the backbone feature extraction network, feature pyramid network (FPN) is utilized to enhance the multi-scale feature fusion capability, and the region of interest alignment (ROI Align) strategy is introduced to improve the candidate box localization precision. The experimental results show that the precision, recall, and mean accuracy (mAP) of the improved model are 91.6%, 89.7%, and 93.8%, respectively, which are improved by 2.7, 2.3, and 3.1 percentage points compared with the original base network, and provide a technical reference for navigation and obstacle avoidance of unmanned agricultural machinery.

摘要

针对复杂农田环境中障碍物的准确检测，采用 ResNet50 作为骨干特征提取网络，利用特征金字塔网络（FPN）提升多尺度特征融合能力，并引入感兴趣区域对齐（ROI Align）策略提高候选框定位精度。实验结果显示，改进模型的精度、召回率和平均精度（mAP）分别为 91.6%、89.7%和 93.8%，相比于原基础网络，提升了 2.7、2.3 和 3.1 个百分点，为无人农业机械的导航避障提供了技术参考。

INTRODUCTION

Driven by the rapid development of agricultural machinery intelligence and automation, unmanned agricultural machinery has made significant progress. In order to ensure that these unmanned agricultural machines can operate safely and efficiently during operation and effectively avoid collision with obstacles such as utility poles, trees, buildings, etc., it is necessary to carry out accurate and fast identification of obstacles in the field. In farmland obstacle detection, deep learning detection algorithms have higher detection accuracy, stronger generalization ability and better adaptability than other detection algorithms, and can more accurately identify complex and changing obstacles in farmland, while maintaining stable detection performance under different light, climate and crop growth conditions.

Deep learning-based detection algorithms are divided into two categories: single-stage object detection algorithms (such as SSD and YOLO series) and two-stage object detection algorithms (such as Sparse R-CNN and Faster R-CNN). The latter, although more complex in process and relatively slower in speed, demonstrates higher accuracy. Research on obstacle detection in the field of machine vision has made significant progress.

He et al., (2022), improved the recognition accuracy by enhancing Mask R-CNN, employing Swin-Le Transformer for feature extraction and ME-PAPN for feature fusion. They integrated a multi-scale enhancement method to boost the detection capability of small targets, achieving a mean Average Precision (mAP) of 91.3% and an average detection time of 4.2 frames per second (FPS).

Rahman et al., (2022) implemented a transfer learning model based on the convolutional neural network MobileNetV2, which can be used on low-configured devices while maintaining a balance between detection speed and processing efficiency. The accuracy of obstacle detection reached 97.00%.

Xue et al., (2022) improved the Faster R-CNN object detection algorithm to identify obstacles in agricultural fields, effectively enhancing the speed of obstacle recognition while reducing false positives and missed detections. This improvement meets the real-time detection requirements for low-speed operations of tractors.

Du et al., (2024), proposed an optimized obstacle detection method based on an improved YOLOv8. By removing the P5 layer and introducing DCNv2 to optimize the bottleneck, the model enhances the detection capability for small and irregular obstacles. The improved model achieved a 3.4% increase in mAP50, a 34.5% reduction in GFLOPs, a 77.4% decrease in parameters, and a 73% reduction in model size.

Han et al., (2024) investigated an autonomous driving obstacle avoidance method based on YOLOv5 monocular vision. This approach combines a deep reinforcement learning path planning algorithm to dynamically generate safe driving paths. By introducing a monocular vision obstacle avoidance aggregation network, the MMA obstacle avoidance method is developed, achieving an accuracy that fluctuates between 78.76% and 88.26%.

Zhao et al., (2024), proposed a real-time high-precision railway obstacle detection model based on a lightweight CNN and an improved Transformer (RH-Net). This model includes a Lightweight Feature Extraction Module (LEM) to minimize computational load, an Improved Transformation Module (IFM) that enhances the capability of extracting global contextual information, and an Enhanced Multi-Scale Feature Fusion Module (EFM) that optimizes the detection of obstacles of varying sizes.

Yang et al., (*2025)* proposed the YOLO-Region model to solve the problem of oversensitive obstacle detection in unmanned electric locomotives in underground coal mines, the model backbone adopts the InceptionNeXt block and the NSPP module, extends the FPN+PAN architecture with the Impro-TSCODE header and introduces the repulsion loss to enhance the detection of occluded targets.

Researchers have made significant progress in the field of obstacle recognition, confirming the feasibility of target detection technology for obstacle detection. However, in today's era of unmanned farming, there is relatively little research on using drone equipment to capture images of agricultural fields for obstacle recognition. To address this gap, this paper selects the second-stage detection algorithm Faster R-CNN, which offers higher detection accuracy, for detecting obstacles in farmland. Given the complexity of agricultural environments, directly applying Faster R-CNN to unstructured agricultural scenes for obstacle detection may lead to decreased model accuracy, particularly under complex and variable weather conditions, where the effectiveness of obstacle detection can be significantly weakened. Therefore, this paper conducts targeted optimization based on the Faster R-CNN model, aiming to address the current issues of insufficient accuracy and poor robustness in deep learning models for obstacle detection in agricultural fields.

MATERIALS AND METHODS
Acquisition of images of farmland obstacles

This study adopts a combined ground and airborne acquisition of the dataset, i.e., combining ground-based cell phone acquisition with airborne UAV acquisition, as shown in Figure 1. Ground-based acquisition can obtain detailed and accurate information about obstacles, including the shape, size, and material of obstacles, which is suitable for complex terrain and dense obstacle areas, and can capture details that may be missed by aerial acquisition. Aerial acquisition can quickly cover a large area and improve acquisition efficiency, and it has unique advantages for high-altitude obstacles or areas that are difficult to reach on the ground. The combination of the two can result in a more complete and accurate obstacle dataset, which is useful for subsequent application and analysis.
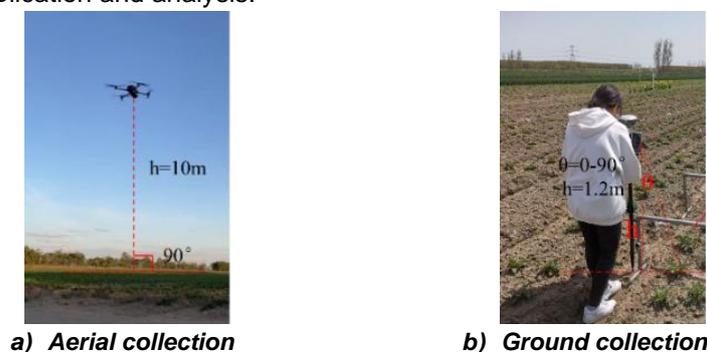

a) Aerial collection   *b) Ground collection*
Fig. 1 – Data collection method

The dataset was collected from March to June 2023 through aerial photography and mobile imaging in Hebei and Shandong provinces, under varying lighting conditions during the morning, afternoon, and evening. The drone used was a DJI Mavic 3, which captured vertical aerial shots of the plots at an altitude of 10 meters and a flight speed of 7 m/s.

The drone automatically collected images at predetermined intervals along a designated flight path. Both longitudinal and lateral overlap rates were set at 70%. Each flight generated approximately 200 images. Simultaneously, mobile phones and cameras were employed to capture multi-angle views of the field obstacles, as illustrated in Figure 2. The types of obstacles were diverse, specifically including buildings, high-voltage power towers, trees, telegraph poles, water wells, personnel, and agricultural machinery, categorized into seven main groups.
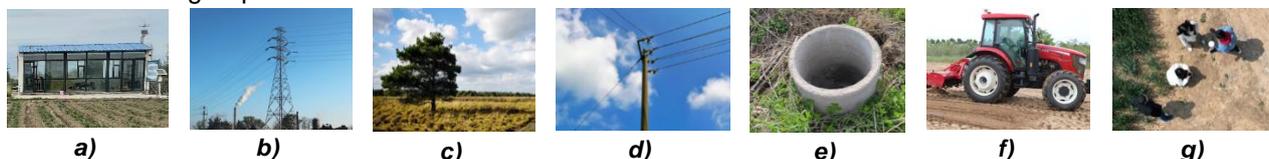

| a) | b) | c) | d) | e) | f) | g) |

Fig. 2 – Example of dataset image

a) building; b) high-voltage power towers; c) tree; d) telegraph pole; e) water well; f) agricultural machinery; g) person

Data expansion

To improve the robustness and generalization performance of the dataset, an image enhancement strategy is implemented. As shown in Fig. 3, data enhancement includes random brightness adjustment to simulate different lighting environments, Gaussian noise addition, horizontal flip, and vertical flip to simulate diverse shooting angles. These methods not only increase the number of images, but also effectively reduce the overfitting problem during model training, and finally an expanded dataset containing 7707 obstacle images is obtained.
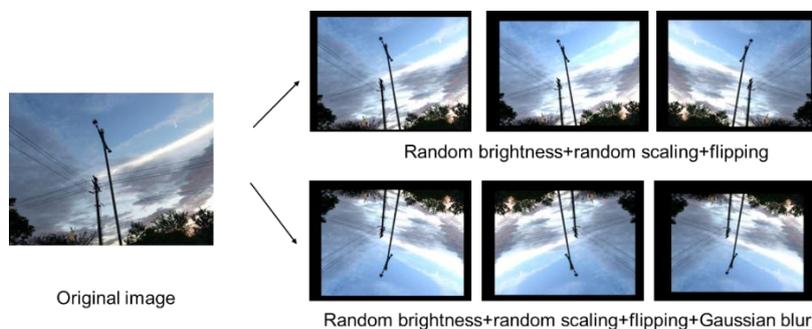

Fig. 3 – Data enrichment

Image annotation and dataset construction

Using the LabelImg image annotation tool, obstacles were annotated on the processed images. The dataset comprises 7,707 images with a total of 11,578 annotation tasks for obstacles. The training data was organized into the standard PASCAL VOC 2007 format. From the entire dataset, 6,823 images were randomly selected as the training set for model learning and parameter tuning, while 884 images were designated as the validation set to monitor the performance of the model in real-time during training, allowing for timely adjustments to the training strategy. A separate test set was established for the final evaluation of the model's recognition accuracy and generalization capability. This division ensures both the thorough utilization of the dataset and the objectivity and accuracy of the evaluation results.

Faster R-CNN model

Faster R-CNN, as a two-stage detection algorithm, has a more complex process and relatively slower running speed compared to one-stage algorithms such as the YOLO series and SSD (Single Shot MultiBox Detector), but it demonstrates a higher level of detection accuracy. Faster R-CNN consists of a feature extraction layer, a Region Proposal Network (RPN), and an RoI Pooling layer (Region of Interest Pooling). The detection process of Faster R-CNN is summarized as follows: First, the training images are resized to a uniform dimension and input into the network, where feature maps are generated via the feature extraction layer; second, the RPN network generates a series of anchor boxes on the feature map based on predefined Intersection over Union (IoU) thresholds; then, the anchor boxes produced by the RPN are combined with the feature map and sent to the RoI Pooling layer to obtain fixed-size (7x7) feature representations of the anchor boxes; finally, these feature representations are input into the classification and regression layers for bounding box regression predictions and object detection classification, resulting in accurate detection outcomes. The entire detection process is illustrated in Figure 4.
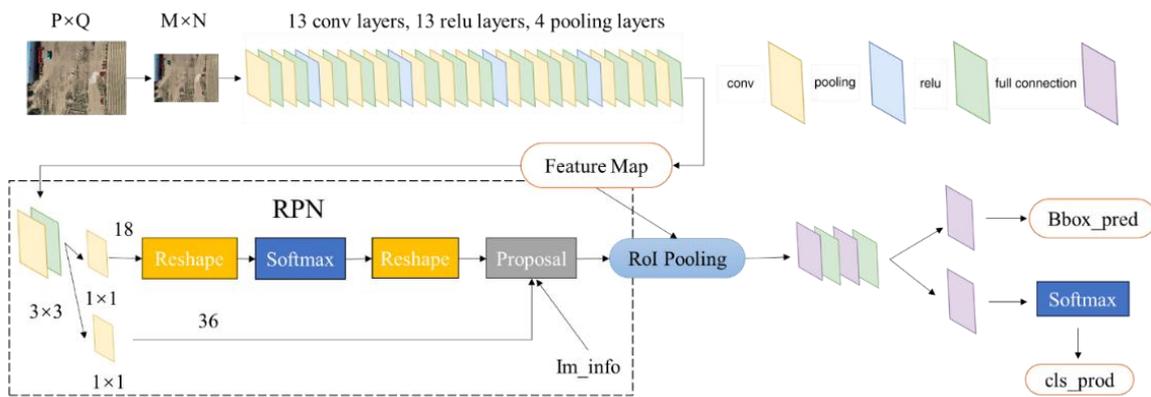
Fig. 4 – Faster R-CNN Detection Flowchart

Improved Faster R-CNN model

To enhance the obstacle detection capability of the Faster R-CNN model in complex agricultural environments, this study implemented the following key improvements: first, the original VGG16 feature extraction backbone used in Faster R-CNN was replaced with the more efficient ResNet50. Second, by introducing a Feature Pyramid Network (FPN), the fusion of high-level and low-level features was achieved, thereby enriching the information content of the feature maps. Finally, the ROI Align strategy was adopted to replace the original ROI Pooling layer, which improved the model's accuracy in candidate box localization. The architecture of the improved Faster R-CNN model is illustrated in Figure 5.

RetNet50 backbone network

To address the hardware resource limitations in agricultural environments, and to significantly enhance the feature extraction capability of the obstacle detection model while optimizing the deployment of the network model in practical production operations, this study adopted ResNet50 as a replacement for the original VGG16, serving as the new backbone network. The ResNet50 network effectively resolves the degradation problem in deep neural networks by introducing residual units. Its architecture, as shown in Figure 5, consists of five core components: conv1, conv2_x, conv3_x, conv4_x, and conv5_x. The conv1 component includes only one convolutional layer and one max pooling layer, which can be considered as the preprocessing stage of the network. The remaining four components (from conv2_x to conv5_x) are structurally similar, consisting of repeated stacks of residual structures such as ReB1 and ReB2.
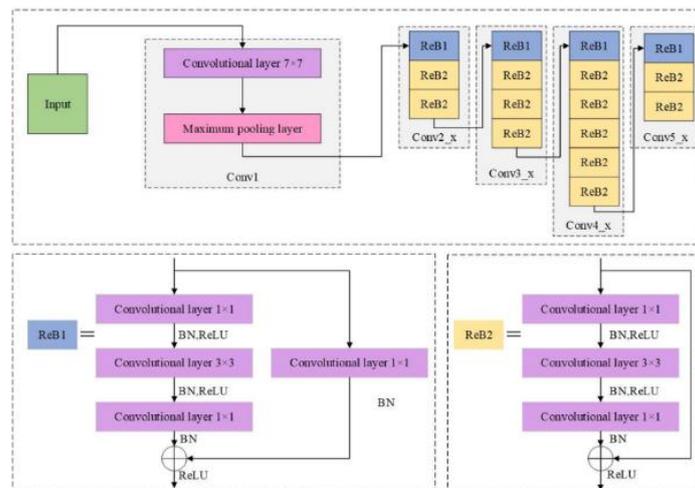

Fig. 5 – ResNet50 structure diagram

Feature pyramid network

Aiming at the problem of complex background and diverse scales of obstacle images jointly captured by UAVs and cell phones, this paper adopts the feature pyramid network FPN to improve the Faster R-CNN model. As shown in Figure 6, FPN fuses the rich semantic information of the high-level feature map with the rich spatial details of the shallow feature map to generate a richer and more accurate feature representation.
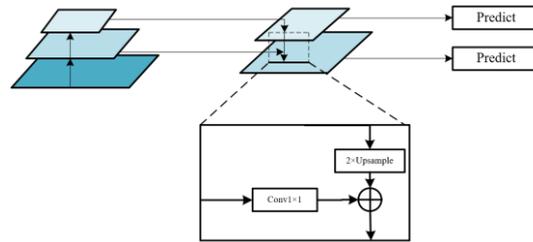
Fig. 6 – Feature pyramid network architecture

The structure combining the FPN network with the ResNet50 network is illustrated in Figure 7. First, the ResNet50 network performs bottom-up convolution operations to extract feature maps of varying scales and different channel numbers, denoted as {C2, C3, C4, C5}. Subsequently, these feature maps are fused through a top-down pathway. During the fusion process, the feature maps are first adjusted using 1×1 convolution operations, resulting in a new set of feature maps {M1, M2, M3, M4}. Then, adjacent feature maps M are fused using upsampling, followed by processing with 3×3 convolution operations, ultimately generating the feature map P. To control computational complexity, this paper only selects four feature maps of different scales for output during the feature extraction phase, namely {P2, P3, P4, P5}. Finally, all feature maps generated by the FPN are input into the Region Proposal Network (RPN), thus achieving effective fusion of high-level and low-level features, further enhancing the model's feature extraction capability.
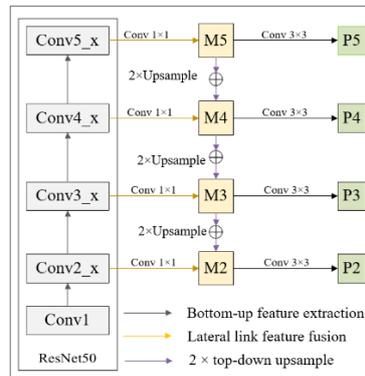

Fig. 7 – Feature pyramid network architecture

ROI Align

The operational process of ROI Align is illustrated in Figure 8. First, the predicted candidate regions are accurately mapped to the corresponding feature levels and thoroughly traversed. At this stage, each candidate region is meticulously divided into k × k uniform small grids, while ensuring the precision of boundary coordinates, effectively avoiding potential errors introduced during the quantization process. Subsequently, within each small grid, the specific values of four key sampling points are accurately calculated using bilinear interpolation. These values are then used as the basis for performing the maximum pooling operation, which determines the comprehensive feature representation of each grid. By integrating the ROI Align strategy, the model is not only able to flexibly adapt to uniform input size requirements when processing candidate regions but also achieve more precise and detailed localization, significantly enhancing the model's object recognition capability.
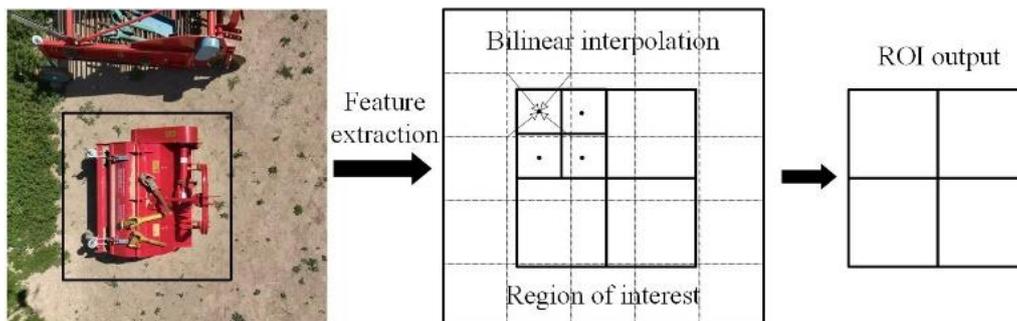

Fig. 8 – ROI Align Principle

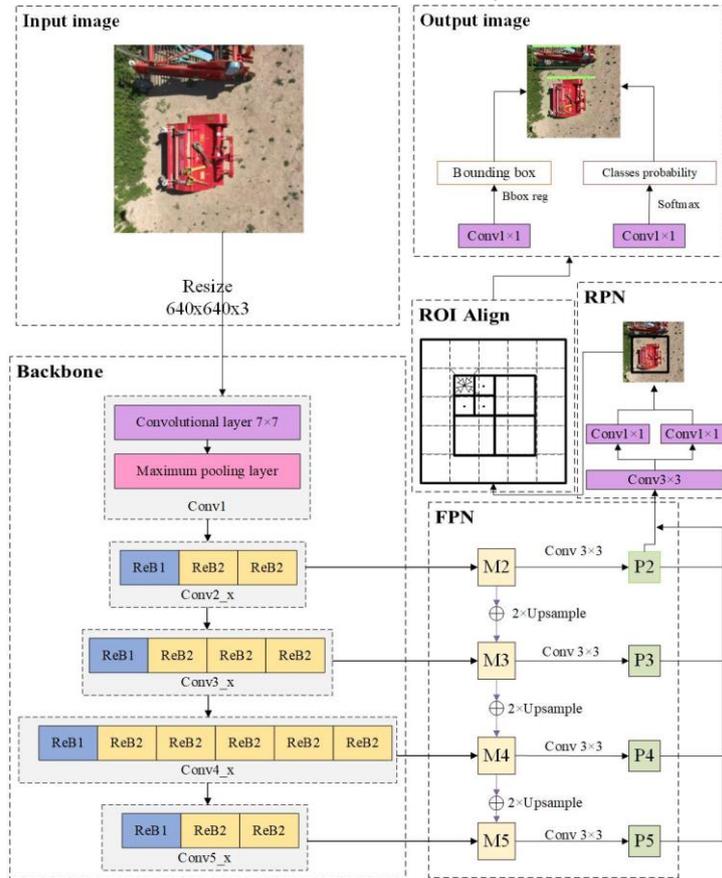The Faster R-CNN model architecture after the above improvements is shown in Fig. 9.

Fig. 9 – Improved Faster R-CNN model architecture

RESULTS AND ANALYSIS
Test environment
The experiments were conducted in the Anaconda environment under the Windows 10 operating system, utilizing the PyTorch framework for development and programming within PyCharm. The hardware setup included an AMD Ryzen 7 4800H processor, a Radeon Graphics 2.0GHz graphics card, an Nvidia GeForce RTX 2060 GPU, and 6GB of RAM.

Test evaluation indicators
In order to test the correctness of the algorithm, Precision (P), Recall (R), mean average precision (mAP) are used as model evaluation metrics. In terms of model complexity, the key factors considered are the parameter and FPS. Parameter indicates the computational memory resources consumed by the model, and FPS indicates the number of images that the model can process per second. The calculation formula is as follows.

$$P = \frac{T_p}{T_p + F_p} \tag{1}$$

$$R = \frac{T_p}{T_p + F_n} \tag{2}$$

$$AP = \int_1^0 P(R)dR \tag{3}$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \tag{4}$$

$$Parameters = r^2 \times a \times v + v \tag{5}$$

In the formula, TP denotes the number of samples where the positive class is predicted to be positive, FP denotes the number of samples where the negative class is predicted to be positive, FN denotes the number of samples where the positive class is predicted to be negative, a is the input size, r is the size of the convolution kernel, and v is the output size.

Comparative test analysis of different algorithms

To validate the effectiveness of the improved Faster R-CNN model used in this study for detecting obstacles in agricultural fields, comparative experiments were conducted under the same training environment and hyperparameter settings, contrasting the improved model with single-stage models in the deep learning domain (SSD, YOLOv8n) and two-stage models (Sparse R-CNN, Cascade R-CNN).

Table 1

Different model experiment results

Models	Precision/%	Recall/%	mAP@0.5%	Parameters/MB	FPS/(img/s)
Faster R-CNN	88.9	87.4	90.7	41.36	8.3
SSD	81.2	82.1	82.7	25.6	4.6
YOLOv8n	86.4	83.4	87.9	2.51	10.5
Sparse R-CNN	87.7	84.5	88.2	43.01	7.8
Cascade R-CNN	87.5	84.8	87.9	40.12	7.1
Our	91.6	89.7	93.8	41.45	8.7

According to the data in Table 1, Faster R-CNN demonstrates the highest accuracy, recall, and mean Average Precision (mAP) when compared to both single-stage and two-stage models, providing a solid foundation for further optimization of subsequent models. The improved Faster R-CNN model surpasses SSD, YOLOv8n, Sparse R-CNN, and Cascade R-CNN in mean Average Precision by 11.1, 5.9, 5.6, and 5.9 percentage points, respectively. In terms of accuracy, it also outperforms these models by 10.4, 5.2, 3.9, and 4.1 percentage points, respectively. Additionally, the improved Faster R-CNN exhibits excellent recall, exceeding these models by 7.6, 6.3, 5.2, and 4.9 percentage points. These data conclusively demonstrate that the improved Faster R-CNN model excels in extracting effective features and accurately predicting the coordinates and category information of obstacles.

Ablation test analysis

To evaluate the specific improvements in the performance of the basic Faster R-CNN model achieved by using ResNet50 as the backbone network, the Feature Pyramid Network, and the Region of Interest Align (ROI Align) strategy, ablation experiments were designed while keeping the dataset and experimental parameters consistent.

Table 2

Results of ablation experiment

Test	ResNet50	FPN	ROI Align	P/%	R/%	mAP@0.5%	Parameters/MB	FPS
1	×	×	×	88.9	87.4	90.7	41.36	8.1
2	√	×	×	89.5	88.1	91.9	41.14	8.3
3	√	√	×	90.7	88.5	93.2	41.14	8.7
4	√	√	√	91.6	89.7	93.8	41.25	8.7

As shown in Table 2, in Experiment 2, ResNet50 was used to replace the original VGG16 as the backbone network. Benefiting from the residual structure of ResNet50, the improved backbone network not only increased the number of convolutional layers but also achieved improvements of 0.6%, 0.7%, and 1.2% in accuracy, recall, and mean Average Precision (mAP), respectively, while reducing the number of parameters. In Experiment 3, the introduction of the Feature Pyramid Network (FPN) for multi-scale feature fusion allowed the feature maps to integrate high-level semantic information and low-level spatial information, thereby enhancing the model's capabilities in multi-scale and small object detection. This improvement resulted in increases of 1.2%, 0.4%, and 1.3% in accuracy, recall, and mean Average Precision, respectively. In Experiment 4, ROI Align was used to replace the original ROI Pooling. ROI Align improved the overall regression performance of the predicted bounding boxes, making them more precise when detecting obstacles. This enhancement led to increases of 0.9%, 1.2%, and 0.6% in accuracy, recall, and mean Average Precision, respectively. Overall, the results of the ablation experiments indicate that these improvements played a positive role in enhancing obstacle detection accuracy. Compared to the original Faster R-CNN model, the improved model not only reduced the number of model parameters and increased detection speed but also achieved increases of 2.7%, 2.3%, and 3.1% in accuracy, recall, and mean Average Precision, respectively.
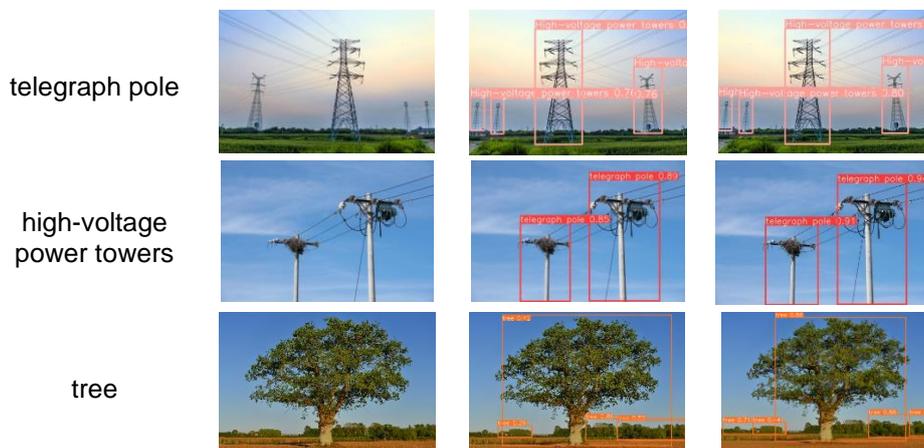
Test results of different obstacles

To clearly present the differences in performance between the original model and the improved model in detecting various categories of obstacles, the test results of both models were compared and the comparison results were displayed in Figure 10. Additionally, the mean Average Precision (mAP) of the obstacle detection results during the training process were statistically analyzed and they are listed in Table 3.

According to the data analysis in Table 3, the improved Faster R-CNN model, compared to its original version, shows only a slight improvement in agricultural machinery detection tasks, while achieving a significant increase in mean Average Precision (mAP) for other obstacle detection tasks. This improvement is reflected not only in the overall optimization of detection performance but also in the notable enhancement of specific obstacle detection. Taking utility pole detection as an example, the mAP@0.5 of the original model was only 82.1%, whereas after adopting the improved YOLOv8 model, the mAP@0.5 for utility pole detection increased to 89.8%. Although there remains a certain gap in detection accuracy for utility poles compared to other obstacles such as vehicles and pedestrians, the improved model has made significant progress compared to the original model. As shown in Figure 10, the improved Faster R-CNN model not only reduces the instances of missed detections and false positives compared to the original model, but it also enhances the confidence level of obstacle detection.

Table 3

Training results for different types of obstacles

Types of obstacles	mAP@0.5/%	
	pre-optimization	post-optimization
telegraph pole	82.1	89.8
high-voltage power towers	92.7	94.5
tree	91.3	93.8
building	92.5	94.6
person	89.4	93.7
agricultural machinery	95.1	95.2
water well	92.0	94.8
all	90.7	93.8

According to the data analysis in Table 3, the improved Faster R-CNN model, compared to its original version, shows only a slight improvement in agricultural machinery detection tasks, while achieving a significant increase in mean Average Precision (mAP) for other obstacle detection tasks. This improvement is reflected not only in the overall optimization of detection performance but also in the notable enhancement of specific obstacle detection. Taking utility pole detection as an example, the mAP@0.5 of the original model was only 82.1%, whereas after adopting the improved YOLOv8 model, the mAP@0.5 for utility pole detection increased to 89.8%. Although there remains a certain gap in detection accuracy for utility poles compared to other obstacles such as vehicles and pedestrians, the improved model has made significant progress compared to the original model. As shown in Figure 10, the improved Faster R-CNN model not only reduces the instances of missed detections and false positives compared to the original model, but it also enhances the confidence level of obstacle detection.
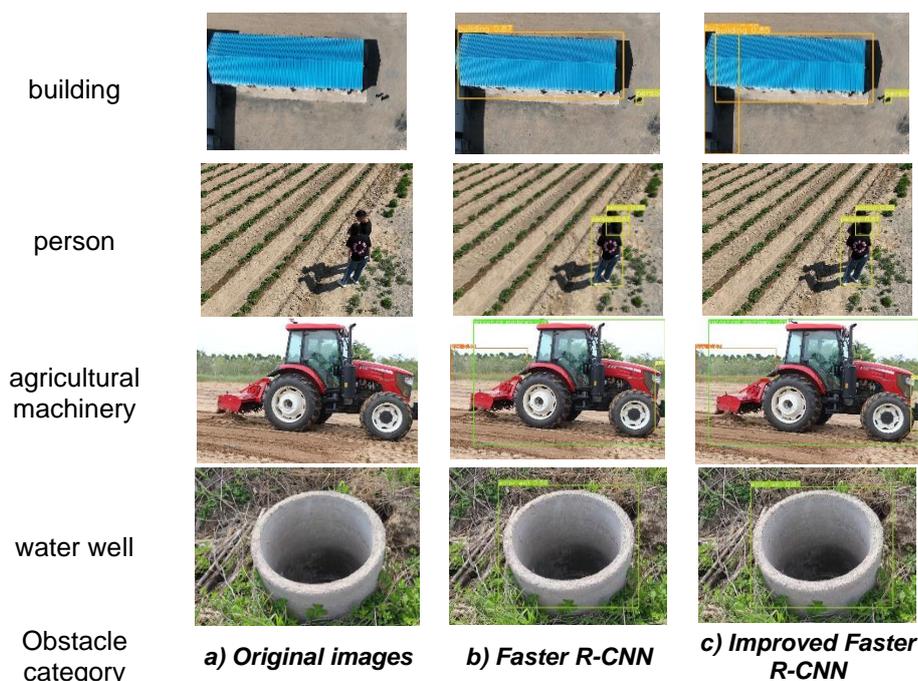

telegraph pole

high-voltage power towers

tree

Fig. 10 –Comparison of Faster R-CNN and Improved Faster R-CNN Detection Results

As shown in Figure 10, the improved Faster R-CNN model not only reduces the instances of missed detections and false positives compared to the original model, but it also enhances the confidence level of obstacle detection.

CONCLUSIONS

This study developed a farmland obstacle detection model based on an improved Faster R-CNN model. Validation using the same obstacle dataset showed that the improved model achieved a recognition accuracy of 91.6% and a recall rate of 93.8%. In practical applications, the model's recognition rate for static obstacles such as utility poles, high-voltage towers, trees, buildings, agricultural machinery, and wells exceeded 90%. Simultaneously, the model's recognition rate for dynamic obstacles, such as people walking at a speed of 0.8 m/s, also remained at 90%, with an average detection frame time of 108 ms, meeting the requirements for real-time detection.

ACKNOWLEDGEMENT

REFERENCES

[1]   Adiuku, N., Avdelidis, N. P., Tang, G., & Plastropoulos, A. (2024). Improved Hybrid Model for Obstacle Detection and Avoidance in Robot Operating System Framework (Rapidly Exploring Random Tree and Dynamic Windows Approach). *Sensors*, 24(7), 2262.

[2]   Ahuja, U., Singh, S., Kumar, M., Kumar, K., & Sachdeva, M. (2023). COVID-19: Social distancing monitoring using faster-RCNN and YOLOv3 algorithms. *Multimedia Tools and Applications*, 82(5), 7553-7566.

[3]   Balasundaram, A., Shaik, A., Prasad, A., & Pratheepan, Y. (2024). On-road obstacle detection in real time environment using an ensemble deep learning model. *Signal, Image and Video Processing*, 1-14.

[4]   Chen, J., Li, D., Qu, W., & Wang, Z. (2024). A MSA-YOLO Obstacle Detection Algorithm for Rail Transit in Foggy Weather. Applied Sciences, 14(16), 7322.

[5]   Cui, J., Zhang, X., Zhang, J., Han, Y., Ai, H., Dong, C., & Liu, H. (2024). Weed identification in soybean seedling stage based on UAV images and Faster R-CNN. *Computers and Electronics in Agriculture*, 227, 109533.

[6]   Dai, X., Mao, Y., Huang, T., Qin, N., Huang, D., & Li, Y. (2020). Automatic obstacle avoidance of quadrotor UAV via CNN-based learning. *Neurocomputing*, 402, 346-358.

[7]   Dev, U., Singh, T., Babu, T., Mandal, A. K., Sharma, M., & Mandal, A. (2025). Enhancing Blood Platelet Counting through Deep Learning Models for Advanced Diagnostics. *SN Computer Science*, 6(1), 1-10.

[8]   Du, Z., Feng, X., Li, F., **an, Q., & Jia, Z. (2024). A Lightweight UAV Visual Obstacle Avoidance Algorithm Based on Improved YOLOv8. *Computers, Materials & Continua*, 81(2).

[9]   EL MALLAHI, I., RIFFI, J., Tairi, H., & Mahraz, M. A. (2024). Efficient Vehicle Detection and Classification Algorithm Using Faster R-CNN Models. *Journal of Automation, Mobile Robotics and Intelligent Systems*, 86-93.

[10]   Han, L., Jiang, Y., Qi, Y., & Altangerel, K. (2024). Monocular visual obstacle avoidance method for autonomous vehicles based on YOLOv5 in multi lane scenes. *Alexandria Engineering Journal*, 109, 497-507.

[11]   He, D., Qiu, Y., Miao, J., Zou, Z., Li, K., Ren, C., & Shen, G. (2022). Improved Mask R-CNN for obstacle detection of rail transit. *Measurement*, 190, 110728.

[12]   Jafri, R., Campos, R. L., Ali, S. A., & Arabnia, H. R. (2017). Visual and infrared sensor data-based obstacle detection for the visually impaired using the Google project tango tablet development kit and the unity engine. *IEEE Access*, 6, 443-454.

[13]   Li, C. J., Qu, Z., Wang, S. Y., & Liu, L. (2021). A method of cross-layer fusion multi-object detection and recognition based on improved faster R-CNN model in complex traffic environment. *Pattern Recognition Letters*, 145, 127-134.

[14]   Nie, J., Zhang, G., Lu, X., Wang, H., Sheng, C., & Sun, L. (2024). Obstacle avoidance method based on reinforcement learning dual-layer decision model for AGV with visual perception. *Control Engineering Practice*, 153, 106121.

[15]   Park, J., Je, Y., Lee, H., & Moon, W. (2010). Design of an ultrasonic sensor for measuring distance and detecting obstacles. *Ultrasonics*, 50(3), 340-346.

[16]   Rahman, F. U., Ahmed, M. T., Hasan, M. M., & Jahan, N. (2022). Real-time obstacle detection over railway track using deep neural networks. *Procedia Computer Science*, 215, 289-298.

[17]   Shahi, T. B., Xu, C. Y., Neupane, A., & Guo, W. (2022). Machine learning methods for precision agriculture with UAV imagery: a review. *Electronic Research Archive*, 30(12), 4277-4317.

[18]   Starostin, I. A., Eshchin, A. V., Godzhaev, T. Z., & Davydova, S. A. (2024). Conceptual directions for the development of driverless agricultural mobile power units. Tractors and Agricultural Machinery, 91(1), 23-38.

[19]   Xue Jinlin, Li Yuqing, & Cao Xinjian. (2022). Obstacle Detection Based on Deep Learning for Blurred Farmland Images (基于深度学习的模糊农田图像中障碍物检测技术). *Transactions of the Chinese Society for Agricultural Machinery*, (03),234-242.

[20]   Yang, T., Guo, Y., Li, D., & Wang, S. (2025). Vision-Based obstacle detection in dangerous region of coal mine driverless rail electric locomotives. *Measurement*, 239, 115514.

[21]   Zhang, Y., Tian, K., Huang, J., Wang, Z., Zhang, B., Zhang, B., Xie, Q. (2024). Field Obstacle Detection and Location Method Based on Binocular Vision. *Agriculture*, 14(9), 1493.

[22]   Zhao, Z., Kang, J., Sun, Z., Ye, T., & Wu, B. (2024). A real-time and high-accuracy railway obstacle detection method using lightweight CNN and improved transformer. *Measurement*, *238*, 115380.