Maturity detection of *Hemerocallis citrina Baroni* based on LTCB YOLO and lightweight and efficient layer aggregation network

Le Chen[1], Ligang Wu[2*], Yeqiu Wu[3]

(1. *College of Coal Engineering, Shanxi Datong University, Datong 037003, Shanxi, China*;
2. *College of Mechanical and Electrical Engineering, Shanxi Datong University, Datong 037003, Shanxi, China*;
3. *College of Architecture and Surveying Engineering, Shanxi Datong University, Datong 037003, Shanxi, China*)

Abstract: *Hemerocallis citrina Baroni* is rich in nutritional value, with a clear trend of increasing market demand, and it is a pillar industry for rural economic development. *Hemerocallis citrina Baroni* exhibits rapid growth, a shortened harvest cycle, lacks a consistent maturity identification standard, and relies heavily on manual labor. To address these issues, a new method for detecting the maturity of *Hemerocallis citrina Baroni*, called LTCB YOLOv7, has been introduced. To begin with, the layer aggregation network and transition module are made more efficient through the incorporation of Ghost convolution, a lightweight technique that streamlines the model architecture. This results in a reduction of model parameters and computational workload. Second, a coordinate attention mechanism is enhanced between the feature extraction and feature fusion networks, which enhances the model precision and compensates for the performance decline caused by lightweight design. Ultimately, a bi-directional feature pyramid network with weighted connections replaces the Concatenate function in the feature fusion network. This modification enables the integration of information across different stages, resulting in a gradual improvement in the overall model performance. The experimental results show that the improved LTCB YOLOv7 algorithm for *Hemerocallis citrina Baroni* maturity detection reduces the number of model parameters and floating point operations by about 1.7 million and 7.3G, respectively, and the model volume is compressed by about 3.5M. This refinement leads to enhancements in precision and recall by approximately 0.58% and 0.18% respectively, while the average precision metrics mAP@0.5 and mAP@0.5:0.95 show improvements of about 1.61% and 0.82% respectively. Furthermore, the algorithm achieves a real-time detection performance of 96.15 FPS. The proposed LTCB YOLOv7 algorithm exhibits strong performance in detecting maturity in *Hemerocallis citrina Baroni*, effectively addressing the challenge of balancing model complexity and performance. It also establishes a standardized approach for maturity detection in *Hemerocallis citrina Baroni* for identification and harvesting purposes.

Keywords: *Hemerocallis citrina Baroni*, maturity detection, YOLOv7, lightweight model, efficient layer aggregation network

DOI: 10.25165/j.ijabe.20251802.9238

1   Introduction

The market demand for *Hemerocallis citrina Baroni* is steadily rising, with the industrial chain becoming increasingly refined. This trend plays a vital role in the development of emerging rural sectors and the revitalization of rural areas.

The unique ripening and harvesting period for *Hemerocallis citrina Baroni* predominantly occurs in the morning and evening. Currently, the determination of maturity heavily relies on manual labor, which presents several challenges. The picking process is highly dependent on the work experience of workers, leading to high labor costs, low efficiency, and inconsistent standards. These issues significantly affect the accuracy and productivity of the harvesting process. Therefore, there is a pressing need for an automated approach to improve the efficiency and standardization

of *Hemerocallis citrina Baroni* picking, making the use of computer vision for maturity detection a promising solution. However, there are some differences in the identification process of manual picking, and the identification standard is not uniform, so the identification with the help of computer vision has a better application prospect.

In recent years, the performance of target detection technology and deep learning algorithms has significantly improved, leading to a wider application in the agricultural industry[1]. Shi et al.[2] realized high precision and efficiency detection by improving the YOLOv4 algorithm for the problem of small variability of fig fruits and dense plants. Wu et al.[3] used the YOLOv5 algorithm to identify cherry trees, detect pests and diseases affecting fruit trees, and monitor cherry ripening. In addition, Wang et al.[4] and Wu et al.[5] incorporated the attention mechanism into the YOLO algorithm, changing the pattern of feature extraction, and accomplished the detection of tomatoes and apples, avoiding pests and improving the yield.

Hui et al.[6] introduced the normalized attention module to improve the YOLOX algorithm, and carried out real-time monitoring of pollination, fertilization, picking, and other phases of strawberry growth. The results of the experiment showed an improvement in both detection precision and recall rate, while keeping the original model size unchanged. Xu et al.[7] used

lightweight Shuffle Net to simplify the model structure, reduce the computer hardware requirements and consumption, and complete aphid detection during the growth of sugarcane, realizing the double improvement of detection efficiency and precision. Liu et al.[8] and Ngugi et al.[9] utilized the YOLO algorithm and convolutional neural network to implement a supplementary residual network and depth-separable convolution within the feature extraction network, which compressed the model volume, improved the real-time recognition efficiency, accurately recognized the growth condition of grapes, and enhanced the efficiency of the grape-picking robot[10].

Besides, in order to assess the yield composition and proportion of wheat, Wu et al.[11] completed the detection and counting of multi-scene and multi-scale wheat seeds based on the deep learning algorithm, and accurately judged the quality and yield of wheat. Zhang et al.[12] introduced lightweight convolution and SE attention mechanism[13] of the YOLOv5 model, completed maturity detection, and optimized the model performance while reducing the model complexity. Different from the above studies, Song et al.[14] innovatively carried out deep learning algorithms on a UAV[15], realizing UAV remote sensing monitoring of maize's growth, reducing the cost and difficulty of monitoring, and improving the monitoring efficiency.

In the field of agriculture, smart farming is advancing due to the widespread use of computer technology and sophisticated deep learning algorithms[16]. However, current neural networks have complex model structures, redundant feature extraction, and unsatisfactory detection precision and recall. Additionally, when faced with challenges such as dense vegetation, mutual obstruction, and the complexity of determining maturity in detecting *Hemerocallis citrina Baroni*, an algorithm for *Hemerocallis citrina Baroni* maturity detection named LTCB (Lightweight efficient layer aggregation networks - Transition module - Coordinate attention mechanism - Bidirectional feature pyramid network) YOLOv7 has been developed. The key advancements presented in this study include:

(1) The Ghost convolution[17] method is presented in the Efficient Layer Aggregation Network (ELAN) and Transition Module (TM) to streamline feature extraction, minimize redundant feature maps, and lower computer hardware demands and energy consumption.

(2) The Coordinate Attention (CA) mechanism[18] is incorporated between the feature extraction and feature fusion network to facilitate two-dimensional feature extraction. This mechanism continuously updates the feature weights, thereby enhancing the feature extraction process and improving the model's localization and classification capabilities.

(3) Within the feature fusion module, the utilization of the Bidirectional feature pyramid network (Bi FPN)[19] is preferred over Concatenate. This approach serves to decrease the quantity of edge nodes while simultaneously enhancing the information fusion channels. Consequently, it facilitates the effective integration of multiple information sources across the channels, thereby enhancing the detection accuracy of the model.
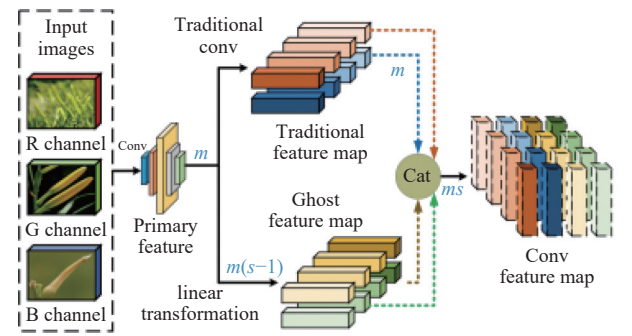
The subsequent sections of this document are organized as follows: Section 1 presents the foundational theory underpinning the utilized modules. Section 2 outlines the enhanced detection algorithm along with the steps taken to achieve improvements. Section 3 assesses the efficacy of different models and scrutinizes the results obtained. Finally, Section 4 provides a summary of the thesis research.

2   Basic theory

2.1   Lightweight Ghost convolution

Compared with traditional convolution, the Ghost convolution with linear transformation has lower computational effort, realizes feature perception, and adaptively activates the detected object. The Ghost convolution can significantly improve network performance and provide better real-time capabilities than other networks under similar conditions.

The lightweight Ghost convolutional structure is shown in Figure 1. Initially, the input images undergo convolution to create the primary feature map, where the key attributes are processed through standard convolution and Ghost linear transformation to generate the corresponding conventional feature map and Ghost feature map. Last, the concatenate information fusion method fuses the two different feature maps to achieve the convolutional feature map output.

Note: RGB refers to the three traditional color channels, red, green, and blue, which represent the basic color components of the input image.

Figure 1   Lightweight Ghost convolutional structure

Assume that in the convolution process, the input image size is $x \in R^{c \times h \times w}$, the output image size is $y \in R^{h' \times w' \times ms}$, the number of input and output channels are c and ms, respectively, and the size of the custom convolution kernel is k. Therefore, the comparison of computational complexity between conventional convolution C_T and lightweight Ghost convolution C_G is presented in the following ratio:

$$r_s = \frac{C_T}{C_G} = \frac{c \cdot k \cdot k \cdot ms \cdot h' \cdot w'}{c \cdot k \cdot k \cdot m \cdot h' \cdot w' + m \cdot k \cdot k \cdot (s-1) \cdot h' \cdot w'} = \frac{c \cdot s}{c + s - 1} \approx s$$

The computational burden of the Ghost convolution, which is a lighter version, is roughly $1/s$ times lower than the traditional convolution method. This results in a notable decrease in computational demands during the convolution operation, facilitating the development of a more resource-efficient design.

2.2   Mechanism of coordinated attention

The Coordinate Attention (CA) mechanism integrates channel information and orientation-specific positional data, effectively combining positional information with channels to retain spatial characteristics and address issues related to long-range dependencies.

The CA mechanism is depicted in Figure 2. Initially, the input characteristics (input feature) are subjected to global average pooling in both the horizontal (X) and vertical (Y) directions, producing feature maps with dimensions of $C \times H \times 1$ in width and $C \times 1 \times W$ in height. Following this, the data is passed through the Batch Normalization (BN) layer and a non-linear activation function, accelerating the model's convergence, preventing overfitting, and mitigating issues related to gradient vanishing. Finally, the intermediate features undergo two parallel stages of 2D

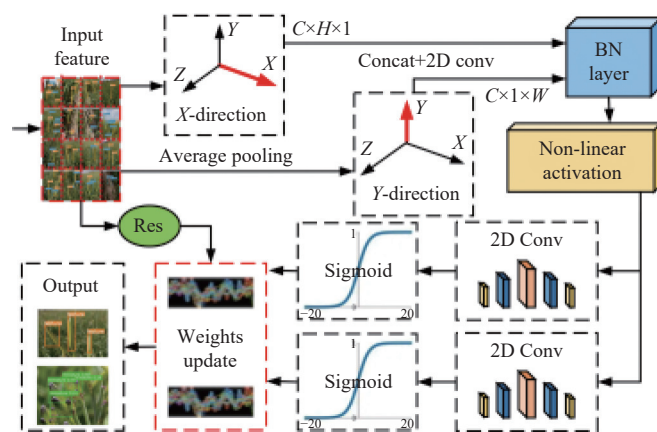convolution and sigmoid activation to complete the feature extraction.

Figure 2   CA mechanism process

2.3   Bi-directional feature pyramid network

Bi FPN is improved and optimized based on a feature pyramid network (FPN) with efficient information fusion capability. Compared with Concatenate, the weights of the network nodes are increased to realize weighted feature fusion.

Figure 3 depicts the Bi FPN, which streamlines the network nodes between layers, diminishes the significance of edge information nodes, enhances the information fusion channel across stages, and assigns varying weights to individual network layers for feature fusion. This approach allows the model to emphasize the importance of each network layer and elevates the precision of *Hemerocallis citrina Baroni* maturity detection[20].
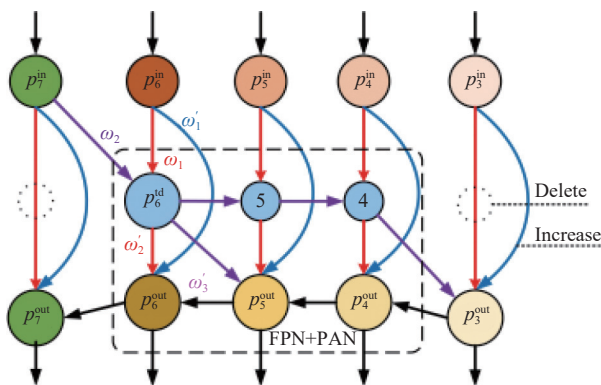

Figure 3   Structure of the Bi FPN module

3   Improved detection method

3.1   Lightweight efficient layer aggregation network

The ELAN consists of two pathways for feature extraction: the backbone pathway and the branch pathway. The improved Lightweight Efficient Layer Aggregation Network (L-ELAN) incorporates a lightweight Ghost convolution at a specific juncture to decrease the model's parameter quantity and computational burden. This adjustment is implemented in consideration of the significance of the feature extraction process and convolution operation.

The L-ELAN module, as depicted in Figure 4, involves a series of five conventional convolutions in the backbone path to perform stacked feature extraction on the input features, resulting in the acquisition of profound semantic features. Second, in the branch path, the input features undergo a lightweight Ghost convolution and produce shallow detail features containing color, position, and

shape. Last, the deep semantic features and the shallow detailed features undergo the concatenate operation to realize the information fusion. The fused feature information undergoes lightweight Ghost convolution to complete the feature output.
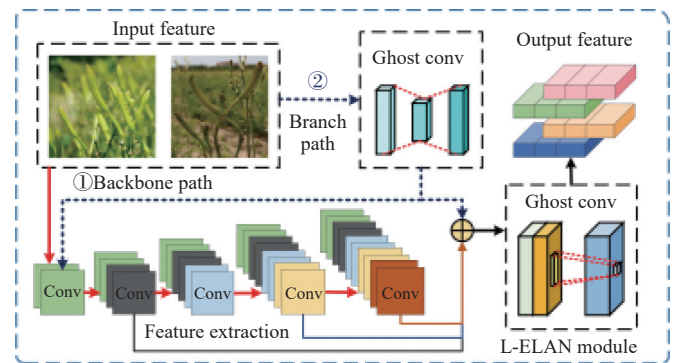

Figure 4   Lightweight efficient layer aggregation network

The improved L-ELAN module not only maintains the capability of the original network feature extraction, but also is lightweight and compresses the model volume.

3.2   Lightweight transition module

The Transition Module (TM) also consists of two feature extraction paths, but it has fewer convolution operations and a simple model structure. Unlike ELAN, the TM adds a maximum pooling operation in the backbone path, which reduces the amount of model data and preserves important feature information.

There exist two categories of transition modules within the context of this study: the feature extraction network transition module and the feature fusion network transition module. Both of these modules are characterized by a streamlined design approach.

As shown in Figure 5, the improved Lightweight Transition Module (L-TM) replaces the traditional conv after maximal pooling with the lightweight Ghost conv, which further filters the feature information after pooling, avoiding redundancy of information while retaining important information.
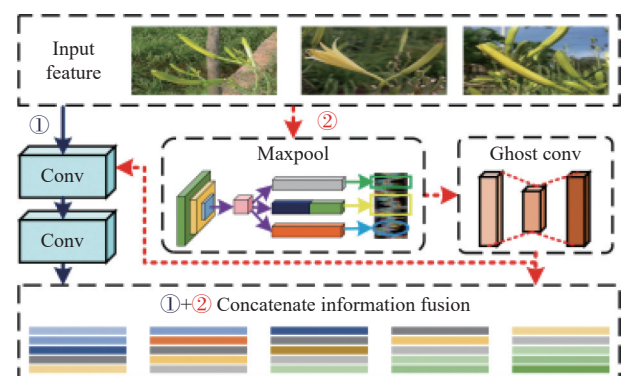

Figure 5   Lightweight transition module

3.3   Increased CA mechanism

The enhanced LTCB YOLOv7 *Hemerocallis citrina Baroni* maturity detection algorithm integrates a CA mechanism between the feature extraction and feature fusion network. This mechanism is implemented to highlight crucial information during feature extraction and alleviate any decrease in performance due to the system's lightweight architecture.

Following the integration of the CA mechanism, the model can dynamically acquire channel weights, allowing it to prioritize both channel and location information simultaneously. This enhancement significantly contributes to the overall generalization capability of

the model.

3.4 Realization of multi-information fusion

During the convolution process, larger objects intended for detection exhibit a higher pixel count, while smaller objects possess a lower number of pixels. As the convolution progresses, the detailed features of larger objects are more likely to be preserved, while those of smaller objects are more likely to be overlooked. Therefore, the application of Bi FPN instead of Concatenate enables cross-channel information fusion to preserve feature information of different sizes at different stages.

In *Hemerocallis citrina Baroni* maturity detection, there are different sizes, and *Hemerocallis citrina Baroni* growing on stalks of different heights have uneven sizes under the same receptive field[21]. Therefore, in order to improve the detection precision, this paper applies Bi FPN in the feature fusion module to perform simple and fast multi-scale feature fusion.

The feature fusion module, depicted in Figure 6, is essential in the process of detecting maturity in *Hemerocallis citrina Baroni*. In this context, the deep feature information within the feature fusion network undergoes expansion and amplification through conventional convolution and up-sampling[22] techniques. Conversely, the shallow feature information within the feature extraction network is responsible for completing the feature extraction process using traditional convolution methods.
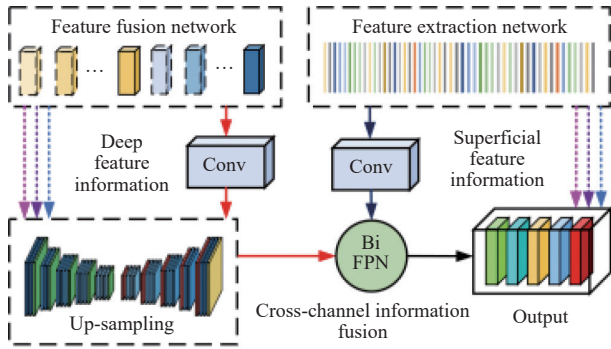

Figure 6   Module for feature fusion

Subsequently, the deep feature information and shallow feature information undergo Bi FPN to complete simple and efficient cross-channel information fusion and realize feature output.
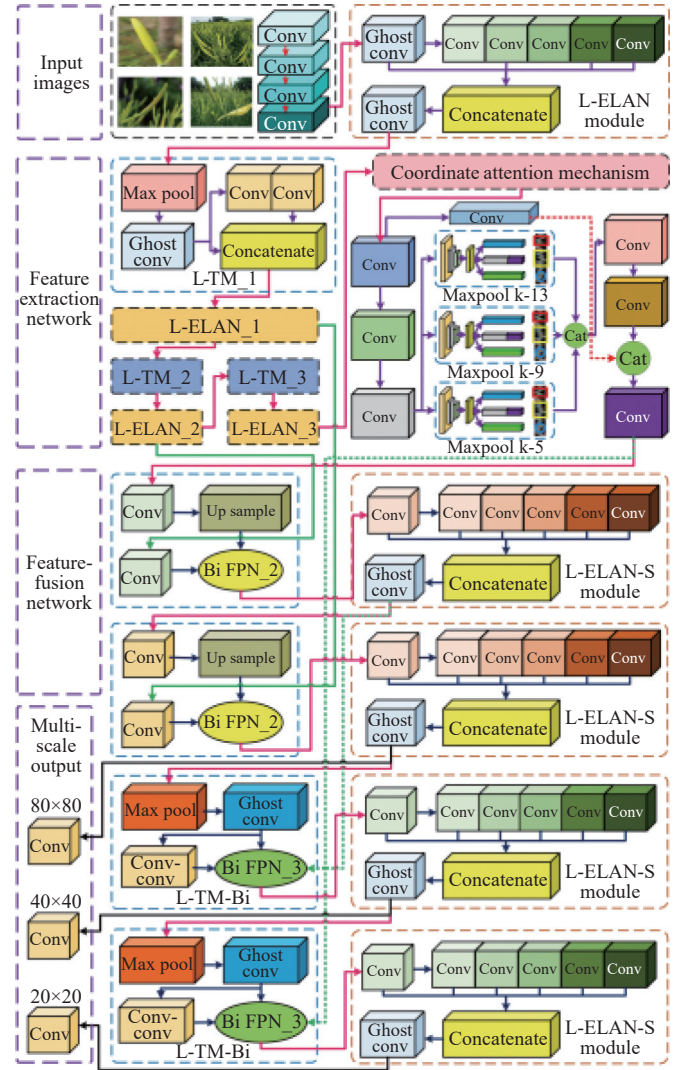
3.5 Improved LTCB YOLOv7 algorithm

The enhanced LTCB YOLOv7 algorithm comprises five key elements: the feature extraction network, the CA mechanism, the Spatial Pyramid Pooling-Cross Stage Partial Channel (SPP-CSPC) module, the feature fusion network, and the multi-scale outputs.

Within the feature extraction network, the computational and parametric burdens are lessened in L-ELAN and L-TM through the incorporation of a lightweight Ghost convolution. Enhancements have been made to the connectivity attention mechanism linking the feature extraction and feature fusion networks, facilitating the effective exchange of information between these elements and enabling the dynamic modification of weights[23].

In the feature fusion network, optimizing the feature fusion technique involves replacing Concatenate with Bi FPN in the transition module to facilitate cross-stage information fusion. As a result, the modified lightweight transition module, after integrating Bi FPN, is denoted as L-TM-Bi.

As illustrated in Figure 7, the input image is randomly cropped, segmented, and spliced to complete the data preprocessing, and feature extraction is completed by the L-ELAN and L-TM modules

alternately after four traditional convolutions in the feature extraction network.

Note: The ways of concatenate information fusion are different between L-ELAN-S and L-ELAN. In the feature extraction path, L-ELAN intermittently fuses three traditional convolutions in the backbone path, while L-ELAN-S sequentially fuses five traditional convolutions in the backbone path.

Figure 7   LTCB YOLOv7 algorithm structure

The CA mechanism is capable of adaptively updating the weights. The SPP-CSPC module bridges the feature extraction and feature fusion networks, and performs maximum pooling by three different sizes of convolutional kernels, 5×5, 9×9, and 13×13, which remove the redundant information and compress and downscale the features. At the multi-scale output prediction end, feature fusion of different depths produces output predictions of different sizes as the convolution continues.

In *Hemerocallis citrina Baroni* maturity detection, different sizes of output prediction frames play different roles. 256-channel 80×80-size prediction frames are able to detect mature and intact *Hemerocallis citrina Baroni* at the shoot stage, while 512-channel 40×40- and 1024-channel 20×20-size prediction frames are able to detect immature *Hemerocallis citrina Baroni* at the leaf spreading stage and the sprouting stage, respectively[24]. The multi-scale output prediction effectively avoids the occurrence of the phenomenon of missed detection, which makes the detection of *Hemerocallis citrina Baroni* maturity under different growth cycles more targeted.

Hemerocallis citrina Baroni maturity detection poses unique

challenges compared to other object detection tasks. These challenges include complex growing environments, severe occlusions, and significant fruit size and shape variations. The dynamic and diverse conditions under which *Hemerocallis citrina Baroni* grows further complicate the detection process, requiring the development of robust models capable of dealing with such variations and environmental factors. These unique characteristics make *Hemerocallis citrina Baroni* maturity detection a particularly difficult task, highlighting the need for tailored methods like LTCB YOLO for effective maturity detection.

4   Experiments and analysis of results

4.1   Experimental environment platform

This study's investigation used the Python 3.9.18 and CUDA 11.8 frameworks. The computational resources utilized were an Intel Core i9-10900k processor operating at 3.7 GHz, an NVidia GeForce RTX 4090, 24564MiB.

The dimensions of the image input were set at 6960×4640 pixels, with an initial learning rate of 0.01 and a final One-Cycle learning rate of 0.1. Image augmentation strategies involved a 50% likelihood of horizontally flipping the image and the application of HSV value augmentation coefficients of 0.4 and saturation enhancement coefficients of 0.7. The training procedure encompassed 200 epochs utilizing a batch size of 10.

4.2   Model training

The images of *Hemerocallis citrina Baroni* were collected on site in Yunzhou District, Datong City, Shanxi Province, the hometown of *Hemerocallis citrina Baroni*, captured by the high-definition equipment Canon EOS 90D, with a resolution size of 6960×4640 pixels. 795 images were selected as the dataset through screening and organizing. This study's dataset comprising 795 images is divided into three subsets: 597 images for training, 148 for validation, and 50 for testing. The loss function is employed as a metric during the model's training to evaluate its performance and guide the learning process.

The characteristics of *Hemerocallis citrina Baroni* are influenced by different weather conditions, lighting, and background complexity. For instance, under warm sunlight or in a complex background, shading and lighting conditions resembling *Hemerocallis citrina Baroni* features can increase identification difficulty. Conversely, the targeted features are more distinct and shaped under low light with a simple background, making feature areas easier to detect.

The annotation process was carried out using LabelImg software. The researchers manually annotated the dataset by drawing bounding boxes around the target areas in each image, strictly following the edge range of the targets to minimize background interference. The annotation mode was set to YOLO format, ensuring compatibility with the model's input requirements. Each annotation file was saved in txt format, corresponding one-to-one with the image file. These files contained the target's category number and normalized coordinates, including the x and y coordinates of the center point and the width and height of the bounding box.

The loss function applied during training directly impacts the model's optimization and prediction precision. Figure 8 demonstrates the convergence behavior of the loss function for the LTCB YOLOv7 algorithm in identifying the maturity of *Hemerocallis citrina Baroni*. The improved algorithm shows a slightly faster convergence rate during the initial pre-training stage compared to the original YOLOv7, while both follow a similar

convergence pattern throughout training. The final convergence value of the loss function is approximately 0.0310.
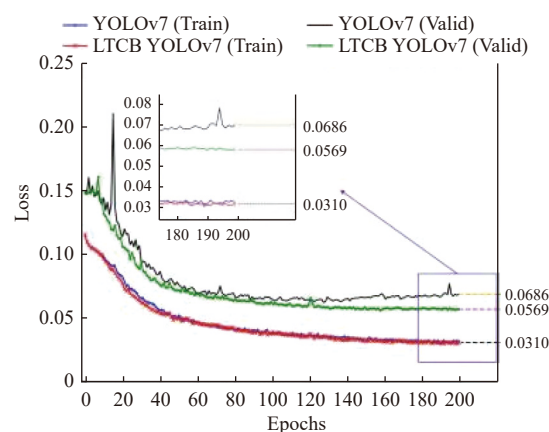

Figure 8   Loss function convergence curve

The loss function during validation further highlights the improved performance of the LTCB YOLOv7 algorithm. As shown in Figure 8, the LTCB YOLOv7 algorithm achieves a superior convergence value of approximately 0.0569, compared to 0.0117 for the original YOLOv7 algorithm. This result indicates that the enhanced LTCB YOLOv7 algorithm provides more accurate and efficient detection of *Hemerocallis citrina Baroni* maturity.

4.3   Ablation experiments

In the realm of deep learning, particularly within intricate deep neural networks, ablation experiments are utilized to demonstrate the essentiality of specific modules[25]. In the study involving *Hemerocallis citrina Baroni* maturity detection, researchers utilized lightweight and efficient layer aggregation networks to conduct a series of seven ablation experiments. These experiments underscore the importance of lightweight convolution, attention mechanism, and cross-channel multi-information fusion in enhancing the detection process.

As shown in Table 1, based on the YOLOv7 model, Model 1 and Model 2 introduce L-ELAN and L-TM in the feature extraction network in turn. Based on Model 2, Models 3 and 4 introduce L-ELAN and L-TM sequentially in the feature fusion network. However, the gradual realization of lightweight detection accuracy has been reduced to varying degrees.

Table 1   Ablation experiments

| Model | Feature extraction network | | Feature fusion network | | CA | Bi FPN |
	L-ELAN	L-TM	L-ELAN	L-TM		
YOLOv7						
Model 1	√					
Model 2	√	√				
Model 3	√	√	√			
Model 4	√	√	√	√		
Model 5	√	√	√	√	√	
LTCB YOLOv7	√	√	√	√	√	√

Based on the framework of Model 4, Model 5 enhances the CA mechanism, leading to a notable enhancement in model performance. This paper introduces the LTCB YOLOv7 model, which integrates Bi FPN into Model 5, resulting in a refined precision in object detection.

4.4   Lightweight analysis

The operability of large-scale detection equipment is seriously affected by the vast cultivation area and dense plants in the field of

Hemerocallis citrina Baroni. Therefore, simplifying structure and reducing computation of the model meets the development needs of smart agriculture, and it is also the future development direction of neural networks[26].

The enhanced LTCB YOLOv7 algorithm offers clear benefits in terms of being lightweight, and a comparison of the lightweight characteristics of various models is presented in Table 2.

Table 2   Model lightweight analysis

Model	Network layers	Parameters/ Millions	Flops/ G	Time/ h	Model volume
YOLOv5	468	46.14	107.9	8.067	92.7
YOLOv7	415	37.20	104.8	8.025	74.8
Model 1	455	35.82	99.0	7.816	72.1
Model 2	470	35.48	98.4	7.616	71.4
Model 3	490	34.68	96.7	7.724	69.8
Model 4	500	34.64	96.7	7.776	69.7
Model 5	508	35.43	97.5	7.690	71.3
Model 6	473	36.17	97.2	8.333	73.7
Model 7	621	71.29	100.3	10.537	124.5
LTCB YOLOv7	508	35.43	97.5	7.703	71.3

The table demonstrates that the YOLOv7 algorithm utilizes a multi-branch stacking strategy, resulting in a decrease of 53 network layers in comparison to the YOLOv5 algorithm. This adjustment also results in a reduction in the number of parameters from 46.14 million to 37.20 million. Furthermore, the floating-point operations decrease from 107.9 G to 104.8 G, and the model size is compressed by 17.9 M.

In the research conducted on identifying maturity in *Hemerocallis citrina Baroni*, the LTCB YOLOv7 algorithm was augmented with an additional 93 network layers in comparison to the original YOLOv7 algorithm. Despite this augmentation, there has been a reduction in the number of model parameters and training duration to different degrees, leading to a more compact model size. Empirical results suggest that the LTCB YOLOv7 algorithm has decreased the count of parameters and floating-point operations by around 4.76% and 6.97%, correspondingly. The model training time has been shortened from 8.025 hours to 7.703 h, and the model size has been compressed from 74.8 M to 71.3 M.

Model 6 and Model 7 are modeled using the backbone network for EfficientNet[19] and MobileNetv3[27] based on the fusion CA mechanism and the updated Bi FPN information fusion mechanism. Compared with Model 6 and Model 7, LTCB YOLOv7 shows significant advantages in terms of modeling efficiency, computational cost, and practical deployment. Although Model 6 has fewer network layers, its training time of 8.333 hours and model volume of 73.7 MB are higher than those of the LTCB YOLOv7 model. Model 7 is not as efficient in lightweight scenarios, with a deeper network, a parameter count of 71.29 M, and a model volume of 124.5 MB, which results in a computation time of 10.537 hours, greatly limiting its performance in resource-limited environments. In contrast, the LTCB YOLOv7 model introduces Ghost convolutional optimization for lightweighting, which is particularly suitable for real-time detection tasks in agricultural environments, combining a lightweight design with high efficiency and robust performance.
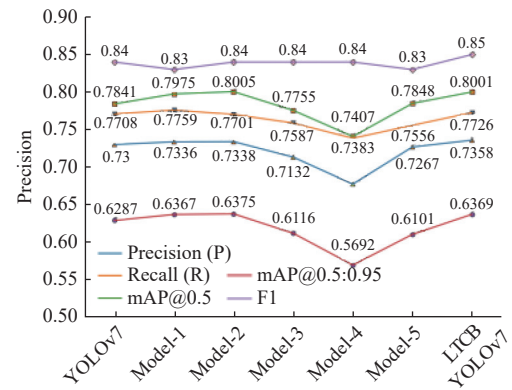
4.5   Precision analysis

The food and medicinal applications of *Hemerocallis citrina Baroni* vary due to different maturity levels, which directly affects the economic benefits. Therefore, it is important to precisely determine the maturity of *Hemerocallis citrina Baroni*[28].

Frequently employed performance evaluation metrics encompass precision (P), recall (R), $F1$ score, receiver operating characteristic (R-P) curve, mean average precision at intersection over union (mAP@0.5), and mean average precision across different intersection over union thresholds (mAP@0.5:0.95). The $F1$ score, a metric derived from the harmonic mean of precision and recall, is sensitive to the confidence threshold and offers a holistic evaluation of the model's effectiveness. The area under the R-P curve signifies the average precision (AP) attained during the training of the model.

The metric mAP@0.5 denotes the mean average precision calculated for different categories where the intersection over union (IoU) ratio between predicted and labeled boxes exceeds 0.5. Conversely, mAP@0.5:0.95 refers to the average precision mean computed across diverse categories using IoU thresholds that vary from 0.5 to 0.95 in increments of 0.05, yielding a set of 10 distinct values.

In Figure 9, the evaluation results of the precision performance index for various models detecting maturity in *Hemerocallis citrina Baroni* are presented. The figure demonstrates that as models are refined and optimized, the performance indices within each model typically exhibit a consistent trend of both increasing and decreasing.

Figure 9   Comparison of the precision of different models

After the gradual introduction of the lightweight Ghost convolution, although Model 1 to Model 4 achieved lightweight, the performance indices showed a decreasing trend[29]. Compared with the YOLOv7 algorithm, the performance of Model 4 after lightweight decreases significantly. The P is reduced from 0.73 to 0.6772, and the R is also reduced from 0.7708 to 0.7383. Furthermore, there is a decrease of approximately 0.0434 in mAP@0.5 and 0.0595 in mAP@0.5:0.95. The $F1$ harmonic mean retains its value of 0.84, in line with the initial YOLOv7 algorithm.

The devised LTCB YOLOv7 algorithm increases the CA mechanism and Bi FPN on the basis of Model 4, which not only maintains the advantage of lightweight, but also exceeds the original algorithm in each precision index. The P, R, mAP@0.5, and mAP@0.5:0.95 of the LTCB YOLOv7 algorithm are 0.7358, 0.7726, 0.8001, and 0.6369, respectively. Compared to the original algorithm, these four indices improved by about 0.0058, 0.0018, 0.016, and 0.0082, respectively. At a confidence level of 0.583, the LTCB YOLOv7 model demonstrates a notable improvement with a harmonic mean $F1$ score of 0.85.

The R-P curve is a visual depiction in which the horizontal axis represents R and the vertical axis represents P. This graphical representation elucidates the correlation between P and R,

demonstrating the fluctuations in model precision with changes in recall.

The relationship between the variables *P* and *R* is illustrated in Figure 10, indicating a negative correlation between the two. The graph illustrates that the LTCB YOLOv7 algorithm shows a significantly greater area under the *R-P* curve in comparison to the conventional YOLOv7 algorithm, leading to an increased average precision. Specifically, the average precision of the YOLOv7 algorithm is 91.3%, whereas the enhanced LTCB YOLOv7 algorithm achieves a higher average precision of 91.7%, surpassing the performance of the original algorithm.
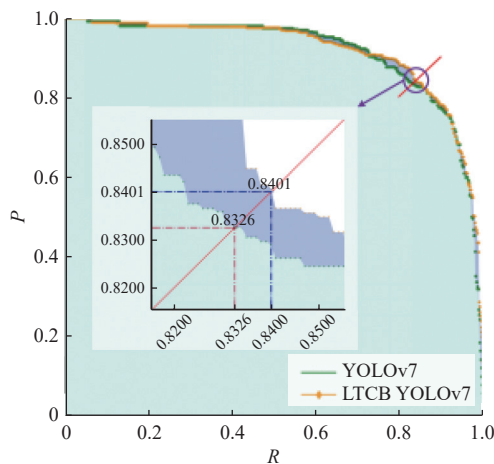

Figure 10   Comparison of *R-P* curves

When the value of *P* is equivalent to *R*, the equilibrium line exhibits a slope of 1 and intersects the *R-P* curve at two points. The equilibrium line intersects the YOLOv7 algorithm at the point (0.8326, 0.8326) and the LTCB YOLOv7 algorithm at the point (0.8401, 0.8401). It can be seen that when the model *P* is balanced with the *R*, the LTCB YOLOv7 model has a better balancing performance, with both *P* and *R* at the same time in a relatively optimal state. Eventually, not only was the detection precision improved[30], but the leakage rate was also reduced.

4.6   Comparison of experimental results

Detecting process maturity can be challenging due to factors such as light exposure, environmental conditions, and growth variations. Hence, the YOLOv7 algorithm incorporates a non-maximum suppression (NMS) technique to enhance the model's local search capability. This strategy effectively eliminates non-maximum elements, thereby enhancing detection accuracy and reducing the occurrence of redundant detections[31].

In order to evaluate the model's performance comprehensively, the detection of *Hemerocallis citrina Baroni* was performed under different lighting conditions, occlusion degrees, and target sizes.

The detection effectiveness was evaluated using a confidence threshold and an Intersection over Union (IoU) threshold set at 0.2 concurrently to validate the model's detection performance. The *Hemerocallis citrina Baroni* was detected in four scenarios: under normal lighting conditions, during early morning, in the evening, and on a rainy day. The detection results are depicted in Figure 11, with the YOLOv7 algorithm's detection outcomes shown on the left side of the diagram and the improved LTCB YOLOv7 algorithm's detection results displayed on the right side.

In the light scene, the LTCB YOLOv7 algorithm introduces the CA mechanism, which is able to detect more occluded *Hemerocallis citrina Baroni* in the complex scene. Meanwhile, the Bi FPN is used to increase the information fusion channel and

reduce the edge leakage detection rate. So, in heat maps, the improved LTCB YOLOv7 algorithm has better performance in detecting large and obviously mature *Hemerocallis citrina Baroni*, and the thermographic detection is more obvious.

In early morning conditions, when light is insufficient, the LTCB YOLOv7 algorithm excels by focusing on mature *Hemerocallis citrina Baroni* with higher confidence. It demonstrates improved detection accuracy with a confidence score of 0.93 for mature *Hemerocallis citrina Baroni*, compared to 0.63 for YOLOv7.

Meanwhile, since early morning is the best time to pick *Hemerocallis citrina Baroni*, the LTCB YOLOv7 algorithm was very effective in heat map detection, realizing a near-fitting detection of mature *Hemerocallis citrina Baroni*.

The LTCB YOLOv7 algorithm decreases the loss function value during both model training and validation. It also enhances the NMS during the detection phase. Consequently, it attains a higher level of accuracy in identifying the maturity of *Hemerocallis citrina Baroni* by precisely determining the ideal bounding box and eliminating redundant localization boxes.

In the evening scene illustrated in Figure 11c, the YOLOv7 algorithm detects an early-stage and fully developed instance of *Hemerocallis citrina Baroni* utilizing a common NMS technique. The immature detection is reported with a confidence level of 0.87, while the mature detection is indicated with a confidence level of 0.41. The improved LTCB YOLOv7 algorithm gives only one mature detection with a confidence level of 0.87. The improved LTCB YOLOv7 algorithm decreases misdetection incidents, featuring a more pronounced heat map detection and enhanced non-maximal suppression screening capability.

In inclement weather conditions, the YOLOv7 algorithm successfully detected the mature stage of *Hemerocallis citrina Baroni* with a confidence level of 0.46. Conversely, the LTCB YOLOv7 algorithm accurately recognized the plant's maturity with a higher confidence level of 0.94. The LTCB YOLOv7 algorithm, with the addition of the lightweight Ghost convolution, streamlines feature extraction, simplifies the process, and prevents false detections.

Subsequently, in heat map detection, combined with image analysis, *Hemerocallis citrina Baroni* was significantly yellower and matured. In contrast to the initial algorithm, the enhanced LTCB YOLOv7 algorithm offers a broader detection scope and a more consistent and well-suited thermal detection distribution.

Furthermore, to demonstrate the effectiveness of the model, tests were carried out on the *Hemerocallis citrina Baroni* dataset. These experiments involved comparing the LTCB YOLOv7 model introduced in this study with the Ghost-SE YOLOv5 model[32], the YOLO-CBAM model from Okafor et al.[33], and the YOLOv7-TP model from Du et al.[34]

Zhang et al. and Okafor et al.[32,33] added SE and CBAM mechanisms to the YOLOv5 algorithm, respectively. However, Du et al.[34] introduced the lightweight pruned method based on the YOLOv7 algorithm. The incomplete findings from the comparison experiments are displayed in Table 3. However, the model structural parameters and experimental results may vary slightly depending on the training equipment and parameters.

As can be seen from Table 3, both Ghost-SE YOLOv5 and YOLO-CBAM models are based on the YOLOv5 algorithm, which has a simple structure and a small number of parameters, but their performance is not ideal. Despite enhancements in all aspects, the YOLOv7-TP model demonstrates precision and recall rates of
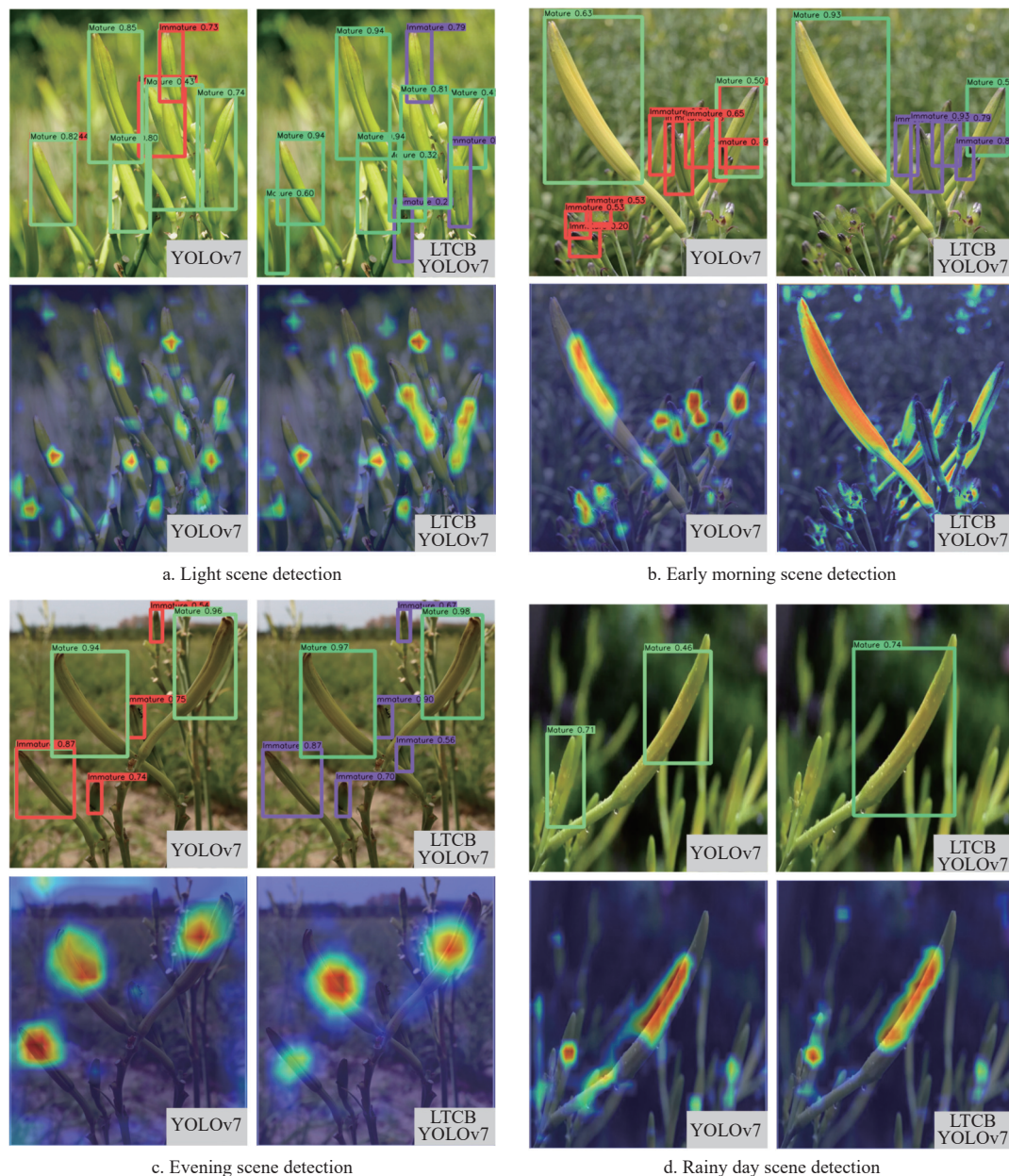
a. Light scene detection

b. Early morning scene detection

c. Evening scene detection

d. Rainy day scene detection

Figure 11   Comparison of different detection results

Table 3   Comparison experiments

Model	Ghost-SE YOLOv5[32]	YOLO-CBAM[33]	YOLOv7-TP[34]	LTCB YOLOv7
GPU/G	6.43	3.45	7.74	7.20
Flops/G	52.5	60.1	96.7	97.5
Parameters/million	25.77	26.52	34.66	35.43
Training times/h	5.102	5.435	7.314	7.703
p	61.45%	65.47%	71.02%	73.58%
R	58.62%	70.28%	73.59%	77.26%
Real-time performance	82.33	91.26	86.53	96.15

71.02% and 73.59%, respectively, which remain slightly inferior to those of the LTCB YOLOv7 model by 2.56% and 3.67%, respectively.

The LTCB YOLOv7 maturity detection method can effectively be implemented in agricultural production to detect *Hemerocallis citrina Baroni* maturity in farmland. The model has been integrated into a practical detection system encapsulated using PyQt5, achieving functions such as maturity detection, counting, and dynamic adjustment of detection parameters. As depicted in

Figure 12, the system interface includes options for detection type, parameter configurations, and threshold modifications, with immediate visualization of detection results.

However, challenges may arise when deploying the model in real-world environments. Factors such as changing lighting conditions, occlusions caused by overlapping plants, and varying growth states in farmland may affect detection accuracy. Additionally, the model's computational efficiency may require optimization for deployment on edge devices with limited processing power. Future work will focus on enhancing the model's robustness against environmental variability and further integrating lightweight hardware solutions to ensure real-time performance.

In summary, the improved LTCB YOLOv7 algorithm combines a lightweight design with enhanced feature extraction capabilities through an attention mechanism and cross-channel information fusion. These improvements effectively reduce false edge detections and increase the accuracy of maturity detection. With further refinements, this approach can serve as a reliable machine vision solution for precision agriculture, including applications in automated harvesting systems.

a. YOLOv7 maturity detection method   b. LTCB YOLOv7 maturity detection method

Figure 12   System interface and detection

5   Conclusions

This research proposes an improved LTCB YOLOv7 algorithm for identifying the maturity level of *Hemerocallis citrina Baroni*. This study's innovations lie in its lightweight design and efficient layer aggregation strategy, which effectively reduce computational requirements while enhancing detection accuracy and real-time performance. The proposed method demonstrates strong practical value, serving as a reliable machine vision solution for automated harvesting systems, such as *Hemerocallis citrina Baroni* harvesting robots.

Experimental results confirm that the LTCB YOLOv7 algorithm achieves superior detection performance compared to the baseline model, offering significant benefits in accuracy, efficiency, and robustness under varying conditions.

Future research will explore integrating multimodal approaches, such as combining visual and spectral information, to improve the detection accuracy further. Additionally, the adoption of reinforcement learning techniques will be investigated to enhance the model's adaptability in complex and dynamic field environments, further advancing intelligent harvesting systems.

Acknowledgements

[References]

[1] Wu L G, Chen L, Zhou Q, Shi J H, Ma Y B. Maturity detection method for *Hemerocallis citrina baroni* based on lightweight and efficient layer aggregation network. Transactions of the CSAM, 2024; 55(2): 268–277. (in Chinese)

[2] Shi L, Yang C K, Sun X Y, Sun J Y, Dong P, Xiong S F, et al. Detection of fusarium head blight using a YOLOv5sbased method improved by attention mechanism. Int J Agric & Biol Eng, 2024; 17(5): 247–254.

[3] Wu L G, Chen L, Liu Z P, Wu Y Q, Shi J H. YOLOv8-ABW based method for detecting *Hemerocallis citrina Baroni* maturity. Transactions of the CSAE, 2024; 40(13): 262–272.

[4] Wang Q F, Cheng M, Huang S, Cai Z J, Zhang J L, Yuan H B. A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed Solanum rostratum Dunal seedlings. Computers and Electronics in Agriculture, 2022; 199: 107194.

[5] Wu D, Lv S, Jiang M, Song H. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. Computers and Electronics in Agriculture, 2020; 178: 105742.

[6] Hui Y M, Wang J, Li B. DSAA-YOLO: UAV remote sensing small target recognition algorithm for YOLOV7 based on dense residual super-resolution and anchor frame adaptive regression strategy. Journal of King Saud University - Computer and Information Sciences, 2024; Paper ID: 101863. DOI: 10.1016/j.jksuci.2023.101863

[7] Xu W, Xu T, Thomasson J, Chen W, Karthikeyan R, Tian G Z, et al. A lightweight SSV2-YOLO based model for detection of sugarcane aphids in unstructured natural environments. Computers and Electronics in Agriculture, 2023; 211: 107961.

[8] Liu B P, Zhang Y Z, Wang J H, Luo L F, Lu Q H, Wei H L, et al. An improved lightweight network based on deep learning for grape recognition in unstructured environments. Information Processing in Agriculture, 2024; 11(2): 202–216.

[9] Ngugi L, Abdelwahab M, Abo-Zahhad M. A new approach to learning and recognizing leaf diseases from individual lesions using convolutional neural networks. Information Processing in Agriculture, 2023; 10: 11–27.

[10] Ahmed I, Dan S, Jeon G, Piccialli F, Fortino G. Towards collaborative

robotics in top view surveillance: A framework for multiple object tracking by detection using deep learning. IEEE/CAA J. Autom. Sinica, 2021; 8(7): 1253–1270.

[11] Wu W, Yang T L, Li R, Chen C, Liu T, Zhou K, et al. Detection and enumeration of wheat grains based on a deep learning method under various scenarios and scales. Journal of Integrative Agriculture, 2020; 19: 1998–2008.

[12] Zhang L, Wu L G, Liu Y Q. *Hemerocallis citrina Baroni* Maturity detection method integrating lightweight neural network and dual attention mechanism. Electronics, 2022; 11(17): 2743.

[13] Hu J, Shen L, Albanie S, Sun G, Wu E H. Squeeze-and-excitation networks. IEEE Trans Pattern Anal. Mach. Intel., 2020; 42: 2011–2023. DOI: 10.48550 /arXiv.1709.01507

[14] Song C Y, Zhang F, Li J S, Xie J Y, Yang C, Zhou H, et al. Detection of maize tassels for UAV remote sensing image with an improved YOLOX Model. Journal of Integrative Agriculture, 2023; 22: 1671–1683. DOI: 10.1016/ j.jia.2022.09.021

[15] Sun P, Li S Q, Zhu B, Zuo Z Y, Xia X H. Vision-based fixed-time uncooperative aerial target tracking for UAV. IEEE/CAA J. Autom. Sinica, 2023; 10(5): 1322–1324.

[16] Yang X, Shu L, Chen J N, Ferrag M, Wu J, Nurellari E, et al. A survey on smart agriculture: development modes, technologies, and security and privacy challenges. IEEE/CAA J. Autom. Sinica, 2021; 8: 273–302.

[17] Han K, Wang Y H, Tian Q, Guo J Y, Xu C J, Xu C. Ghost Net: More Features from Cheap Operations. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020; pp.1577–1586. DOI: 10.48550/arXiv.1911.11907

[18] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021; pp.13708–13717. DOI: 10.48550/arXiv.2103.02907

[19] Tan M, Pang R M, Le Q. EfficientDet: Scalable and efficient object detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020; pp.1077–10787. DOI: 10.48550/arXiv.1911.09070

[20] Li G J, Huang X J, Ai J Y, Yi Z R, Xie W. Lemon-YOLO: An efficient object detection method for lemons in the natural environment. IET Image Processing, 2021; 15(9): 1998–2009.

[21] Liu H Y, Duan X H, Chen H N, Lou H T, Deng L X. DBF-YOLO: UAV small targets detection based on shallow feature fusion. IEEJ Transactions on Electrical and Electronic Engineering, 2023; 18(4): 605–612.

[22] Zhang R, Wen C B. SOD-YOLO: A small target defect detection algorithm for wind turbine blades based on improved YOLOv5. Advanced Theory and Simulations, 2022; 5(7): 2100631.

[23] Ganesan G, Chinnappan J. Hybridization of ResNet with YOLO classifier for automated paddy leaf disease recognition: An optimized model. Journal of Field Robotics, 2022; 39(7): 1085–1109.

[24] Junos M, Khairuddin A, Thannirmalai S, Dahari M. An optimized YOLO-based object detection model for crop harvesting system. IET Image Processing, 2021; 15(9): 2112–2125.

[25] Wang X Y, Hu Q, Cheng Y S, Ma J Y. Hyperspectral image super-resolution meets deep learning: A survey and perspective. IEEE/CAA J. Autom. Sinica, 2023; 10(8): 1668–1691.

[26] Liu Z Y, Li Y, Shuang F, Huang Z M, Wang R C. EMB-YOLO: Dataset, method and benchmark for electric meter box defect detection. Journal of King Saud University - Computer and Information Sciences, 2024; Paper ID: 101936. DOI: 10.1016/j.jksuci.2024.101936

[27] Cao Z G, Li J B, Fang L, Li Z Q, Yang H X, Dong G H. Research on efficient classification algorithm for coal and gangue based on improved MobilenetV3-small. International Journal of Coal Preparation and Utilization. 2024 May 20: 1–26. DOI: 10.1080/19392699.2024.2353128

[28] Han D X, Hu H R, Yang J, Liang X L, Ai J M, Abula A, et al. The ideal harvest time for seed production in maize (Zea mays L.) varieties of different maturity groups. Journal of The Science of Food and Agriculture, 2022; 102(13): 5867–5874.

[29] Tan J B, Wan J F, Xia D. Automobile component recognition based on deep learning network with coarse-fine-grained feature fusion. International Journal of Intelligent Systems, 2023; Article ID 1903292, 14 pages. DOI: 10.1155/2023/1903292

[30] Zermas D, Nelson H, Stanitsas P, Morellas V, Mulla D, Papanikolopoulos N. A methodology for the detection of nitrogen deficiency in corn fields using high-resolution RGB imagery. IEEE Transactions on Automation Science and Engineering, 2023; 18(4): 1879–1891.

[31] Wu L G, Zhang L, Shi J H, Zhang Y, Wan J F. Damage detection of grotto murals based on lightweight neural network. Computers and Electrical Engineering, 2022; 102: 108237.

[32] Zhang M C, Hao S, Zhang Y, Yu Y, Zhou M S. Deep learning-based damage detection of mining conveyor belt. Measurement, 2021; 175: 109130.

[33] Okafor E, Mojeed O, Motaz A. Deep reinforcement learning with light-weight vision model for sequential robotic object sorting. Journal of King Saud University - Computer and Information Sciences, 2024; Paper ID: 101896. DOI: 10.1016/j.jksuci.2023.101896

[34] Du W S, Jia Z, Sui S S, Liu P. Table grape inflorescence detection and clamping point localization based on channel pruned YOLOV7-TP. Biosystems Engineering, 2023; 235: 100–115.