RESEARCH ON LOCUST TARGET DETECTION ALGORITHM BASED ON YOLO V7 -MOBILENETV3-CA

/

基于 YOLOv7-MobileNetV3-CA 的蝗虫目标检测算法研究

Dongfang QIU [1)]
College of Information Science and Engineering, Shanxi Agricultural University, Jinzhong, Shanxi, China
Tel: +86 177034433; E-mail: qdf_0106@126.com
DOI: https://doi.org/10.35633/inmateh-74-24

Keywords: *locust detection, YOLO v7, MobileNetV3, CA attention mechanism*

ABSTRACT

To accurately detect various kinds of locusts in real-time and make locust detection more universal, a locust data set that contained different species of locusts was created through the Internet crawler and public dataset IP102, and a locust target detection algorithm YOLOv7-MobileNetV3-CA was proposed in this paper. Firstly, to reduce the size of model parameters, the backbone of YOLOv7 was replaced by MobileNetV3, secondly, a CA (Coordinate Attention) attention mechanism was added to further improve the detection accuracy of locusts after feature enhancement. The experiment showed that the precision of locusts identification was 95.96%, the recall rate was 92%, the AP was 95.74%, and the F1 was 0.92. Compared with YOLOv7, the model size was reduced by 27%, and the AP was improved by 4.48%. Compared with YOLOv4, YOLOv4 MobileNetV3, YOLOv5, and SSD algorithms, AP has improved by 51.16%, 26.81%, 11.9%, and 11.75%, respectively. Experiments have shown that this algorithm performs well in detecting locusts of different scales, scenes, and types, and can provide reference for real-time locust detection.

摘要

为了能实时准确地检测各类蝗虫目标，使得蝗虫检测更具有普适性，本文通过互联网爬虫及公有数据集IP102形成蝗虫数据集，提出了YOLOv7-MobileNetV3-CA的蝗虫目标检测算法。首先，为了降低模型参数量，使用MobileNetV3替换YOLOv7骨干网。其次，在特征加强后加入了CA（Coordinate Attention）注意力机制，以进一步提高蝗虫的检测精度。实验表明，蝗虫的检测精确率为95.96%，召回率92%，mAP为95.74%，F1为0.92，与YOLOv7相比，模型大小降低27%，mAP提高了4.48%。与YOLOv4、YOLOv4-MobileNetV3、YOLOv5、SSD算法相比，mAP分别提高了51.16%、26.81%、11.9%、11.75%。试验表明本算法对不同尺度、不同场景及不同种类的蝗虫检测效果较好，可以为蝗虫实时检测提供参考。

INTRODUCTION

There are various types of locusts in the world, such as desert locusts, rice locusts, grass locusts, flying locusts, and so on. Locust plague caused by large-scale locusts can cause destructive damage to agricultural, forestry. Locust plague can even cause animal husbandry production, further cause serious economic losses and famine due to food shortages (*Kang et al, 2019*). Locust plague has been the focus of agricultural pest control all over the world. Therefore, there is an urgent need to establish a locust detection system with higher accuracy and that can detect more types of locusts (*Yu et al, 2021*).

At present, the detection methods for locusts include artificial ground investigation, climate prediction, phenological prediction, radar detection technology, GPS/GIS detection, etc. However, due to the small size of locusts, methods that rely on manual detection methods cost highly and perform weakly in real time. Detection relying on remote sensing satellites can only be achieved by making models from historical data, resulting in low real-time accuracy. With the development of information technology, establishing locust disaster detection by image processing and pattern recognition provides a new method that can improve the efficiency of locust control.

Early locust recognition was mainly based on image processing methods. Someone used the frame difference method to determine the motion area in the image, and then counted locusts attempting to further extract locust information using chromaticity and morphological features (*Mao et al, 2008*). Some researchers used fuzzy patterns to identify locusts (*Zheng et al, 2010*).

[1] *Qiu Dongfang, As. Lect. M.S. Msc*

However, the traditional image processing greatly relies on the manually designed features, the accuracy of locust detection based on traditional image process is low due to the small size, variety and complex growth environment of the locust. The development of deep learning and computer vision provides a new way for automatic detection, and convolutional neural networks can better express image features compared with traditional *image processing.* With high detection accuracy, the computer vision based on the deep learning is widely used in agriculture, including intelligent harvesting (*Chen et al, 2024; Yu et al, 2024; Zhi et al, 2023; Huang et al, 2024; Wang et al, 2023; Wang et al, 2023; Matache et al, 2022),* crop yield estimation *(Ma et al, 2024; Wang et al, 2024; Xu et al, 2022; He et al 2021),* weed recognition *(Zhao et al, 2021; Cao et al., 2023; Fan et al, 2021),* animal facial recognition *(Ning et al,2023; Wang et al,2023),* pest and disease detection *(Zhang et al, 2023; Liu et al, 2023; Zhou et al, 2022; Mu et al, 2022; Ma et al, 2023*) and other agricultural fields. In recent years, locust detection has been successfully addressed using various deep learning-based object detection models; for example,

Ma et al., (2022), implemented locust detection in grasslands utilizing the YOLOv5 algorithm, while *Li et al., (2021),* achieved video detection of flying locusts using the SSD object detection algorithm. *Bai et al., (2022),* applied a combination of MOG2 and YOLOv4 for the detection and recognition of flying locusts. Additionally,

Kumar et al., (2021), employed deep learning techniques for the early detection of locust populations. However, these methods could all detect a single species or a certain growth stage of locusts.

Due to the limitations in diagnosing various locust species, the current locust detection systems cannot be widely promoted and applied. The wide variety, small size, complex and various growth environments, and the scale of locust images, make it difficult to collect them. To improve the locust detection accuracy, a locust data set that contained various kinds of locusts was created by internet crawler and public dataset IP102 in this paper. Furthermore, the data set was expanded by mosaic method. A YOLOv7-MobileNetV3-CA algorithm was proposed to identify locusts in the dataset images. To reduce the size of the detection model, it was replaced the backbone of YOLOv7 with MobileNetV3. Then, to further improve the detection accuracy, a coordination attention (CA) attention mechanism was added after feature enhancement. The proposed algorithm can further provide a new way to detect locusts.

MATERIALS AND METHODS

Locust Data Set and Its Pre-processing

The image data set created in this paper is part from the public data set IP102 (*Wu et al, 2019*) (crop pest data set) and part from the Internet. After filtering and choosing, 544 valid images were finally included in the data set, including 96 images from IP102 and 448 from the Internet. Afterwards, 84 images were expanded in the data set with the Mosaic method. The data set included grass locusts, rice locusts, desert locusts, and other locust species. Fig.1 shows partial images in the dataset.

a. Grass locust *b. Rice locust* *c. Desert locust*

Fig. 1 - The examples of the locust data

Images from the data set were manually labelled. The label format was XML. When creating the image label in the dataset, the following principles were followed: (1) Annotating the entire locust target; (2) Giving annotation with occlusion but clear visible locust (3) Not labelling locust targets with unclear targets. One example regarding label annotation is presented in Fig.2.

Fig. 2 - Example of the labelled image

YOLOv7-MobileNetV3-CA model

The existing convolutional neural networks based on deep learning have high detection accuracy and high model complexity, but they cannot meet the needs of real-time detection. To further improve detection efficiency while ensuring detection accuracy, a YOLOv7-MobileNetV3-CA model was proposed. Firstly, the backbone of YOLOv7 was replaced with a lightweight MobileNetV3 network, and the coordination attention was added after feature enhancement. The improved YOLOv7 MobileNetV3-CA model structure is shown in Fig.3. The improved model not only behaves better; but also has smaller model size than the original model YOLOv7.

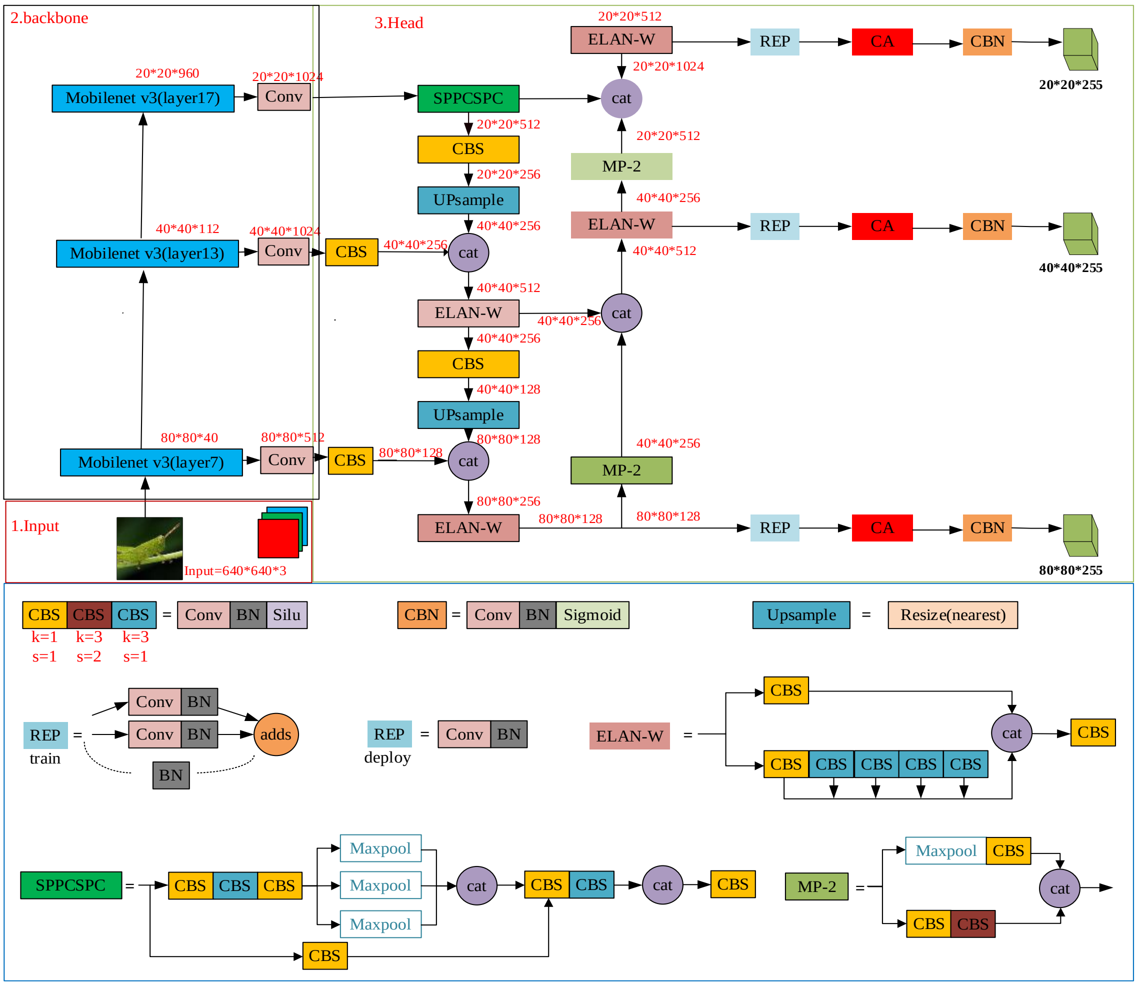


Fig. 3 - YOLOv7-MobileNetV3-CA model structure

MobileNetV3 Model

MobileNet is a kind of lightweight network model that includes MobileNetV1 (*Howard et al, 2017*), MobileNetV2, and MobileNetV3. MobileNetV3 is widely used in image recognition (*Mao et al, 2023; Li et al, 2023*), and its overall architecture follows the design of MobileNetV2, adopting lightweight structures such as depth wise separable convolution and residual blocks, and optimizing and upgrading modules, including bottleneck structure, SE module, and NL module. It has performed well in tasks such as image classification, object detection, and semantic segmentation on mobile devices. The network structure is shown in Table 1.

Table1

The structure of the MobileNetV3

Input	Operator	exp size	#out	SE	NL	s
2242×3	Conv2d	-	16	-	HS	2
1122×16	bneck,3*3	16	16	-	RE	1
1122×16	bneck,3*3	64	24	-	RE	2
562×24	bneck,3*3	72	24	-	RE	1
562×24	bneck,5*5	72	40	√	RE	2
282×40	bneck,5*5	120	40	√	RE	1
282×40	bneck,5*5	120	40	√	RE	1
282×40	bneck,3*3	240	80	-	HS	2
142×80	bneck,3*3	200	80	-	HS	1
142×80	bneck,3*3	184	80	-	HS	1
142×80	bneck,3*3	184	80	-	HS	1
142×80	bneck,3*3	480	112	√	HS	1
142×112	bneck,3*3	672	112	√	HS	1
142×112	bneck,5*5	672	160	√	HS	2
72×160	bneck,5*5	960	160	√	HS	1
72×160	bneck,5*5	960	160	√	HS	1
72×160	conv2d,1*1	-	960	-	HS	1
72×960	pool,7*7	-	-	-	-	1
12×960	conv2d,1*1,NBN	-	1280	-	HS	1
12×1280	conv2d,1*1,NBN	-	k	-	-	1

Note: Operator represents the block structure that the feature layer will operate on, and exp size represents the number of channels after the inverse residual structure rises within the bneck; # Out represents the number of channels in the feature layer when inputting bneck, and SE represents whether to use SE attention mechanism; NL represents which activation function to use, Hs represents h-swish, and RE represents Rule.

Coordination Attention

The Coordination Attention (CA) is a lightweight network attention method (Hou, et al) that can capture channel and location information, helping to more accurately locate and identify the target of detection. It mainly consists of two steps: coordinate information embedding and coordinate attention generation. To obtain attention to image width and height and encode precise positional information, CA first divides the input feature map into two directions: width and height, and performs global average pooling to obtain feature maps in both directions, as shown in formulas (1) and (2):

$$z_c^h = \frac{1}{W}\sum_{0\le i\le W} x_c(h,i) \tag{1}$$

$$z_c^w = \frac{1}{H}\sum_{0\le i\le H} x_c(j,w) \tag{2}$$

z_c^w, z_c^h represent the width and height of the feature maps, respectively, W and H represent the width and height of the image, and $x_c(i,j)$ represent the pixel of (i,j). Next, the feature maps in both the width and height of the global are concatenated together. Then, they are fed into a shared convolution module with a kernel, reducing their dimensionality to the original C/r. Then, the batch normalized feature map F_1 is used to obtain a feature map f with a size of 1 x (W+H) x C/r by the sigmoid activation function, as shown in formula (3).

$$f = \delta\left(F_1\left(\left[z^h, z^w\right]\right)\right) \tag{3}$$

Subsequently, to obtain feature maps F_h *and* F_w with the same number of channels as before, the feature map f is convolved with a 1x1 kernel according to its original height and width. Then, the attention weights g_h and g_w in the width direction of the feature map are obtained by the sigmoid function, which is shown in formulas (4) and (5).

$$g^h = \sigma\left(F_h(f^h)\right) \tag{4}$$

$$g^w = \sigma\left(F_w(f^w)\right) \tag{5}$$

After the above calculation, the attention weights of the input feature map in the height and width direction will be obtained. Finally, the final feature map with attention weights in the width and height will be obtained by multiplying and weighting the original feature map, which is shown in the formula (6). The total structure is shown in Fig.4.

$$y_c(i,j) = x_c(i,j) \times g_c^h(i) \times g_c^w(j) \tag{6}$$

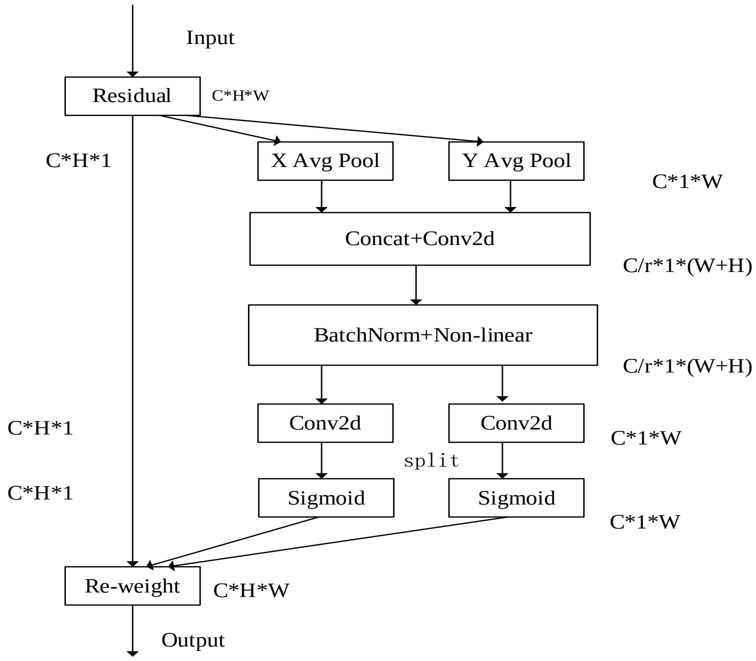


Fig. 4 - The structure of the CA

Model Evaluation Metrics

The locusts' identification required consideration of detection accuracy and the real-time detection ability of the model. Precision, Recall, $F1$ score, and Average Precision (AP) were chosen as evaluation metrics to evaluate the detection accuracy. The calculations are as follows.

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$\mathrm{Re}call = \frac{TP}{TP+FN} \tag{8}$$

$$F1 = \frac{2TP}{2TP+FP+FN} \tag{9}$$

where:

TP (True Positive): represents the number of correct predictions as positive samples.

FP (False Positive): represents the number of incorrect predictions as positive samples.

TF (False Positive): represents the number of correct predictions as negative samples.

FN (False Negative): represents the number of incorrect predictions as negative samples.

AP (Average Precision): It is calculated by computing the precision-recall curve and the area under the curve (AUC). AP represents the average precision across all levels. Besides, model Size represents the model complexity.

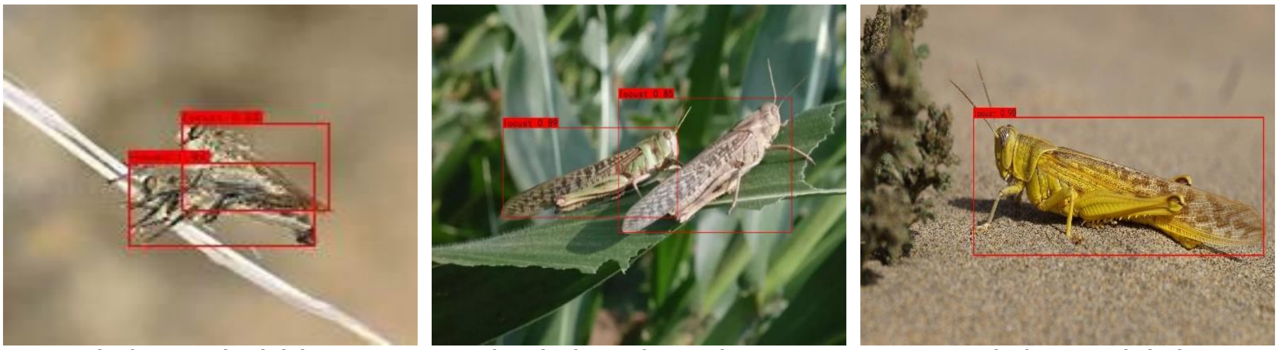
RESULTS

The model was trained on a server with a GPU model of NVIDIA GeForce RTX-3090, 4 cores CPU, 35GB memory, Python 3.6.9 software, and a deep learning framework using Python 1.12.

The Result of the YOLOv7- MobileNetV3-CA

To verify the performance of YOLOv7-MobilenetV3-CA, 54 locust images in the test set were tested and evaluated. The model achieved Precision of 95.96%, Recall of 92%, AP of 95.74%, and F1 of 0.92, model size being 27.37 MB.

Some detection examples are shown in Fig.5, showing that YOLOv7-MobilenetV3-CA performs well in locust detection and recognition under complex conditions such as various of types, scenes, number of targets, and scales.



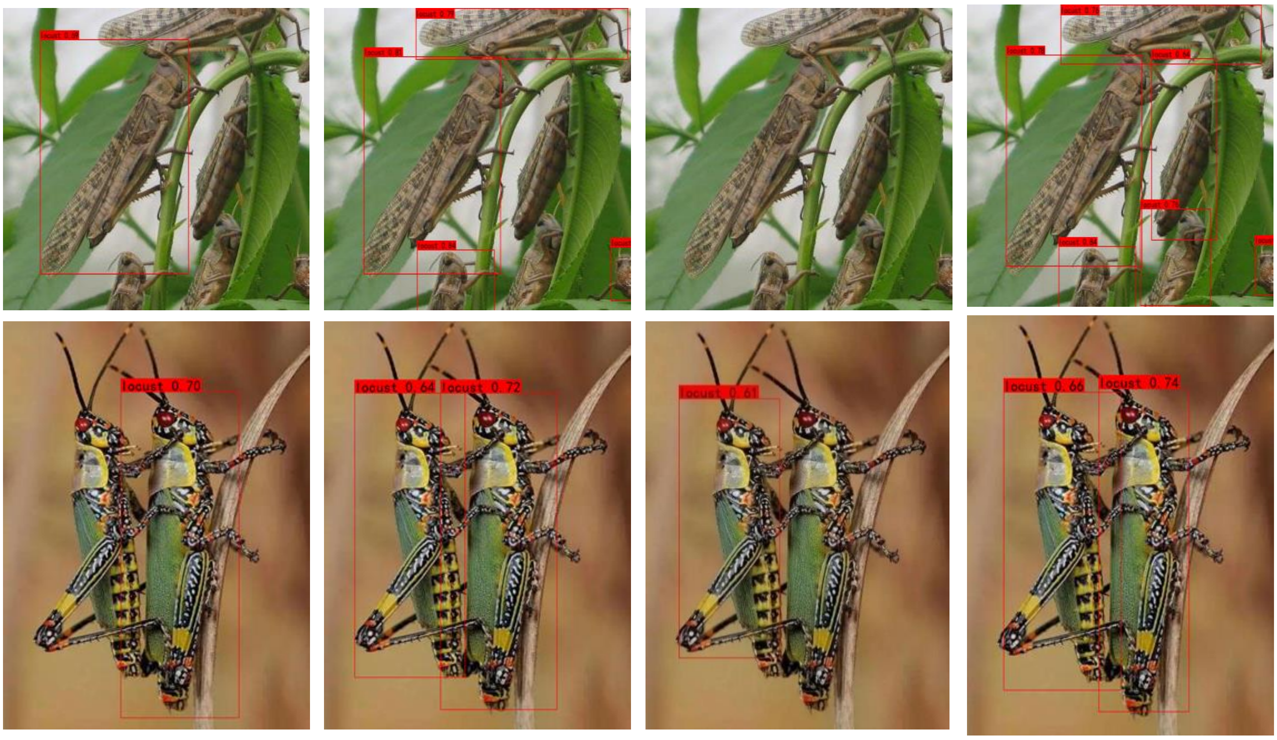
a. Multi-objective and occluded scene b. Multi-objective locust on leaves c. Single-objective on the land

d. Multi-objective on the rock e. Single-objective on the flower f. Multi-objective and multiscale scene

Fig. 5 - The detection result of the YOLOv7-MobilenetV3-CA

Ablation Experiment Results

To further verify the effectiveness of the model, ablation experiments were conducted to demonstrate the effectiveness of the model, comparing detection results and evaluation indicators were used to demonstrate the effectiveness. The detection results of YOLOv7, YOLOv7-CA, and YOLOv7-MobilenetV3 are shown in Fig.6, and it was found that the final model YOLOv7- MobileNetV3-CA performed best on occlusion images with various dimension scales of the locusts.



a. YOLOv7 b. YOLOv7-CA c. YOLOv7-MobileNetV3 c. YOLOv7-MobileNetV3-CA

Fig. 6 - The detection results of the ablation experiment

The evaluation indicators of the model YOLOv7, YOLOv7-CA, YOLOv7-MobileNetV3, and YOLOv7 MobileNetV3-CA are shown in Table 2. It was found that adding the CA attention mechanism to YOLOv7, made the recall rate, precision and F1 to achieve the following values: 0.99%, 0.91%,0.01, respectively, AP decreased by 5.98%, because the addition of the attention mechanism resulted in overfitting of the model. After replacing the backbone network of YOLOv7 with MobileNetV3, the detection performance of the model significantly decreased. The reason for this is that the model underwent depth wise separable convolution, which not only greatly reduced the number of parameters but also significantly reduced the detection accuracy of the model. Finally, by adding the CA attention mechanism, the detection accuracy of the model was greatly improved. YOLOv7-MobileNetV3-CA has increased AP by 4.4%, recall by 18.61%, detection accuracy by 1.48%, F1 by 0.12, while the model size decreased by 27% compared to YOLOv7.

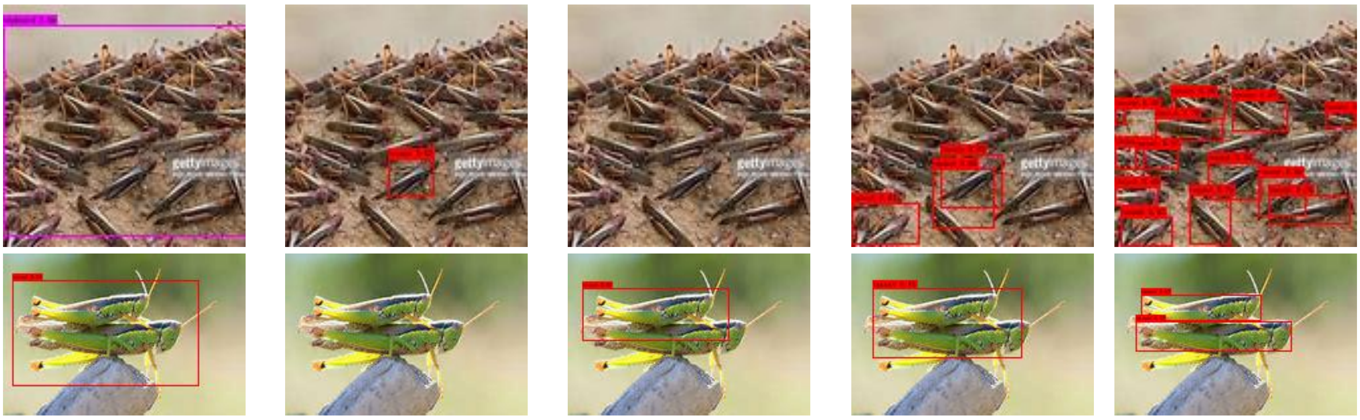
Table 2

The evaluation of ablation experiment

model	AP	Recall	Precision	F1	Model size/MB
YOLOv7	91.26%	69.35%	93.48%	0.80	37.62
YOLOv7-CA	86.24%	70.34%	94.39%	0.81	37.88
YOLOv7-MobileNetV3	65.37%	48.65%	83.08%	0.61	27.31
YOLOv7-MobileNetV3-CA	95.74	87.96%	95.96%	0.92	27.36

Comparative Experimental Results

To further validate the performance of the model YOLOv7-MobileNetV3-CA, it was compared with YOLOv4, YOLOv4-MobileNetV3, YOLOv5, and SSD models. The detection results are shown in Fig.7. By comparison, it was found that when multiple targets and occlusion exist in the image, YOLO v4 can detect the full target area, not the single target, while YOLOv5 and YOLOv4-MobilenetV3 did not perform well. SSD could detect some locusts, and YOLOv7-MobileNetV3-CA presented in this paper had the best detection performance, detecting more locusts compared with other models.



a. YOLO v4 b. YOLOv4-MobileNetV3 c. YOLO v5 d. SSD e. YOLOv7-Mobilenetv3-CA

Fig.7- The comparing detection result

To objectively evaluate the performance of the model, YOLOv7-MobileNetV3-CA was compared with YOLOv4, YOLOv4 MobileNetV3, YOLOv5, SSD, and YOLOv7, using evaluation indicators including precision, F1, Recall, Average Precision (AP), and the model size. The results are shown in Table 3. Overall, the YOLOv7-MobileNetV3-CA behaved best on the detection ability compared with other models. Its model size was smaller compared with YOLOv4, YOLOv5, SSD, and YOLOv7.

Table 3

The evaluation of the models

Model	Precision	F1	Recall	AP	Model size/MB
YOLO v4	90.48%	0.68	55.07%	73.00%	64.36
YOLO v4-MobileNetV3	80.00%	0.36	23.53%	70.01%	12.69
YOLO v5	88.57%	0.67	53.45%	83.83%	47.05
SSD	88.68%	0.82	75.81%	83.99%	26.28
YOLO v7	93.48%	0.80	69.35%	91.26%	37.62
YOLOv7-MobileNetv3-CA	95.96%	0.92	87.96%	95.74%	27.36

CONCLUSIONS

In this paper, a locust images data set was created from the public dataset IP102 and Internet crawlers and then images were manually labelled using LablelImage software. Besides, an improved locust detection algorithm YOLOv7-MobileNetV3-CA was proposed, which replaces the backbone of YOLOv7 with MobileNetV3 to catch model features and to reduce the parameters of the model. After the features were enhanced, a CA attention mechanism was added to improve the detection accuracy of the model. Finally, by comparing with other target detection algorithms, the following conclusions were drawn: (1) Compared with YOLOv7 model, the AP of the proposed model is increased by 4.48%; (2) Compared with YOLOv7, the model size has been reduced by 27%, achieving a lightweight model. (3) Compared with the classic object detection algorithms YOLO v4, YOLO v4 MobilenetV3, YOLO v5, and SSD, the average detection accuracy has been improved by 22.74%, 25.73%, 11.9%, and 11.75%, respectively, and the detection behaves best compared with other models. (4) This model shows good detection ability for different kinds of locusts in various scenes. Therefore, the proposed algorithm can provide a new way for locust detection.

ACKNOWLEDGEMENT

This paper was supported by the Science and Technology Innovation Fund of Shanxi Agricultural University (Grant number: 2019022)

REFERENCES

[1] Bai, Z., Tang, Z., Diao, L., Lu, S., Guo, X., Zhou, H., Liu, C., Li, L. (2022). Video target detection of East Asian migratory locust based on the MOG2-YOLOv4 network. *International Journal of Tropical Insect Science.* Vol.42, pp.793–806. USA.

[2] Cao, Y., Zhao, Y., Yang, L. (2023). Weed Identification Method in Rice Field Based on Improved DeepLabv3 (基于改进 DeepLabv3+的水稻田间杂草识别方法). *Transactions of the Chinese Society for Agricultural Machinery*. Vol.54, pp.242-252, Beijing/China.

[3] Cheng, F., Cheng, C., Zhu, X., Shen, D.&Zhang, X. (2024) Detection of camellia oleifera fruit maturity based on improved YOLOv7 (基于改进 YOLOv7 的油茶果实成熟度检测). *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE),* Vol.40, pp.177-186, Being/China.

[4] Fan, X.; Zhou, J., Xu, Y., Li, K., Wen, D. (2021). Identification and Localization of Weeds Based on Optimized Faster R——CNN in Cotton Seedling Stage (基于优化 Faster R-CNN 的棉花苗期杂草识别与定位). *Transactions of the Chinese Society for Agricultural Machinery*, Vol.52, pp.26-34, Beijing/China.

[5] He, X., Luo, H., Qiao, M., Tian, Z., Zhou, G. (2021). Yield estimation of winter wheat in China based on CNN-RNN network (基于 CNN-RNN 网络的中国冬小麦估产). *Transactions of the Chinese Society of Agricultural Engineering*, Vol.37, pp. 124-132, Beijing/China.

[6] Hou, Q., Zhou, D., Feng, J. (2021), Coordinate Attention for Efficient Mobile Network Design. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* pp.13708-13717, Nashville/USA.

[7] Howard A.G., Zhu M., Chen B. et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, *Computer Science*. USA. https://doi.org/10.48550/arXiv.1704.04861.

[8] Huang, H., Zhang, H., Hu, Xiao. &Nie, X. (2024) Recognition and Localization Method for Pepper Clusters in Complex Environments Based on Improved YOLO v5 (基于改进 YOLO v5 的复杂环境下花椒簇识别与定位方法). *Transactions of the Chinese Society for Agricultural Machinery.* Vol.55, pp.243-251. Beijing/China.

[9] Kang, L., Wei, L. (2022). Progress of acridology in China over the last 60 years (中国蝗虫学研究 60 年). *Journal of Plant Protection,* Vol.49, pp.4-16, Beijing/China.

[10] Kumar, K.S., Abdul Rahman, A. (2021). Early Detection of Locust Swarms Using Deep Learning. *Advances in Machine Learning and Computational Intelligence. Algorithms for Intelligent Systems.* pp.303–310. Springer, Singapore. DOI: 10.1007/978-981-15-5243-4_27

[11] Li, L., Bai, Z., Diao, L., Tang, Z., Guo, X. (2021). Video Detection and Counting Method of East Asian Migratory Locusts Based on K SSD F (基于 K-SSD-F 的东亚飞蝗视频检测与计数方法). *Transactions of the Chinese Society for Agricultural Machinery*, Vol.52, pp.262-266, Beijing/China.

[12] Li, R., Mamat, S, Sheng,Y., He, X.(2023). Identification and application flaunt based on CA-Mobilenet-V2(基于 CA-MobileNet-V2 的核桃病害识别与应用). *Acta Agriculturae Zhejianggensis*, Vol.35, pp.2977-2987, Zhejiang/China.

[13] Liu, F., Wang, S., Pang, S., Han, Z. (2024), Detection and recognition of tea buds by integrating deep learning and image‐processing algorithm. *Journal of Food Measurement and Characterization*, Vol.18, pp.2744–2761, America.

[14] Liu, S., Hu, B., Zhao, C. (2023). Detection and identification of cucumber leaf diseases based improved YOLOv7 (基于改进 YOLOv7 的黄瓜叶片病虫害检测与识别). *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, Vol.39, pp.163-171, Beijing/China.

[15] Ma, C., Zhang, H., Ma, X., Wang, J., Zhang, Y., Zhang, X. (2024). Method for the lightweight detection of wheat disease using improved YOLOv8(基于改进 YOLOv8 的轻量化小麦病害检测方法). *Transactions of the Chinese Society of Agricultural Engineering*, Vol.40, pp.187-195, Beijing/China

[16] Ma, N., Li, Y., Xu, M., Yan, H. (2023). Improved YOLOV8-based automated detection of wheat leaf diseases, *INMATEH - Agricultural Engineering*, Vol. 71, no.3, pp.499-510, Bucharest / Romania.

[17] Ma, H., Zhang, M., Dong, K., Wei, S., Zhang, R., Wang, S. (2022). Research of Locust Recognition in Ningxia Grassland Based on Improved YOLO v5 (基于改进 YOLOV5 的宁夏草原蝗虫识别模型研究) *Transactions of the Chinese Society for Agricultural Machinery,* Vol.53, Beijing/China.

[18] Mu J., Wang J, Liu S, Wang Z, Jiang H, Ma B, Zhang Z, Hu X (2022), A pest accurate segmentation method based on critical point nonlinear enhancement, *INMATEH - Agricultural Engineering*, vol. 68, pp.21-31. Bucharest / Romania. https://doi.org/10.35633/inmateh-68-02

[19] Mao, T, Yu, l, Zhou, X, Yao, T., Wan, W., Xiong, B., Ou, Q. (2023),. Human behavior recognition method in infrared image based on improved MobileNet V1 (基于改进 MobileNetV1 的红外图像人体行为识别方法). *Journal of Liaoning Technical University (Natural Science)*, Vol.42, pp.362-369, Liaoning/China.

[20] Mao, W., Zheng, Y., Zhang,Y., Yuan, Y., Zhang, X. (2008). Grasshopper detection method based on machine vision (基于机器视觉的草地蝗虫识别方法) *Transactions of the Chinese Society of Agricultural Engineering.* Vol.24, pp.155-158, Beijing/China.

[21] Matache, M.G., Marin, F.B., Gurauc., Gurau, G., Marin M. (2022), Găgeanu I., Ionescu A., Neural network testing for spot-application of phytosanitary substances in vegetable crops using a self-propelled electrical sprayer, *INMATEH - Agricultural Engineering*, Vol. 68, pp.471-480, Bucharest / Romania. DOI : https://doi.org/10.35633/inmateh-68-46

[22] Ning, J., Lin, J., Yang, S., Wang,Y., Lan, X. (2023), Face Recognition Method of Dairy Goat Based on Improved YoLo V5s(基于改进 YOLO v5s 的奶山羊面部识别方法). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 4, pp.331-337. Beijing/China.

[23] Wang, L., Liu, Q, Cao, Y., Hao, X. (2023). Posture recognition of group-housed pigs using improved Cascade Mask R-CNN and cooperative attention mechanism (基于改进 Cascade Mask R-CNN 与协同注意力机制的群猪姿态识别). *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, Vol,39. Beijing/China.

[24] Wang, J., Zhou, J., Zhang, Y., Hu, H. (2023). Multi-pose dragon fruit detection system for picking robots based on the optimal YOLOv7 model (基于优选 YOLOv7 模型的采摘机器人多姿态火龙果检测系统) *Transactions of the Chinese Society of Agricultural Engineering*, Vol.39, pp.276-283. Beijing/China.

[25] Wang, X., Xu, Y., Zhou, J., Chen, J. (2023). Safflower picking recognition in complex environments based on an improved YOLOv7(基于改进 YOLOv7 的复杂环境下红花采摘识别) *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE),* Vol.39, pp.169-176. China.

[26] Wang, P., Du, J., Zhang, Y., Liu, J., Li, H., Wang, C. (2024). (基于遥感多参数和 CNN-Transformer 的冬小麦单产估测). *Transactions of the Chinese Society for Agricultural Machinery*, Vol.55, pp.173-182, Beijing/China.

[27] Wu, X., Zhan, C., Lai, K., et al (2019). "IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition,* DOI: 10.1109/CVPR.2019.00899.

[28] Xu, D., Ma, W., Tan, Y., Liu, X., Zheng, Y., Tian, Z. (2022) Yield estimation method for tea buds based on YOLOv5 deep learning (基于 YOLO v5 深度学习的茶叶嫩芽估产方法). *Journal of China Agricultural University*, Vol.27, pp.213-220. Beijing/China.

[29] Yu, H., Shi, W. (2021) Outbreak monitoring and control technology of desert locust Schistocerca gregaria (沙漠蝗灾发生、监测及防控技术进展). *Journal of Plant Protection,* Vol.48, pp.28-36. Beijing/China.

[30] Yu, J., Chen, W., Guo, Y., Mu, Y., Fan, C. (2024). Improved Oriented R-CNN-based model for oriented wheat ears detection and counting (基于改进 Oriented R-CNN 的旋转框麦穗检测与计数模型). *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE),* Vol.40, pp.248-257, Beijing/China.
[31] Zhang, G., Lyu, Z., Liu, H., Liu, W., Long, C., Huang, C. (2023). Model for identifying lotus leaf pests and diseases using improved DenseNet and transfer learning (基于改进 DenseNet 和迁移学习的荷叶病虫害识别模型). *Transactions of the Chinese Society of Agricultural Engineering.* Vol.39, pp.188-196. Beijing/China
[32] Zhao, H., Cao, Y., Yue, Y., &Wang, H. (2021). Field weed recognition based on improved DenseNet (基于改进 DenseNet 的田间杂草识别). *Transactions of the Chinese Society of Agricultural Engineering*, Vol.37, pp.136-142, Beijing/China.
[33] Zheng, Y., Wu, G., Wang, Y.& Mao, W. (2010) Locust images detection based on fuzzy pattern recognition (基于模糊模式的蝗虫图像识别方法). Transactions of the Chinese Society of Agricultural Engineering Vol.26, pp. 21-25, Beijing/China.
[34] Zhou, W., Niu, Y., Wang, Y, Li, D. (2022). Rice pests and diseases identification method based on improved YOLOv4-GhostNet *Jiangsu J. of Agr.* Vol.38, pp.685-695. Jiangsu/China.