Online diagnosis platform for tomato seedling diseases in greenhouse production

Xin Jin[1,2], Xiaowu Zhu[1], Jiangtao Ji[1,2*], Mingyong Li[1,2], Xiaolin Xie[1,2], Bo Zhao[3]

(1. *College of Agricultural Equipment Engineering, Henan University of Science and Technology, Luoyang 471003, Henan, China*;
2. *Science & Technology Innovation Center for Completed Set Equipment, Longmen Laboratory, Luoyang 471023, Henan, China*;
3. *Institute of Mechatronics Technology and Application, China Mechanized Agriculture Science Research Institute Group Co., Ltd, Beijing 100020, China*)

Abstract: The facility-based production method is an important stage in the development of modern agriculture, lifting natural light and temperature restrictions and helping to improve agricultural production efficiency. To address the problems of difficulty and low accuracy in detecting pests and diseases in the dense production environment of tomato facilities, an online diagnosis platform for tomato plant diseases based on deep learning and cluster fusion was proposed by collecting images of eight major prevalent pests and diseases during the growing period of tomatoes in a facility-based environment. The diagnostic platform consists of three main parts: pest and disease information detection, clustering and decision-making of detection results, and platform diagnostic display. Firstly, based on the You Only Look Once (YOLO) algorithm, the key information of the disease was extracted by adding attention module (CBAM), multi-scale feature fusion was performed using weighted bi-directional feature pyramid network (BiFPN), and the overall construction was designed to be compressed and lightweight; Secondly, the *k*-means clustering algorithm is used to fuse with the deep learning results to output pest identification decision values to further improve the accuracy of identification applications; Finally, a detection platform was designed and developed using Python, including the front-end, back-end, and database of the system to realize online diagnosis and interaction of tomato plant pests and diseases. The experiment shows that the algorithm detects tomato plant diseases and insect pests with mAP (mean Average Precision) of 92.7%, weights of 12.8 Megabyte (M), inference time of 33.6 ms. Compared with the current mainstream single-stage detection series algorithms, the improved algorithm model has achieved better performance; The accuracy rate of the platform diagnosis output pests and diseases information of 91.2% for images and 95.2% for videos. It is a great significance to tomato pest control research and the development of smart agriculture.

Keywords: pest and disease detection, YOLO, diagnosis platform, *k*-means clustering, facility production base
DOI: 10.25165/j.ijabe.20241701.8433

Citation: Jin X, Zhu X W, Ji J T, Li M Y, Xie X L, Zhao B. Online diagnosis platform for tomato seedling diseases in greenhouse production. Int J Agric & Biol Eng, 2024; 17(1): 80–89.

1 Introduction

Tomatoes are affected by different pests and diseases at different stages of growth, which is the main reason for reduced yield. Accurate identification and diagnosis of pests and diseases in tomato seedlings during the growing period, as well as early detection and treatment, will not only provide a healthy growing environment for tomatoes but will also effectively increase tomato yields to a large extent[1]. In particular, the high relative temperature and humidity, poor lighting conditions, and poor circulation in the facility base provide excellent environmental conditions for the rapid spread of pathogens, greatly increasing the chances of pathogenic infestation and leading to a disaster. The most serious diseases of tomato seedlings during the growing period are Bacterial Spot[2], Early Blight[3], Late Blight[4], and Leaf Mold[5], which have a greater than 50% chance of developing[6,7]. The disease mainly affects the leaves, starting with individual seedlings and then spreading rapidly in all directions with these plants as the center, infecting neighboring plants. The main pest species are Aphid[8], Helicoverpa Armigera[9], Spider Mite[10] and White Fly[11], all of them are highly reproductive, fast-growing and widespread when the environment is suitable, and in addition to direct damage, can also spread directly or promote secondary infection of the disease[12]. According to statistics, approximately 15% of global tomato production is affected by pests and diseases each year, with average yield reductions in severe regions capable of reaching 40% to 80%[13,14]. Careful control of pests and diseases is a key task to reduce losses and increase crop yields. Once a pest or disease has invaded a field, it must be detected in time for farmers to treat it and prevent it from spreading[15]. Therefore, it is necessary to select pests and diseases that cause serious damage to tomatoes as research objects, to collect and collate relevant information, to achieve accurate identification and detection, and to provide a theoretical basis for targeted early warning and prevention.

Traditional detection methods no longer meet the needs of research and production in terms of identification efficiency, accuracy, and application scenarios. With the continuous development of the Internet, the application of information technology has provided new methods and ideas for the

Received date: 2023-07-20 **Accepted date:** 2023-12-06
Biographies: Xin Jin, PhD, Associate Professor, research interest: vegetables intelligent production and agricultural sensors, Email: jx.771@163.com; **Xiaowu Zhu**, MS candidate, research interest: agricultural engineering and information technology, Email: zhuxiaowu620@163.com; **Mingyong Li**, PhD, research interest: agricultural information and intelligent machine, Email: mingyonglee@163.com; **Xiaolin Xie**, PhD, research interest: agricultural intelligent power equipment, Email: xiaolinxie1985@hotmail.com; **Bo Zhao**, Research Fellow, research interest: image processing and agricultural machinery intelligence, Email: zhaobo@caams.org.cn.
***Corresponding author: Jiangtao Ji**, PhD, Professor, research interest: agricultural utility robot technology. Henan University of Science and Technology, Luoyang 471003, Henan, China. Tel: +86-379-64877817, Email: jjt0907@163.com.

identification of crop pests and diseases. The successful application of deep learning in other fields has attracted the attention of many agricultural scholars and applied it to the agricultural field[16,17]. Deep learning (DL) methods, especially those based on convolutional neural networks (CNN), are widely used in object detection and classification in the agricultural field, demonstrating excellent performance and classification in applications such as plant pests detection and plant identification[18].

Current mainstream object detection algorithms can be divided into two types. One is a two-stage object detection based on the candidate region method, which requires a proposal (a pre-selected box that may contain the object to be detected) followed by fine-grained object detection. Such as RCNN (Regions with CNN features), Fast RCNN, and Faster RCNN, etc. Deng et al.[19] proposed a multiple pests detection technology based on federated learning (FL) and improved fast regional convolutional neural network (R-CNN). The improved R-CNN has an average accuracy of 90.27% for multiple pests detection in orchards, and the detection time of each image is only 0.05 s, realizing the accurate identification of small pests and diseases in complex environments. Jiao et al.[20] proposed an anchor-free area proposal network (AFRPN) and combined it with Fast R-CNN to detect 24 types of pests in an end-to-end manner. The mAP and recall of the improved model are 7.5% and 15.3% higher than that of Faster R-CNN, the running time can reach 0.07 seconds per image, which meets real-time detection. Zhang et al.[21] proposed an improved Faster RCNN algorithm to detect tomato diseased leaves, using ResNet101 instead of VGG16 for feature extraction and *k*-means clustering algorithm for clustering bounding boxes, the accuracy rate increased by 2.71%, which can effectively detect and recognize tomato diseases. Xie et al.[22] proposed Faster DR-IACNN algorithm on the self-built grape leaf disease dataset (GLDD). Inception-v1 module, InceptionResNet v2 module and SE module are introduced, mAP is 81.1%, and detection speed reached 15.01 fps. One is a single-stage object detection based on regression methods, which requires only one input to the network to predict all bounding boxes, and extracts feature directly from the network to predict the classification and location of objects. Such as Single Shot MultiBox Detector (SSD), You Only Look Once (YOLO), etc. Wang et al.[23] proposed a Deep Block Attention SSD (DBA_SSD) method for plant leaf disease identification by combining an improved VGG network and a channel attention mechanism, achieving 92.2% accuracy on the plant Village dataset. Sun et al.[24] built a new apple leaf disease detection model based on the Mobile AppleNet SSD algorithm using the MEAN module and the Inception module, which could achieve 83.12% mAP and 12.53 fps in a complex background. Liu and Wang[25] constructed a dataset of tomato pests and diseases in a real natural environment, used image pyramids to optimize the feature layer of the YOLO V3 algorithm, realized multi-scale feature detection, and could accurately, quickly detect the location and type of tomato pests and diseases. Wang et al.[26] achieved early real-time detection of tomato pests and diseases with an F1 value of 94.77% and an AP value of 91.81%, with a false detection rate of only 2.1%, based on YOLOv3 with fused expanded convolution and convolution factor decomposition. Liu et al.[27] proposed a tomato pests identification algorithm based on the improved YOLOv4 fusion triple attention mechanism (YOLOv4-TAM) by introducing the focal loss function and the *k*-means + clustering algorithm, with an average identification accuracy of 95.2%. Qi et al.[28] added the Squeeze and Excite (SE) module to the YOLOv5 model for the detection of tomato virus diseases in natural backgrounds, with an accuracy rate of 91.07%. Chen et al.[29] integrated the involute bottleneck module and SE module on the basis of the original YOLOv5 network algorithm. The detection accuracy of the algorithm for powdery mildew and anthracnose is 86.5% and 86.8%, respectively.

As mentioned above, deep learning-based object detection algorithms can better extract features from images and show good performance in the identification of agricultural objects, but relatively little research has been done on the identification of tomato pests and diseases in the growth period. This study created a tomato growing period pests and disease dataset consisting of 2388 images by collecting images of tomato pests and diseases in a facility environment, using the plant public dataset and internet crawler technology for additions (The dataset has been shared at https://drive.google.com/file/d/1V1cRPVwtqrJuBJifGyQj0z2b0N0Dh_OS/view?usp=sharing). It can provide data to support research on tomato pests and diseases during the growing period.

To improve the efficiency and accuracy of identifying tomato pests and diseases, the YOLOv5s-CBC algorithm was proposed, which achieves high accuracy and lightweight by adding the attention module, weighted bi-directional feature pyramid network module, and GhostNet module. Designed and developed an online diagnosis platform for pests and diseases, using *k*-means clustering algorithm to analyze and make decisions on the detection results, and output the optimal identification results and confidence to the platform interface to achieve real-time detection of pictures or videos. The online platform was applied in the tomato facility production base, and the test results met the need for rapid and accurate detection of diseases and pests during the growth period of tomato seedlings. It is of great significance to tomato pest control research and the development of smart agriculture.

2 Construction of pest and disease datasets

2.1 Data acquisition

Eight tomato pests were selected for this study: Aphid (AP), Bacterial Spot (BS), Early Blight (EB), Helicoverpa Armigera (HA), Late Blight (LB), Leaf Mold (LM), Spider Mite (SM), and White Fly (WF). The data was collected from the facility production base in the Luolong District, Luoyang City, Henan province, China, 112°28'27.09"E, 34°38'17.19"N, the scene is shown in Figure 1. There is sufficient light in the facility, and a DSC-H300 camera (Sony Corporation, Japan) was used to capture the images. A total of 1978 images were obtained, with a pixel size of 640×480 pixels. In order to ensure the diversity of data, 231 images were collected through the plant public dataset[30] and 179 images were collected using web crawler technology[31], with pixels of 256×256 pixels.

a. Exterior b. Interior

Figure 1 Facility production base

These images contain different light intensities and object sizes. After screening and sorting, a total of 2388 effective multi-scenario and multi-scale images of pests and diseases were obtained. Figure 2 are examples of some pest and disease images in the dataset.

2.2 Data expansion and enhancement

To avoid overfitting problems due to the small number of datasets, the pests in the dataset were expanded using CycleGAN (Cycle-Consistent Generative Adversarial Networks)[32], using contrast enhancement, rotation and flipping to expand the pests.

CycleGAN is trained iteratively against each other to achieve the effect of expanding the dataset. The coordinates and categories of the dataset were manually annotated using the lightweight image annotation tool LabelImg (CSAIL, USA) to obtain the corresponding label files for the dataset images, and the annotated dataset was divided into a training set and a test set. The distribution of the training and test sets and the labels for each category are listed in Table 1.

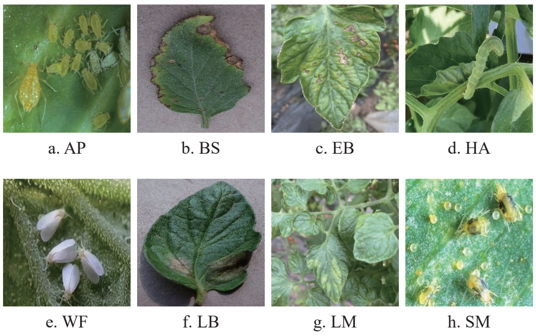


Figure 2 Example pests and disease dataset

The training data is enhanced by the use of a data loader during the data input process. The data loader uses a variety of data enhancement methods, including color transformation, random rotation and panning, Mosaic[33], Mixup[34], and a random combination of these methods stitched together to enrich the object and background. The enhancement process is shown in Figure 3.

Table 1 Distribution of the number of pests and disease images

Category	Train/enhance	Test/enhance	Sum	Label
AP	259/777	45/135	1216	0
BS	249/747	45/135	1176	1
EB	253/759	45/135	1192	2
HA	256/768	45/135	1204	3
LB	250/750	45/135	1180	4
LM	250/750	45/135	1180	5
SM	251/753	45/135	1184	6
WF	260/780	45/135	1220	7
Sum	8112	1440	9552	--

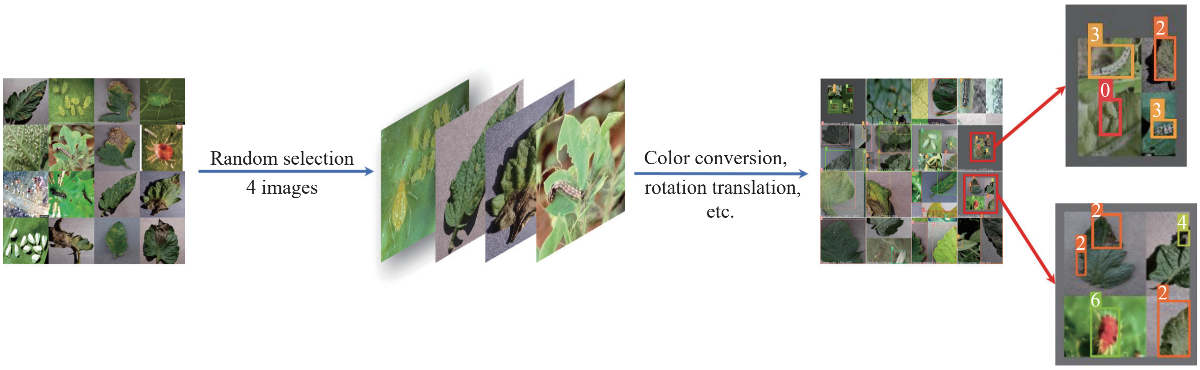


Figure 3 Data enhancement process

3 Detection platform design

In order to achieve a facility-based pests detection application, an algorithm based on the YOLOv5 was proposed for the detection of eight pests and diseases during the growing period of tomatoes. By clustering the identified information using the *k*-mean clustering algorithm, the frequency of pests and disease occurrence and the confidence level are statistically analyzed to achieve rapid data clustering and fusion and online pests and disease diagnosis and decision-making. Design and develop a pests and disease detection webpage to introduce information on various types of pests and disease and prevention methods, and output of the algorithm identification results on the platform after analysis by clustering.

Figure 4 shows the architecture of the online pests and disease diagnosis process. After the algorithm is trained and saved, it is connected to the online diagnosis platform, where users upload images or videos of tomato pests and diseases, and the platform will identify and output prediction results.

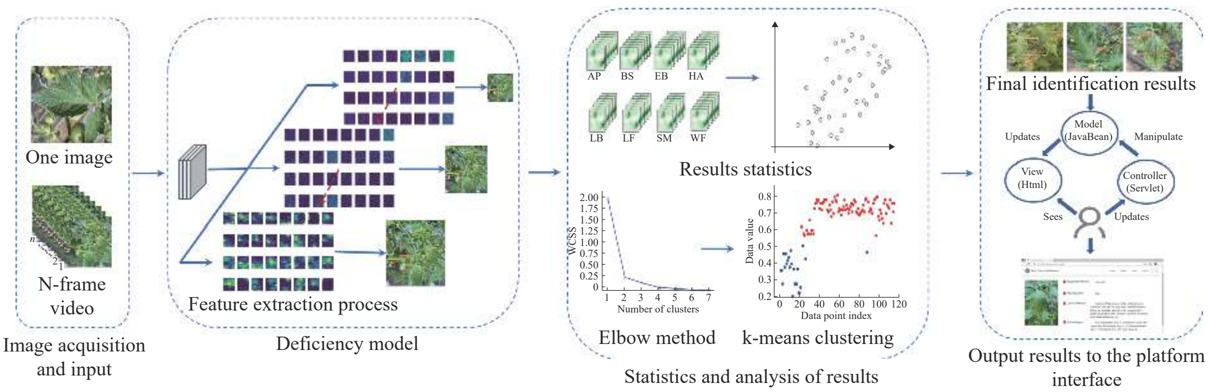


Figure 4 Online diagnostic process architecture

3.1 Identification algorithm model

The pest and disease diagnostic platform needs to meet the real-time needs of users. Accurate and real-time detection is a prerequisite for easy use of the webpage, and YOLOv5 was chosen as the detection model, taking into account factors such as real-time pest and disease detection, identification accuracy, and future deployment. YOLOv5 is a more mature single-stage object detection model in the You Only Look Once (YOLO)[35] family, capable of locating and classifying objects by directly regressing the relative positions of candidate frames. There are three main components: the backbone network (Backbone), the neck network (Neck), and the detection network (Head). The backbone network consists of convolution module (Conv), C3 module, and a modified spatial pyramid module (SPPF) for extracting features from the input image and passing them to the Neck. Neck uses the structure of FPN (Feature Pyramid Network) and PAN (Path Aggregation Network), where FPN uses a high resolution of low-level features and semantic information of high-level features in top-down delivery of semantic information, and PAN is bottom-up delivery of localization information, making it easier to propagate low-level information to the top level and enhancing the ability of network feature fusion. Head is responsible for classifying the features extracted after compression and fusion, generating detection frames, and classifying them into the appropriate coordinates, categories, and confidence levels.

A tomato pest detection algorithm is proposed based on the YOLOv5s algorithm. In Bankbone, the CBAM attention mechanism has been introduced to help the network locate the location of pests and diseases more accurately, thus reducing the interference of object background information. In Neck, the use of the BiFPN structure instead of the original PANet structure improves the detection of small object pests by increasing the weights to adjust the contribution at each scale of the algorithm. The C3Ghost module is used instead of the C3 module to reduce the number of parameters, achieve a lighter model, and balance identification accuracy with detection speed. Figure 5 shows the structure of the improved algorithm.

Improvement 1: The attention mechanism can focus on important information with high weights and ignore irrelevant information with low weights, and can continuously adjust the weights so that important information is selected in different situations[36]. There are problems of multiple objects, small size and complex backgrounds in the homemade pest and disease dataset. A CBAM module[37] was added to the original feature extraction network part, which consists of a channel attention module and a spatial attention module. It enables the network to better extract feature information of pests and diseases and improve the characterization capability.

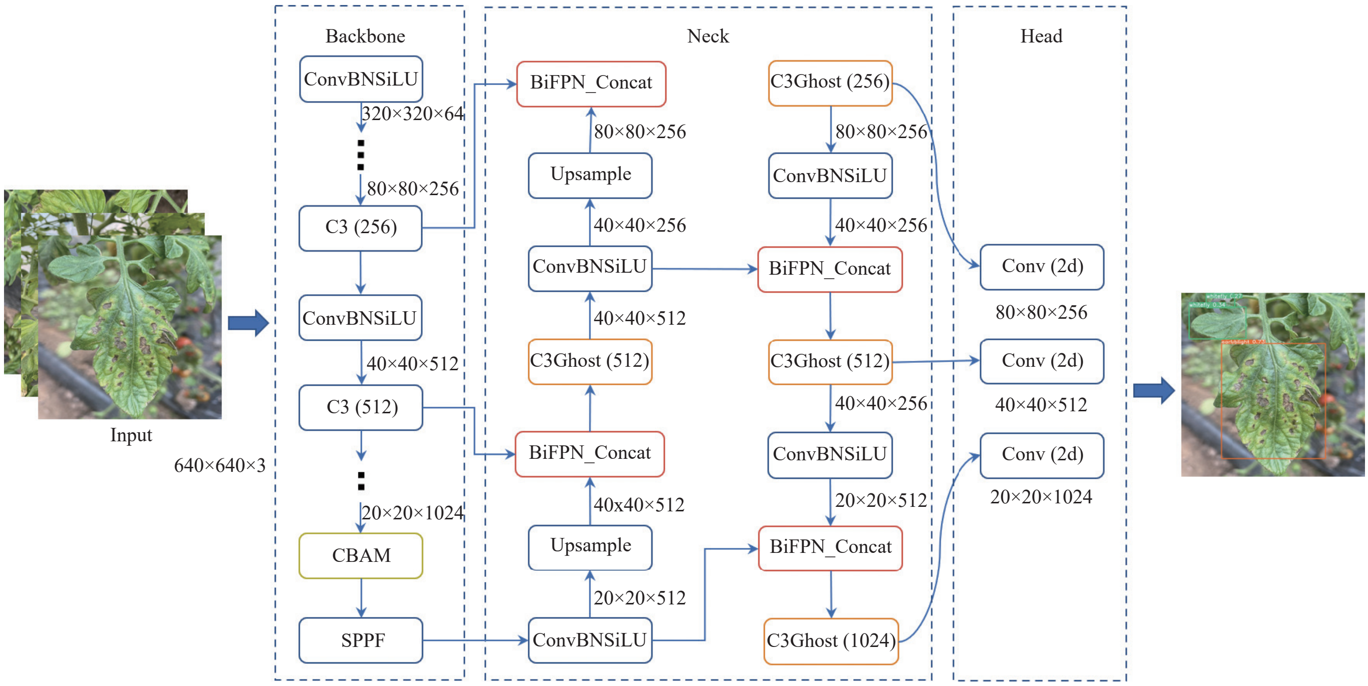


Figure 5 Structure of pests detection algorithm

Improvement 2: The neck of the original YOLOv5s simply adds up different scale pest features in the feature fusion process, the method does not make full use of feature information between different scales, and the importance of feature information varies at different scales, which can lead to lower detection accuracy of the network. To solve this problem, the BiFPN network structure[38], a combination of Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) structures, was introduced for enhancing the depth of information mining and further improving the feature extraction capability of the algorithm.

Improvement 3: During the extraction of pest and disease features, the standard convolution extracts feature maps that are very similar on many feature channels, with some duplication. Therefore, the desired feature maps can be obtained without the need for a full convolution operation. To reduce the number of operations, the Ghost module[39] is used, and the C3Ghost module is used to replace the C3 module in the network in a two-step operation: first, the ordinary convolution calculation is used to obtain a feature map with a smaller number of channels; then, more feature maps are obtained using linear operations, and the output results are stitched with the constant mapping in the channel direction to obtain the output feature map.

3.2 Cluster analysis model

As an unsupervised clustering algorithm in data mining, k-means is simple and effective[40]. This algorithm is a distance-based clustering algorithm, which uses the distance between samples as the sample similarity measure, and its basic process is as follows:

1) Randomly select k value from the known model data as the initial cluster center, and the j^{th} mean value is recorded as u_j;

2) Classify the categories by calculating the distance d from

each cluster object to the cluster center;

$$d = \sqrt{\sum_{i=1}^{n}(x_i - u_j)^2} \tag{1}$$

where, $j = 1,2\ldots,k, i = 1,2\ldots,n, d$ is the distance from the i^{th} point to the j^{th} mean, and x_i is the i^{th} data value;

3) Calculate each cluster center again;

$$u_k = \frac{\sum_{i=1}^{n_k} X_i}{n_k} \tag{2}$$

where, n_k denotes the number of objects in the k^{th} cluster, u_k denotes the number of cluster centers;

4) Repeat steps 2) and 3) until the clustering centers no longer change or the maximum number of iterations is reached, then stop.

To perform unsupervised *k*-means clustering on the given data in this study, it is first necessary to determine the optimal number of clusters, using the Elbow Method, by plotting the intra-cluster sum of squares (WCSS) against the number of clusters and selecting the number of clusters where WCSS decreases at a slower rate and an inflection point occurs.

3.3 Interface design module

The platform adopts B/S (Browser/Server) system structure and is divided into three parts: front-end, back-end and database. The development mainly uses Python language and the front end uses html combined with js, bootstrap and css to display the page and interact with the back end. The back end uses the SSM (Spring+SpringMVC+MyBatis) framework, the SSM framework is Spring MVC, Spring and Mybatis framework integration, using Spring MVC is responsible for request forwarding and view management, Spring to achieve business object management, Mybatis as the data object persistence engine. MVC is composed of Model (JavaBean), View (Html), and Controller (Servlet). Model is a module used to carry data and perform calculations on requests submitted by users. View provides a interface for the user and interacts directly with the user. The Controller is used to forward the user request to the corresponding Model for processing and provide the corresponding response to the user with the calculation result of the processing Model. Use the server session to cache the currently logged-in user, upload the file by copying the file to the tomcat path, and then use the URL to access it, and the database uses mysql to store the database.

4 Results and analysis

4.1 Experimental environment

To ensure the reliability of the experimental data, all algorithms used in this study were tested under the same experimental conditions. The hardware and software details are listed in Table 2.

Table 2 Hardware and software environment

Item	Type
CUDA	11.7
CPU	Intel(R) Core(TM) i5-10100F CPU@2.9GHz
GPU	NVIDIA GeForce GT 1030
Operating system	Windows 10
Deep learning frame	Pytorch
Python	3.9

The model was trained using stochastic gradient descent (SGD) for optimization, with the learning scheduler set to Poly, the momentum parameter set to 0.937, the weight decay set to 0.0005, the learning rate set to 0.001, the epochs set to 270, and the batch size set to 16. The aspect ratio of the image subject was set to 1:1, and for consistency, part of the image was cropped from the original image was cropped out of the original image for consistency.

4.2 Algorithm performance analysis

In order to objectively assess the model performance, precision (Pr), recall (Re), mean average precision (Intersection over Union = 0.5), operational parameters, and algorithm weights are used as evaluation metrics for comparison experiments.

$$\mathrm{Pr} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \times 100\% \tag{3}$$

$$\mathrm{Re} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100\% \tag{4}$$

$$\mathrm{AP} = \int_0^1 \mathrm{PrdRe} \tag{5}$$

$$\mathrm{mAP} = \frac{1}{C}\sum_{i=1}^{C} \mathrm{AP}i \tag{6}$$

where, TP is the number of true positive samples; FP is the number of false positive samples; FN is the number of false negative samples; C is the type of pest and disease, and C = 8 in this study. mAP represents the average value of average precision (AP) under each detection type, that is, the area of the curve formed by the precision and recall of each pest category. mAP@0.5 means the average AP of all types when IoU (Intersection over Union) is set to 0.5.

4.2.1 Ablation experiments

Ablation research observes the impact on performance by removing or adding some features of the algorithm[41]. To verify the effectiveness of the introduced module in improving the performance of the algorithm, four algorithms were built to perform ablation experiments on the dataset of this study, all in the same training environment.

Among them, YOLOv5s-CB indicates the introduction of the CBAM module into the Backbone tail. YOLOv5s-BF indicates the introduction of the Bi-directional Feature Pyramid Network (BiFPN) structure into the Neck. YOLOv5s-CG indicates the replacement of the C3 module in the Neck with the C3Ghost module. YOLOv5s-CBC is the algorithm proposed in this study. It can be seen from Table 3 that compared with YOLOv5s, the precision rate decreased by 0.1% after the fusion of the CBAM module, while the recall rate and mAP increased by 1.1% and 0.8% respectively, indicating that after adding the attention mechanism, the model is more focused on the object of pests and diseases area, the detection ability of the model is increased, but the algorithm weight is increased by 0.4M. After integrating the BiFPN module, the recall rate decreased by 1.6%, and the precision rate and mAP increased by 1.2% and 1.2%, indicating that the feature fusion ability of the network is improved after adding BiFPN, and the feature consistency of pests and diseases at different scales is ensured. After integrating the C3Ghost module, the recall rate is reduced by 0.8%, the precision rate and mAP are increased by 0.3% and 1.5%, and the weight is reduced by 1.9M. It shows that the weight and operation parameters of the algorithm are reduced after adding C3Ghost, but the detection accuracy of the algorithm is not affected. Combining the above three not only improves the performance of the algorithm, but also reduces the weight of the algorithm to a certain extent, balancing the increase in the weight of the network after the attention mechanism is added, and the final trained algorithm can ensure high accuracy while maintaining a fast detection speed.

Table 3 Comparison of ablation experiment results

Algorithm	CBAM[1]	BiFPN[2]	C3Ghost[3]	Pr[4]/%	Re[5]/%	mAP[6]@ 0.5/%	Weights/ M	Parameter
YOLOv5s	×	×	×	90.3	87.5	89.6	14.3	7 031 701
YOLOv5s-CB	√	×	×	90.2	88.6	90.4	14.7	7 228 407
YOLOv5s-BF	×	√	×	91.5	85.9	90.8	14.3	7 031 701
YOLOv5s-CG	×	×	√	90.6	86.7	91.1	**12.4**	**6 069 413**
YOLOv5s-CBC	√	√	√	**90.6**	**88.9**	**92.7**	12.8	6 266 128

Note: 1. Convolutional Block Attention Module; 2. Bi-directional Feature Pyramid Network; 3. C3+ Ghost module; 4. Precision; 5. Recall; 6. mean Average Precision (Intersection over Union=0.5); The bold part denotes that the algorithm achieves the best results under this metric compared to other algorithms.

Figure 6 shows the change curve of mAP with epoch. The mAP curve fluctuates greatly at the beginning of training from 0 to 100 times, indicating that the convergence speed of the model training is fast in the early stage, which meets the requirements of model training. After 100 times, it is stable and the change is small, indicating that the model is well-trained and there is no overfitting. After 240 times, the curve basically tends to be stable, indicating that the training of the pest detection model is basically completed at this time. The best results of the improved model mAP were obtained when the curves of the five models became stable.

4.2.2 Comparison of different detection models

In order to verify the performance of the improved algorithm, the proposed algorithm was tested against single-stage object detection algorithms such as YOLOv3 tiny, YOLOv4-tiny, YOLOv5n and YOLOv7. The specific results are listed in Table 4. Compared to other algorithms, the improved algorithm has 0.1%-2.0% higher accuracy, 1.0%-6.1% higher recall, 1.9%-3.8% higher mAP, and the lowest Parameters and Weights. mAP is 3.1% higher compared to YOLOv5s, and Parameters and algorithm Weights are 10.9% and 1.5M lower, respectively. The improved algorithm has a significant advantage over other models in terms of detection accuracy and speed while reducing the algorithm parameters.

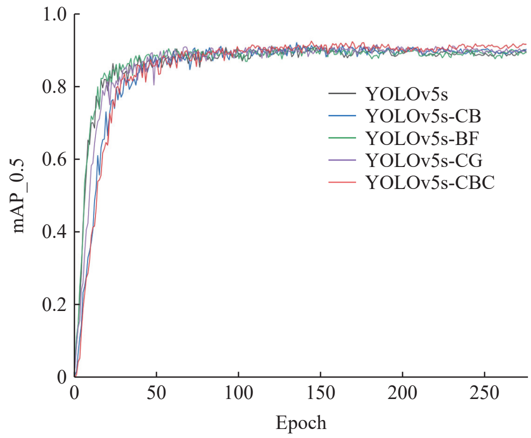


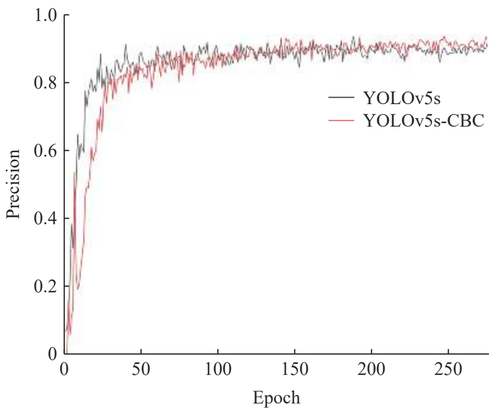
Figure 6 mAP-epoch

Table 4 Comprehensive performance of the network under different models

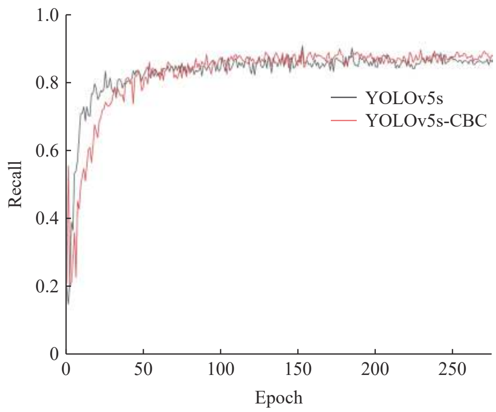
Algorithm	Pr[1]/%	Re[2]/%	mAP@ 0.5[3]/%	Parammeters	Inference time/ms	Weights/M
YOLOv3-tiny	89.4	83.9	87.5	8 682 862	38.6	17.4
YOLOv4-tiny	89.6	84.8	89.2	9 129 509	66	18.6
YOLOv5s	90.3	87.5	89.6	7 031 701	43.4	14.3
YOLOv5n	90.5	87.9	90.8	20 881 221	83.5	42.1
YOLOv7	88.6	82.8	88.9	37 236 527	135.1	72.8
YOLOv5s-CBC	**90.6**	**88.9**	**92.7**	**6 266 128**	**33.6**	**12.8**

Note: 1. Precision; 2. Recall; 3. mean Average Precision (Intersection over Union = 0.5); The bold part denotes that the algorithm achieves the best results under this metric compared to other algorithms.

Figure 7 is a comparison curve of precision and recall. It can be seen from the figure that both show a rapid growth trend within 100 epochs. After 240 epochs of training, the network has stabilized on each evaluation index. The YOLOv5s algorithm is stable at 90.3% and 87.5%, and the YOLOv5s-CBC is stable at 90.6% and 88.9%.



a. Precision comparison



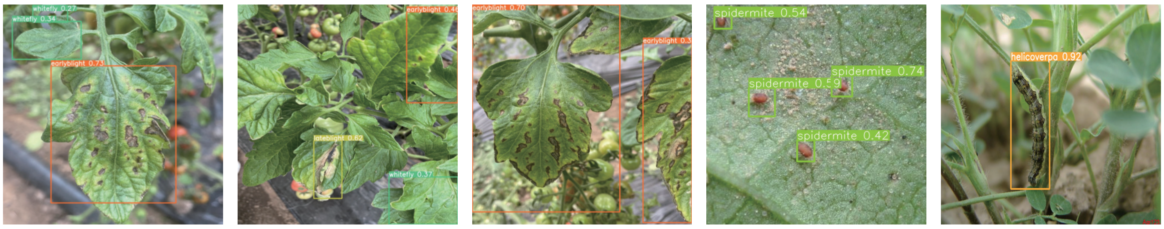
b. Recall comparison

Figure 7 Comparison of precision and recall with epoch

4.2.3 Visual comparison of results

To verify the practical effectiveness of this improved algorithm, an application test was carried out in a tomato facility production base and the results are shown in Figure 8. Some of the white flies in the first column are missed by the original algorithm and are successfully identified by the improved algorithm. In the second column, the original algorithm incorrectly identifies late blight as helicoverpa and lacks identification of late blight, while the improved algorithm identifies the disease but incorrectly identifies late blight as early blight. In the third column, the improved algorithm accurately identifies the right half of the early blight that the original algorithm failed to identify. In the fourth column, the original algorithm misses spider mite and misidentifies the background, but the improved algorithm identifies it accurately. The original algorithm in the fifth column misdetects the background as a helicoverpa, whereas the improved algorithm does not make the error. The results show that the improved algorithm is able to identify pests and diseases missed by the original algorithm, as well as increase the detection of small objects and complex background objects.

a. YOLOv5s



b. YOLOv5s-CBC

Note: Red arrows denote missed pests and diseases;

Figure 8 Visualization comparison

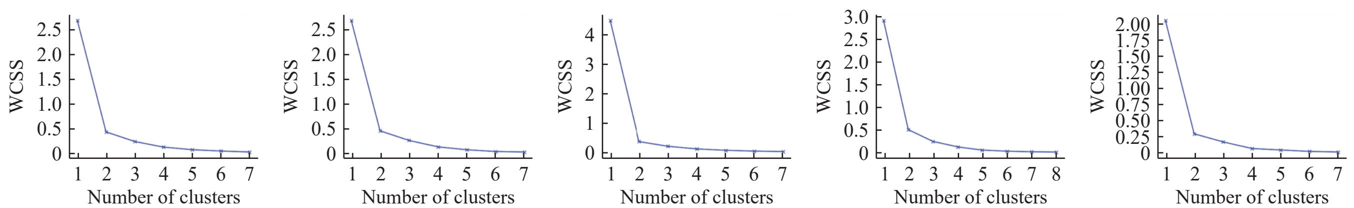
4.3 Analysis of clustering results

Five groups of videos of 10s duration were selected for detection, 1 s is 12 frames, a total of 120 images can be detected, the detection results are listed in Table 5.

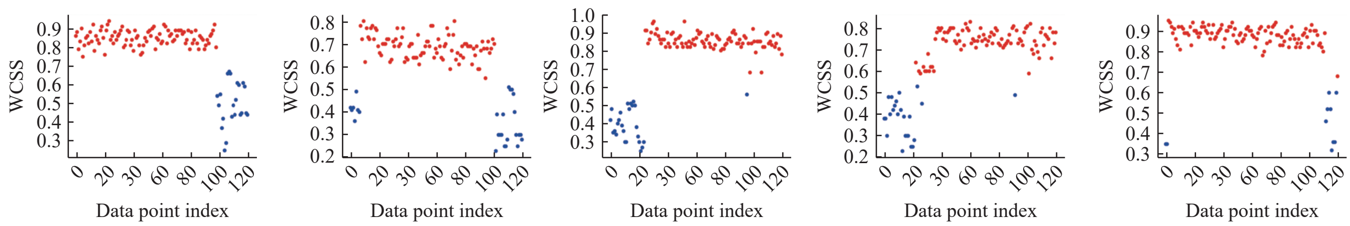
Table 5 Detection result

Number/ Category	AP	BS	EB	HA	LB	LM	SM	WF	Total
1	98	0	0	3	2	2	3	12	120
2	0	7	92	0	9	5	1	4	118
3	10	2	4	2	0	0	6	95	119
4	3	9	9	0	13	81	0	0	115
5	2	0	0	109	4	0	3	2	120

Before clustering, the data needs to be preprocessed, and the value under each label is regarded as a dimension of the vector. Eight kinds of diseases and insect pests correspond to labels 1-8 respectively, and multiple one-dimensional arrays are connected together, and they are converted to a single-column two-dimensional array for use in calculations. The optimal number of clusters are first determined using the Elbow Method based on the data obtained from the algorithm identification, plotting the intra-cluster sum of squares (WCSS) versus the number of clusters. The WCSS is the sum of the squares of the distances between each point and the center of mass of its assigned cluster, and as the number of clusters increases, the WCSS usually decreases, followed by a 'twist'. The WCSS-Number of clusters obtained by the Elbow Method is shown in Figure 9a, with the five data sets clustering best at k=2. Figure 9b shows the data rendering after using k-means clustering. The results obtained by clustering the five groups of data are AP (0.94), EB (0.80), WF (0.86), LM (0.83), and HA (0.95), respectively corresponding to the highest red scatter points in the figure, indicating that k-means clustering can accurately output the results predicted by the algorithm. The results of multiple experiments prove that k-means clustering can successfully output the most frequently occurring and high-confidence pest species, which is in line with the expected effect.



a. WCSS-Number of clusters

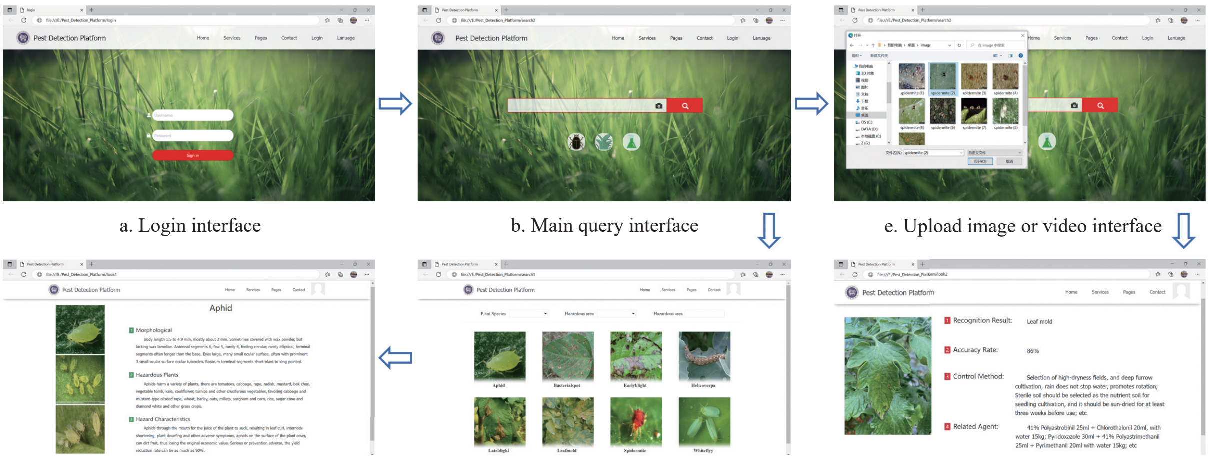


b. Clustering display

Figure 9 Clustering implementation

4.4 Platform realization

Based on the data detected by the algorithm and the results of the cluster analysis, a detection webpage was designed and developed, with some additional features added to make this webpage even better. Figure 10 shows the detection web interface. Figure 10a is the login interface, where the user enters an account and password before entering. Figure 10b is the main query interface, where the user can search for various types of tomato pests and diseases by directly entering the name of the pest or by clicking on the three images below. Figure 10c is the rough query interface, where you can search for pests and diseases based on their characteristics or by clicking on the images directly. Figure 10d is the details interface, which includes the form of the pest, the plant affected, the characteristics of the pest and the prevention and treatment methods. Figure 10e is the detection interface, where images or videos of the pests to be detected are uploaded by clicking on the camera button. Figure 10f is a display of the detection results, showing the optimal identification results and confidence levels calculated by the *k*-means clustering algorithm, and the corresponding control method and related agent. Finally, the tomato pest and disease detection platform was integrated.



a. Login interface b. Main query interface e. Upload image or video interface

d. Detailed information interface c. Rough query interface f. Display of detection results interface

Figure 10 Pest and disease detection platform

5 Discussion

To verify the feasibility of the platform, 125 photographs were taken at the tomato facility production site using a Sony DSC-H300 camera, as well as 10 s of video of different angles of each image at 640×480 pixels, uploaded and tested, and the platform showed accuracy as listed in Table 6.

Table 6 Platform experiment results

Detection type	YOLOv5s	YOLOv5s-CBC
Image/Video	109/112	114/119
Accuracy rate /%	87.2/89.6	91.2/95.2

It was found that when performing image detection, the original algorithm correctly detected 109 images with an accuracy rate of 87.2% and the improved algorithm correctly detected 114 images with an accuracy rate of 91.2%, an improvement of 4 percentage points. When performing video detection, the original algorithm correctly detected 112 images with an accuracy of 89.6% after using the *k*-means clustering algorithm to analyze the results output, and the improved algorithm correctly identified 119 images with an accuracy of 95.2%, an improvement of 5.6 percentage points. It is proved that the improved algorithm YOLOv5s-CBC combined with the *k*-means clustering algorithm in this study can achieve good results and improve the detection ability of pests and diseases.

Figure 11 shows the comparison of the experiment results, after changing the separate image detection to video detection, both can improve the identification accuracy, indicating that video detection can effectively avoid the false detection and leakage problems caused by the interference of light, background, and shooting angle.

According to the characteristics of various types of pests and diseases, the YOLOv5s-CBC algorithm is proposed. Using the CBAM module to suppress irrelevant features, the BiFPN module integrates multi-scale information, and the Ghost module streamlines the network. The improved method enhances the detection capability of the algorithm and reduces the algorithm parameters and weights, which reduces the hardware device requirements and detection time of the identification procedure. Compared with other deep learning object detection algorithms, the improved algorithm can identify the type and location of pests on the image in a timely and effective manner, with small weights and inference time while maintaining accuracy, to satisfy the task requirements for the work of identifying diseases during the growing period of tomato in real life. This method can replace traditional manual identification, and through the designed online pest detection web page, the *k*-means clustering algorithm is used to analyze the identification results. Users can query pest information online and upload images or videos to detect unknown pests and diseases, which greatly improves identification efficiency.

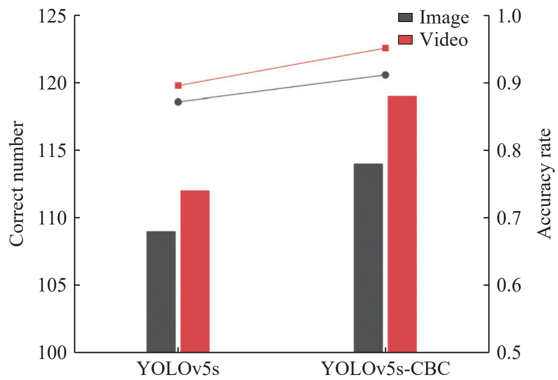


Figure 11 Comparative analysis of results

Regarding misjudgments and missed detections in the identification process, after comprehensive analysis, the misjudgments and missed detections in the identification process are related to the scattered image collection, overly complex background environment, and limited computer performance. Compared to other studies that used data from the same study area, the data from this experiment were selected from 3 sections, with differences in the data, but the results are more adaptable. Due to the limitations of clustering algorithms, only a single pest or disease can be output during video detection. The next step of work can be to study the following two aspects: firstly, in terms of tomato pest and disease datasets, although the homemade dataset used in this study includes pests and backgrounds of different complexity, it still lacks images of different species and scenes. Future datasets with more extensive and diverse multi-species backgrounds should be used, and the functionality of this diagnostic platform should be continuously enriched. Secondly, the recognition algorithms are constantly validated and improved, incorporating the latest benchmark networks and modules of the day to develop a more robust algorithm, making improvements to the clustering algorithms, such as adding weights to the detection results, clustering and outputting multiple pest and disease detections at the same time, and improving the output accuracy.

6 Conclusions

By collecting images of eight major prevalent pests and diseases during the tomato growth period in the facility environment, a data set was created to solve the application problem of object detection in the detection of tomato pests and diseases. To realize online diagnosis and interaction of plant diseases during the growing period of tomatoes, and to successfully carry out the application of an online diagnosis platform for tomato pests and diseases in tomato facility production bases. The following conclusions are drawn from the experiments described in this study:

1) Based on the fusion of the improved YOLOv5 algorithm and the *k*-means clustering algorithm, the online pest and disease diagnosis platform is able to detect the disease information and cluster and make decisions on the results, which are finally displayed on the interactive interface. At the same time, a database of eight pest and disease characteristics during the growing period of tomatoes was established (The dataset has been shared at https://drive.google.com/file/d/1V1cRPVwtqrJuBJifGyQj0z2b0N0Dh_OS/view?usp=sharing), which can provide data support for the detection of pests and diseases during the growing period of tomato.

2) The ablation experiments demonstrate that the CBAM module improves multi-scale feature extraction, the BiFPN module improves the detection of small-scale objects, and the Ghost module effectively reduces the algorithm weights and parameters. The mAP, precision, and recall of the improved algorithm are 92.7%, 90.6%, and 88.9%, respectively. Compared with the current mainstream single-stage detection series algorithms, the mAP is improved by 1.9%-5.2%, and the weights are reduced by 1.5 M-60 M.

3) The improved algorithm is fused with the *k*-means clustering algorithm to achieve real-time monitoring of pests and diseases, and the accuracy of the image and video diagnostic outputs of the online diagnostic platform is 91.2% and 95.2%, which is 4% and 5.6% higher than that of the original algorithm, respectively. It can be used to provide a reference for the prediction and control of pests and diseases.

Acknowledgements

This study was supported by the National Key Research and Development Program of China Project (Grant No. 2021YFD2000700), the Foundation for University Youth Key Teacher of Henan Province (Grant No. 2019GGJS075), and the Natural Science Foundation of Henan Province (Grant No. 202300410124).

[References]

[1] Jia Z H, Zhang Y Y, Wang H T, Liang D. Identification method of tomato disease period based on Res2Net and bilinear attention mechanism. Transactions of the CSAM, 2022; 53(7): 259–266.

[2] Osdaghi E, Jones J B, Sharma A, Goss E M, Abrahamian P, Newberry EA, et al. A centenary for bacterial spot of tomato and pepper. Molecular Plant Pathology, 2021; 22(12): 1500–1519.

[3] Hu Z, Zhang Y. Effect of dimensionality reduction and noise reduction on hyperspectral recognition during incubation period of tomato early blight. Spectroscopy and Spectral Analysis, 2023; 43(3): 744–752.

[4] Wu K T, Gevens A J, Silva E M. Exploring grower strategies and needs for enhancing organic disease management of tomato late blight. Renewable Agriculture and Food Systems, 2022; 37(5): 382–398.

[5] Zhao T T, Pei T, Jiang J B, Yang H H, Zhang H, Li J F, et al. Understanding the mechanisms of resistance to tomato leaf mold: A review. Horticultural Plant Journal, 2022; 8: 667–675.

[6] Hu W Y, Hong W, Wang H K, Liu M Z, Liu S. A study on tomato disease and pest detection method. Applied Sciences, 2023; 13(18): 10063.

[7] Mansour A, Al-Banna L, Salem N, Alsmairat N. Disease management of organic tomato under greenhouse conditions in the Jordan Valley. Crop Protection, 2014; 60: 48–55.

[8] Ramírez C C, Gundel P E, Karley A J, Leybourne D J. To tolerate drought or resist aphids? A new challenge to plant science is on the horizon. Journal of Experimental Botany, 2023; 74(6): 1745–1750.

[9] Reddy G, Tangtrakulwanich K. Module of integrated insect pest management on tomato with growers' participation. Journal of Agricultural Science, 2014; 6(5): 1619–9752.

[10] de Souza Marinke L, de Resende J T V, Hata F T, Dias D M, de Oliveira L V B, Ventura M U. Selection of tomato genotypes with high resistance to Tetranychus evansi mediated by glandular trichomes. Phytoparasitica, 2022; 50(3): 629–643.

[11] Saleem M, ul Hasan M, Sagheer M, Atiq M. Determination of insecticide resistance in *Bemisia tabaci* (Hemiptera: Aleyrodidae) populations from Punjab, Pakistan. International Journal of Tropical Insect Science, 2021; 41: 1799–1808.

[12] Mateos Fernández R, Petek M, Gerasymenko I, Juteršek M, Baebler Š, Kallam K, et al. Insect pest management in the age of synthetic biology. Plant Biotechnology Journal, 2022; 20(1): 25–36.

[13] Arunnehru J, Vidhyasagar B S, Anwar Basha H. Plant leaf diseases recognition using convolutional neural network and transfer learning. In: International Conference on Communication, Computing and Electronics Systems, Springer, 2020; pp.221–229.

[14] Singh R, Ao N T, Kangjam V, Rajesha G, Banik S. Plant growth promoting microbial consortia against late blight disease of tomato under natural epiphytotic conditions. Indian Phytopathology, 2022; 75(2): 527–539.

[15] Li Y F, Wang H X, Dang L M, Sadeghi-Niaraki A, Moon H. Crop pest recognition in natural scenes using convolutional neural networks. Computers and Electronics in Agriculture, 2020; 169: 105174.

[16] Liu J, Wang X W. Plant diseases and pests detection based on deep learning: a review. Plant Methods, 2021; 17: 22.

[17] Kamilaris A, Prenafeta-Boldú F X. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 2018; 147: 70–90.

[18] Lu J Z, Tan L J, Jiang H Y. Review on convolutional neural network (CNN) applied to plant leaf disease classification. Agriculture, 2021; 11(8): 707.

[19] Deng F M, Mao W, Zeng Z Q, Zeng H, Wei B Q. Multiple diseases and pests detection based on federated learning and improved faster R-CNN. IEEE Transactions on Instrumentation and Measurement, 2022; 71: 1–11.

[20] Jiao L, Dong S F, Zhang S Y, Xie C J, Wang H Q. AF-RCNN: An anchor-free convolutional neural network for multi-categories agricultural pest detection. Computers and Electronics in Agriculture, 2020; 174: 105522.

[21] Zhang Y, Song C L, Zhang D W. Deep learning-based object detection improvement for tomato disease. IEEE Access, 2020; 8: 56607–56614.

[22] Xie X Y, Ma Y, Liu B, He J R, Li S Q, Wang H Y. A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks. Frontiers in Plant Science, 2020; 11: 751.

[23] Wang J, Yu L Y, Yang J, Dong H. DBA_SSD: A novel end-to-end object detection algorithm applied to plant disease detection. Information, 2021; 12(11): 474.

[24] Sun H N, Xu H W, Liu B, He D J, He J R, Zhang H X, et al. MEAN-SSD: A novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks. Computers and Electronics in Agriculture, 2021; 189: 106379.

[25] Liu J, Wang X W. Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network. Frontiers in Plant Science, 2020; 11: 898.

[26] Wang X W, Liu J, Zhu X N. Early real-time detection algorithm of tomato diseases and pests in the natural environment. Plant Methods, 2021; 17: 43.

[27] Liu J, Wang X W, Miao W Q, Liu G X. Tomato pests recognition algorithm based on improved YOLOv4. Frontiers in Plant Science, 2022; 13: 814681.

[28] Qi J T, Liu X N, Liu K, Xu F R, Guo H, Tian X L, et al. An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease. Computers and Electronics in Agriculture, 2022; 194: 106780.

[29] Chen Z Y, Wu R H, Lin Y Y, Li C Y, Chen S Y, Yuan Z N, et al. Plant disease recognition model based on improved YOLOv5. Agronomy, 2022; 12(2): 365.

[30] Hughes D P, Salathé M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv, 2015; In press. doi: 10.48550/arXiv.1511.08060.

[31] Shi Z J, Shi M Y, Lin W G. The implementation of crawling news page based on incremental web crawler. In: 2016 4th Intl Conf on Applied Computing and Information Technology/3rd Intl Conf on Computational Science/Intelligence and Applied Informatics/1st Intl Conf on Big Data, Cloud Computing, Data Science & Engineering (ACIT-CSII-BCD): IEEE, 2016; pp.348–351. doi: 10.1109/ACIT-CSII-BCD.2016.073.

[32] Zhu J-Y, Park T, Isola P, Efros A A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, Venice: IEEE, 2017; pp.2223–2232. doi: 10.1109/ICCV.2017.244.

[33] Bochkovskiy A, Wang C-Y, Liao H-Y M. Yolov4: Optimal speed and accuracy of object detection. arXiv, 2020; In press. doi: 10.48550/arXiv.2004.10934.

[34] Zhang H Y, Cisse M, Dauphin Y N, Lopez-Paz D. mixup: Beyond empirical risk minimization. arXiv, 2017; In press. doi: 10.48550/arXiv.1710.09412.

[35] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas: IEEE, 2016; pp.779–788. doi: 10.1109/CVPR.2016.91.

[36] Hao S Y, Lee D-H, Zhao D. Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system. Transportation Research Part C:Emerging Technologies, 2019; 107: 287–300.

[37] Woo S, Park J, Lee J-Y, Kweon I S. Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV), Munich: Computer Vision Foundation, 2018; pp.3–19. doi: 10.1007/978-3-030-01234-2_1.

[38] Tan M X, Pang R M, Le Q V. Efficientdet: Scalable and efficient object detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle: IEEE, 2020; pp.10778–10787. doi: 10.1109/CVPR42600.2020.01079

[39] Han K, Wang Y H, Tian Q, Guo J Y, Xu C J, Xu C. Ghostnet: More features from cheap operations. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle: IEEE, 2020; pp.1580–1589. doi: 10.1109/CVPR42600.2020.00165.

[40] Liu H F, Chen J X, Dy J, Fu Y. Transforming complex problems into *k*-means solutions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023; 45(7): 9149–9168.

[41] Zhang Y J, Ma B X, Hu Y T, Li C, Li Y J. Accurate cotton diseases and pests detection in complex background based on an improved YOLOX model. Computers and Electronics in Agriculture, 2022; 203: 107484.