Abnormal Crops Image Data Acquisition Strategy by Exploiting Edge Intelligence and Dynamic-Static Synergy in Smart Agriculture

Xiaomin Li, *Member, IEEE*, Bingfa Hou, Hao Tang, *Senior Member, IEEE*, Bandeh Ali Talpur, Zeeshan Zeeshan, Uzair Aslam Bhatti, *Senior Member, IEEE*, Juan Liao, Jinru Liu, Bayan Alabdullah, and Imad Saud Al Naimi

***Abstract*—Abnormal crops image data play crucial role in controlling crop diseases and pest for smart agriculture. However, current agricultural image acquisition methods suffer from low-value data. This article presents a new strategy for collect high-quality image data for abnormal crops. First, a novel agricultural Internet of Things (IoT) image acquisition system is proposed, that integrates edge intelligence, motion–static synergy, which enables both coarse and fine crop image acquisition. To enhance image acquisition efficiency and data value in the agricultural IoT, this article proposes an image acquisition method based on edge intelligence and static and motion collaboration, using banana plantations as the example object. The method comprises three phases. In the first phase, the edge server deploys the YOLO-FDAC target detection model to detect abnormal crops from the images captured by static nodes. In the second phase, the coordinate solution method of abnormal crops and the quantification method of the degree of abnormality is presented. In the third phase, based on the severity of abnormality and the ant colony optimization, a path optimization algorithm for the image acquisition robot is designed. Finally, this article evaluates the performance of each level of the proposed method by comparing it with traditional methods. The experimental results demonstrate that the proposed image acquisition strategy has high acquisition efficiency and high image data value.**

***Index Terms*—Edge intelligence, image acquisition, motion collaboration, path optimization target detection.**

Manuscript received 6 May 2024; revised 29 May 2024; accepted 10 June 2024. Date of publication 15 July 2024; date of current version 24 July 2024. This work was supported in part by the NSFC Foreign Scholars Research Fund Project under Grant 62350410483 and in part by the Princess Nourah bint Abdulrahman University Researchers under Grant PNURSP2024R440. *(Corresponding author: Uzair Aslam Bhatti.)*

Xiaomin Li and Bingfa Hou are with the School of Mechanical and Electrical Engineering, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China (e-mail: xiaocao0624@163.com; houbingfa53@163.com).

Hao Tang, Uzair Aslam Bhatti, and Jinru Liu are with the School of information and Communicaiton Engineering, Hainan University, Haikou 510642, China (e-mail: melineth@hainanu.edu.cn; uzair@hainanu.edu.cn; liujr07@163.com).

Bandeh Ali Talpur is with the School of Computer Science and Statistics, Trinity College Dublin, Dublin 2 Dublin, Ireland (e-mail: talpurba@tcd.ie).

Zeeshan Zeeshan is with the IMPINJ INC, Seattle, WA 98109 USA (e-mail: zeesystech@gmail.com).

Juan Liao is with the College of Engineering, South China Agricultural University, Guangzhou 510642, China (e-mail: liaojuan0529@scau.edu.cn).

Bayan Alabdullah is with the Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia (e-mail: Bialabdullah@pnu.edu.sa).

Imad Saud Al Naimi is with the National University of Science and Technology Muscat 620, Oman (e-mail: imadalnaimi@nu.edu.om).

Digital Object Identifier 10.1109/JSTARS.2024.3414306

I. Introduction

CROPS plays a crucial role in the national economy and feeding world population [1]. To expedite the advancement of agricultural informatization and modernization, researchers have proposed the smart agriculture, unmanned farm paradigm [2], which utilizing Internet of Things (IoT) and artificial intelligence technologies. Specially, crop images data are crucial data for these new agricultural framework [3]. These image data can be used to visually monitor crop growth and obtain detailed information about the crop, such as leaf area [4], [5], plant nutritional status [6], pests diseases [7], [8], plant height [9], [10], and fruit size [11]. Obviously, how to obtain image data becomes a key link.

Recently, some researchers use cameras in the field or laboratory to capture crop images [12], [13], [14], primarily to study the feasibility of automatic detection or diagnostic algorithms in detecting agricultural pests and diseases. Others use handheld portable image capture devices to collect crop images [15], [16], [17]. However, these methods are not suitable for capturing images of crop pests and diseases in hard-to-reach locations, and they lack real-time capabilities. To address this issue, some researchers have installed network cameras in agricultural fields based on IoT [18], [19], [20], [21], [22]. However, due to their immobility, they can only capture images of crop canopies in specific areas. Furthermore, some researchers have installed image capture nodes on mobile platforms, such as unmanned vehicles [23], [24], [25] or drones [26], [27], [28].

A synergistic approach to image acquisition in both motion and static can provide more comprehensive, accurate, and real-time agricultural monitoring and management. Furthermore, the utilization of edge servers for integrating and processing sensed data from end devices can alleviate the burden of data processing in the cloud. Deploying webcams in farmland provides overall crop growth, while mobile robots can perform individualized monitoring and management. By combining the two image acquisition methods, a real-time monitoring and early warning system based on edge intelligence can be established. This system can enable refined agricultural management, improve crop growth efficiency and yield, and reduce resource waste.

This article presents a case study of a banana plantation as an entry point, focusing on abnormal banana trees (e.g., a banana

tree with a change in appearance due to disease), edge computing is utilized to decrease the expense of image transmission and improve the effectiveness of the data, and robot is used to getting more finer crop images. Finally, a strategy for acquiring high-quality image data of abnormal crops based on edge intelligence and motion–static synergy is proposed. The main contributions of this article are as follows.

1) The traditional single-form image acquisition method, usually by a robot or a surveillance camera to take pictures at a fixed time, not only generates a large amount of image data, but also the availability of data is extremely low. Therefore, this article proposes an image acquisition framework based on edge intelligence and dynamic and static collaboration to obtain fine images that can reflect more crop details, while reducing the acquisition of images that are not interesting to users. This method improves the efficiency and data value density of agricultural IoT image acquisition.
2) A fast detection method for abnormal crop based on YOLOv7 (YOLOV7-FDAC) was proposed to screen the images collected by static acquisition nodes (fixed cameras) and retain the image data of abnormal crop. This method not only reduces the data transmission, but also provides the target for the robot to carry out the fine image acquisition task.
3) Aiming at the problems that binocular stereo vision, depth camera, laser ranging, and other positioning methods are faced with, such as large calculation amount, easy to be affected by light, and high hardware cost, a method of abnormal crop location based on monocular vision RGB images is proposed to quickly calculate the actual location of abnormal plants in farmland, and provide basic support for robot work.
4) In order to enable the mobile image acquisition node (robot) to acquire fine images of abnormal crops efficiently, reduce the path length and time in the acquisition process, and reduce the collection of useless data, this article proposes a heuristic path optimization algorithm based on the severity of abnormal crops, so that the robot can quickly and accurately reach the location of abnormal crops. The robot path planning method balances the relationship between path length and abnormal crops, and obtains an efficient data acquisition path.

The rest of this article is organized as follows. Section II provides an overview of relevant research on image data in agricultural IoT, while Section III explains the overall framework of the proposed image acquisition method and the metrics used to evaluate the system's performance. Section IV proposes a fast method for detecting and localizing abnormal crops, as well as an intelligent scheduling method for robots. Section V presents experiments and analysis of the proposed method. Finally, Section VI concludes this article.

II. Related Works

This section briefly introduces representative image acquisition methods and applications in agricultural IoT.

A. Task-Oriented Image Acquisition Method

Literature [29] describes a modern farm prototype based on IoT and image processing technologies using multiple cameras to capture images and assess crop growth. Literature [20] describes a high-quality image acquisition system for the farm environment based on a DSLR camera for remote camera access, shooting control, and image transmission. Literature [30] develops an automated image acquisition system for grape yield prediction to provide a solution for automated grape yield prediction. An image acquisition method for kiwifruit picking robot is discussed in [31], which proposes a look-up capture method to acquire fruit images. Literature [32] uses RGB images acquired by UAV imaging system and divides the images into different blocks for weed detection. Literature [33] describes an image acquisition system for pig pens and also gives the process of analyzing the images. Literature [34] reviewed the application of small UAS in precision agriculture and emphasized on UAV imagery as a low-cost alternative to high-resolution satellite imagery.

B. Banana Plant Detection Methods

Literature [35] discussed the detection and prevention methods of banana streak virus disease in banana plants and proposed an economic threshold level algorithm to detect banana streak virus disease. Literature [36] introduced an artificial intelligence banana pest and disease detection system based on deep convolutional neural network and transfer learning, and developed six different models for comparison. Literature [37] proposed a deep learning model for banana wilt detection. Literature [38] discussed a banana disease detection and classification method based on local binary pattern and support vector machine (SVM) and successfully classified leaf spot disease. Literature [39] combined deep learning models for banana classification and disease detection by using multispectral and RGB image data, while class activation maps were used to understand the prediction results of the classification models. Literature [40] combined machine learning and deep learning techniques for banana disease detection in two parts, CNN for feature extraction of images through convolutional layers and pooling operations and then classification by SVM classifier.

C. Data Processing Methods for Edge Intelligence

Literature [41] proposes a hierarchical data processing architecture that reduces communication bottlenecks and energy consumption, applies edge computing to sensor data processing and analysis in precision agriculture, and aggregates and reconstructs data through fog computing nodes. Literature [42] explores the application of intelligent edge computing in satellite IoT image target detection, which performs data processing and analysis at the edge of the satellite to achieve real-time target detection, reduces data transmission delay and bandwidth consumption, and improves processing efficiency. Literature [43] uses lightweight edge mining algorithms to compress agricultural data within WSNs to address the challenges of poor Internet connectivity and memory constraints of WSNs devices

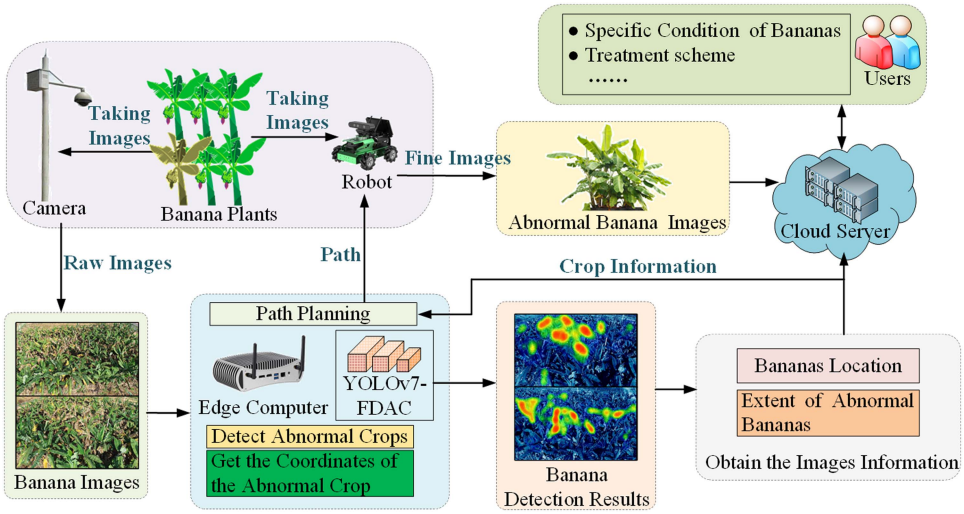


Fig. 1. High-quality image data acquisition model for abnormal crops based on edge intelligence and motion–static synergy.

in agricultural environments. Literature [44] proposed an edge computing framework for collaborative video processing in multimedia IoT systems that leverages the computational and communication capabilities of resource-rich mobile devices to extract features from videos and avoid bandwidth constraints. Literature [45] proposed an edge computing based target detection architecture for distributed and efficient target detection over wireless communication. Literature [46] proposed a data filter based on edge intelligence to solve the problem of large amount of useless data causing congestion in the communication link. Literature [47] combines edge computing and IoT to build a lightweight crop lifecycle data sensing framework for multi-parameter and mobile sensing, and proposes a data-driven algorithm to optimize the sensing parameters, reduce redundancy, and improve the correlation between the sensed data and crop growth stages.

III. System Model

This section describes the methodology for acquiring high-quality image data of abnormal crops and then explains the methodology used to assess the degree of crop abnormality and the density of image value acquisition.

A. Design of Image Data Acquisition Framework

The architecture of the image data acquisition system proposed in this article is shown in Fig. 1, taking banana abnormal plants as an example. The workflow is as follows: first, the static nodes are used to collect the whole image of the field, and then the edge server uses the improved YOLOv7-FDAC target detection model to quickly detect the abnormal crops that require focused attention, and then solves the geographical location from the image information and camera parameters. At the same time, the severity of each abnormal plant is quantified based on the data generated by the target detection model, such as the confidence level and abnormal area. Finally, the edge server will send the path to the robot so that it can move next to the abnormal banana, collect a fine image containing the details of the banana disease, and transmit it to the user or the cloud so that the user can further analyze the specific condition of the abnormal banana plant (such as what disease is present) and adjust the planting strategy in time.

B. Acquisition Methods for Fine Images

In order to quickly acquire fine-grained images with crop details, this article uses a mobile robot carrying a camera to capture images. The traditional method of using robots to acquire images generally adopts a patrolling approach, where the robot acquires images of all crops in its path and transmits them to the user. Not only is the acquisition process very inefficient, but it also generates a large amount of data traffic, placing a heavy burden on edge devices and networks. In addition, only some of the images captured by the robot are needed by the user, such as images of abnormal banana tree.

This article proposes a solution. When abnormal crops are detected in the images collected by static nodes, the edge server calculates their coordinates and plans the robot's moving path, and the robot collects images of abnormal crops along the path. In the path planning process, it is assumed that X is the ensemble of crops that the robot needs to observe, namely, $X = \{X_1, X_2, \ldots, X_i, \ldots, X_n\}$. To enable the robot to observe the crops X_i with relatively high severity as soon as possible, a heuristic path optimization algorithm based on the severity of the abnormal crops is proposed to obtain an optimal acquisition path and improve the efficiency of image acquisition.

C. Methods for Quantifying the Severity of Abnormal Crops

Since the severity of anomalous crops directly affects the order of destination selection by the path planning method proposed in this article, which ultimately affects the efficiency of image acquisition, the crop anomaly degree index (CADI) is proposed.

Definition 1: CADI is the result of combining the size of the area occupied by each abnormal crop in the image with its

corresponding confidence level, and its result is proportional to the area and the confidence level, reflecting the severity of the abnormal crop.

Specifically, if an image contains *n* abnormal banana trees, and *i* is the number of one of the banana trees, its CADI expression is

$$K_i = P \times S_i \times W_i \ , W \in [W_{\min}, 1] \tag{1}$$

where *W* denotes the confidence level of the target detection result. $W_{\min}$ is the confidence threshold, the crops with confidence lower than the threshold are considered normal, and the default confidence threshold in this paper is $W_{\min} = 0.3$. *P* denotes the accuracy of the target detection model, i.e., the performance evaluation index of the target detection model with 0.5 intersection average accuracy (AP@0.5). *S* is the area occupied by the abnormal crop in the image, which is calculated by the pixel length (*dx*) and width (*dy*) of the image, and the number of pixels occupied by the crop. The number of pixels *c* is calculated and obtained as follows:

$$S = c \times (dx \times dy). \tag{2}$$

D. Image Data Value Analysis

Among the image data acquired by the robot, the images of abnormal crops are the main target of interest, and in order to evaluate the acquisition efficiency of the image acquisition method proposed in this paper for the images of abnormal crops, the image value acquisition density (IVAD) is proposed.

Definition 2: The ratio of the image data value acquired by the robot to the amount of data in a unit of time, the result of which depends on the total data value, the total amount of data, and the execution time of the image acquisition task performed by the robot.

Specifically, the robot performs an image acquisition task where images of *n* banana trees need to be captured, and *i* denotes the number of one of the banana trees. The value of the image corresponding to banana tree *i* is denoted as $G(K_i)$, the time for the image to be transmitted to the user is $T_s(i)$, the size of the data volume of the image is Q_i, and the total time taken by the robot to perform one image acquisition task is denoted as T_m. Then, the expression of IVAD is

$$M = {T_m}^{-1} \sum_{i=1}^{n} \left(T_i \times {Q_i}^{-1} \times G(K_i)\right) \tag{3}$$

where $G(K_i) \propto K$, i.e., the higher the CADI score of the crop, the higher the value of its corresponding image.

Furthermore, assuming that the shooting time of the image of crop *i* is $T_p(i)$, $T_s(i)$ is the elapsed time of the image transmission to the cloud, and the walking time of moving from crop *i* to crop $i+1$ is $T_w(i)$. Considering that the robot, after capturing a set of images, can simultaneously transmit images during the process of moving to the next target point, and that the image transmission time is much smaller than the robot moving time, i.e., $T_s(i) \subseteq T_w(i)$. In this case, $T_s(i)$ does not need to be counted in elapsed time, then T_m can be further expressed as

$$T_m = \sum_{i=1}^{n} T_w(i) + \sum_{i=1}^{n} T_p(i). \tag{4}$$

The final expression for IVAD is

$$M\left(T_w(i), T_p(i), Q_i, G(K_i)\right) = \left(\sum_{i=1}^{N} T_w(i) + \sum_{i=1}^{N} T_p(i)\right)^{-1} \sum_{i=1}^{N} \left({Q_i}^{-1} \times G(K_i)\right). \tag{5}$$

Considering the methodology defined in this article for acquiring high-quality image data of abnormal crops, the ultimate goal is to maximize the IVAD, which can be expressed as

$$\underset{G(K_i)}{Max}\ M\left(T_w(i), T_p(i), Q_i, G(K_i)\right) \tag{6}$$

$$\sum_{i=1}^{N} T_w(i) + \sum_{i=1}^{N} T_p(i) \le T \tag{6b}$$

where the constraint *T* in (6b) is the maximum operating time that the battery carried by the robot can support, within which the robot must complete the image acquisition task.

IV. High-Quality Image Data Acquisition Method for Abnormal Crops Based on Edge Intelligence and Kinetic Collaboration

The maximization of IVAD needs to be achieved by two aspects. First, accurately detecting the abnormal crop and obtaining its position information. Second, taking the value of the image as the starting point, optimizing the path for image acquisition by the ground robot. This section focuses on the specific implementation methods, first introducing the construction of fast detection model based on lightweight neural network, then explaining the crop position mapping method based on image and coordinate system transformation, and finally giving the heuristic path optimization algorithm based on the degree of crop abnormality. The specific process is shown in the Fig. 2.

A. Construction of YOLOv7-FDAC Target Detection Model

YOLOv7-tiny is a simplified network model based on the structure of YOLOv7 [48], which mainly consists of backbone, neck, and head. In the backbone, the more concise ELAN [49] is used instead of E-ELAN. The convolution operation in MPConv is eliminated and only pooling is used for downsampled. The optimized SPP structure is retained for inputting richer feature maps for the neck. In the neck, the PANet structure is still used for feature aggregation. In the head, the standard convolution (SConv) is used instead of REPConv for channel number adjustment. YOLOv7-tiny has some advantages in terms of speed and lightness, but it sacrifices some accuracy. However, the network still has some shortcomings: First, due to the use of a large number of ELAN modules in the backbone, which results in too many network parameters and too much computation, there is still room for optimization of the detection speed. Second, the choice of using the LeakyRelu activation function, which is less effective when the model hierarchy is deeper, resulting

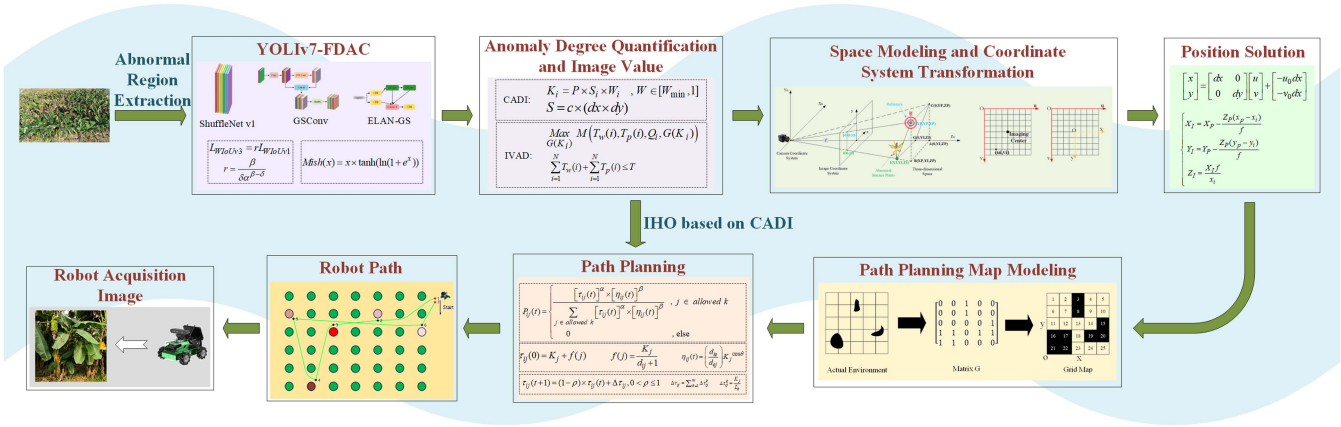


Fig. 2. Process for maximizing the IVAD.

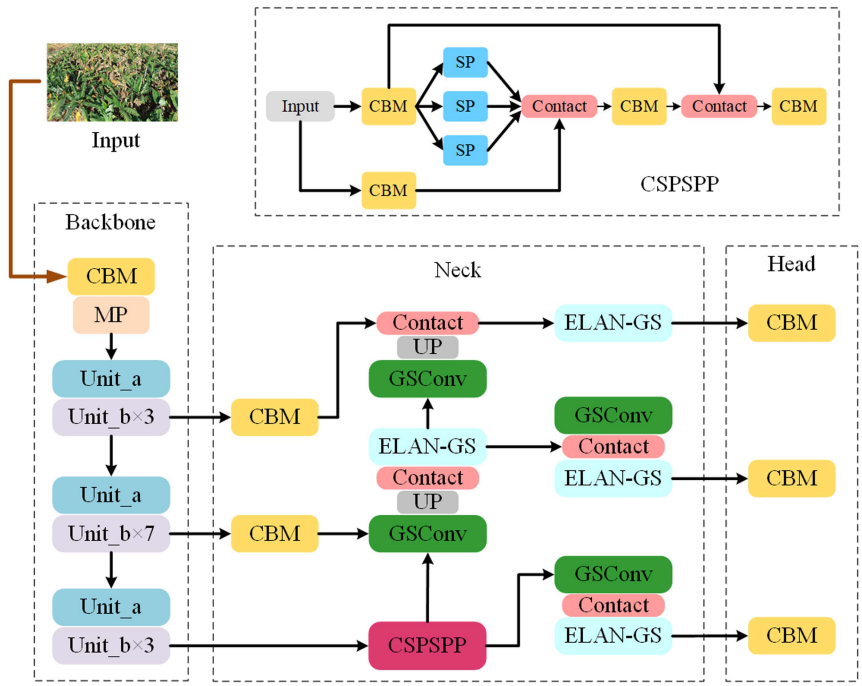


Fig. 3. Network structure of YOLOv7-FDAC.

in the classification accuracy is affected. Third, in the neck for feature fusion, the ELAN module is also overused, which is easy to cause feature redundancy. Therefore, in this article, based on YOLOv7-tiny, the YOLOv7-FDAC model is obtained after lightweight improvement, which reduces the number of parameters and computation of the model under the premise of guaranteeing the accuracy, so as to run in the edge server.

The network structure of the improved YOLOv7-FDAC model is shown in Fig. 3. In the backbone, YOLOv7-FDAC adopts the basic modules of ShuffleNet v1 [50] network, i.e., Unit_a with step size of 1 and Unit_b with step size of 2, to minimize the dense connectivity and reduce the computation of the network, and at the same time, combines multiple units to increase the network depth and ensure feature richness [51]. In the improvement of Neck, the lightweight module GSConv [52] is used to fuse the features, while the improved ELAN-GS is used to further reduce the parameters and computation of the model. WIoUv3 is selected to replace the original IoU metric to achieve high-precision detection of abnormal banana trees in complex environments, and a more efficient activation function Mish [53] is introduced to replace the LeakyRelu function to ensure that the features can be efficiently executed in the backward transfer process. Finally, the WIoUv3 position loss function is used to improve the model localization accuracy.

ShuffleNet v1: Fig. 4 shows the basic module of ShuffleNet v1, which consists of Unit_a with step size 1 and Unit_b with step size 2 stacked on top of each other. Unit_a is divided into two branches, where the left branch is not processed, the right branch is subjected to group convolution (GC) and depth-separable convolution, normalization and channel mixing, and the feature

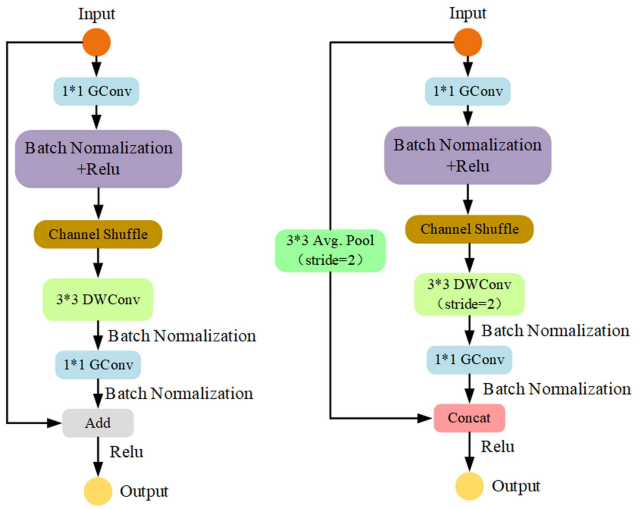


Fig. 4. Architecture of the ShuffleNet v1 module.

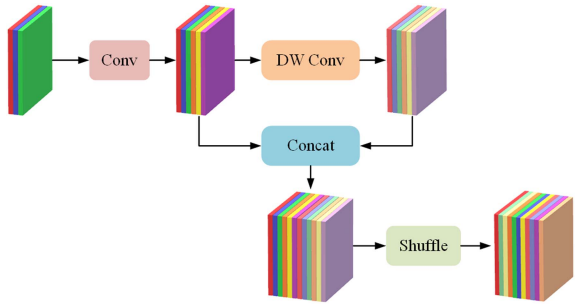


Fig. 5. GSConv module.

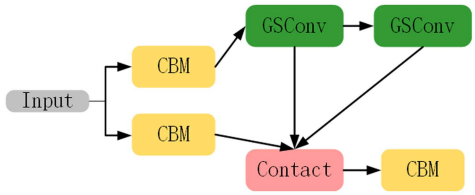


Fig. 6. Improved ELAN module (ELAN-GS).

maps of the two branches are summed. Unit_b is also divided into also branches, both downsampled, the left branch halves the size of the feature map using an average pooling operation, the right branch performs a convolution operation with a step size of 2, and finally latitude doubling is performed using a splicing operation.

The number of parameters resulting from each convolution operation is calculated as follows. Assume that the input feature map has a width of W_i, a height of H_i, the number of channels is C_i, the number of output channels is C_o, and the size of the convolution kernel is $K \times K$. Therefore, the number of parameters under SConv is

$$P_{\text{sc}} = C_i \times C_o \times K \times K. \tag{7}$$

GC is based on the Sconv, the convolution kernel and the number of input channels are grouped, the number of groups is g. Each group of convolution kernel is individually convolved on the feature map, so the number of parameters is

$$P_{\text{gc}} = \frac{C_i}{g} \times C_o \times K \times K \tag{8}$$

Depth-wise (DW) Separable Convolution is a channel-by-channel convolution operation on the input feature map with the number of convolution kernels equal to the number of input channels, so the number of parameters for depth-separable convolution is

$$P_{\text{dw}} = C_i \times K \times K. \tag{9}$$

Therefore, among the three convolution operations, the SConv has the largest number of parameters, and the GC is $1/g$. The depth-separable convolution has the smallest number of parameters, which is only $1/C_o$ of the SConv, so the combination of GC and depth-separable convolution can greatly reduce the number of parameters of the network and reduce the amount of computation.

GSConv and ELAN-GS: The structure of GSConv is shown in Fig. 5. Let the number of input channels be C_1 and the number of output channels be C_2. The input image first undergoes a SConv, the number of channels becomes $C_2/2$, and then undergoes a depth-separable convolution, the number of channels remains unchanged, and finally the results of the two convolutions are spliced and mixed, and finally C_2 is obtained.

GSConv is introduced into the ELAN module for improvement, as shown in Fig. 6. The two convolutions before the Concat layer use the GSConv, which reduces the number of parameters of the model while maintaining the detection accuracy, and achieves a slight improvement of the ELAN.

WIoUv3 loss function: The loss function is an important part of the target detection model, and the detection performance of the model depends on the design of the loss function, and a good bounding box loss function will bring significant performance improvement for the target detection model. The bounding box loss of YOLOv7-tiny is calculated by CIoU_Loss function, and the classification loss and confidence loss are calculated by BEC_Loss function. The agricultural environment is complex and variable, and banana trees often have branches and leaves that shade each other, which makes the model detection difficult.

Therefore, WIoUv3 with dynamic nonmonotonic focusing mechanism is chosen instead of CIoU_Loss as the bounding box loss calculation function of the improved algorithm model in this article. Positive and negative sample imbalance is unavoidable in the training dataset, which inevitably leads to the emergence of low-quality samples, and the previous loss function will exacerbate the punishment of low-quality samples, thus reducing the generalization ability of the model. The dynamic nonmonotonic focusing mechanism in WIoUv3 can effectively avoid the negative impact of low-quality samples during the training process by balancing the proportion of high- and low-quality samples, focusing the results of bounding box regression on the target object, and solving the problem that banana leaves are difficult to detect due to occlusion between them. WIoUv3 is based on

the addition of the dynamic nonmonotonic focusing mechanism to WIoUv1, and the specific computational expression is

$$L_{\text{WioUv1}} = R_{\text{WIoU}} L_{\text{IoU}} \tag{10}$$

$$R_{\text{WIoU}} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \tag{11}$$

where W_g and H_g are the width and height of the minimally closed region of the predicted and true frames, and * denotes the separation of W_g and H_g from the graph and the localization of constants to prevent the creation of gradients that prevent convergence

$$L_{\text{WIoUv3}} = rL_{\text{WIoUv1}} \tag{12}$$

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{13}$$

where the mapping of nonmonotonic focusing coefficients r and outliers β is controlled by hyperparameters α, δ, and δ.

Mish activation function: To ensure the accuracy of the feature extraction network, the Mish activation function is chosen as an alternative to the LeakeyReLU activation function because it has a minimum value at the zero point, which can effectively buffer the weights and maintain the stability of the network. Compared to LeakeyReLU, the Mish function exhibits smoother and stronger derivability near the zero point, which is helpful for gradient computation and updating, and speeds up the convergence of the model. In addition, Mish adds more nonlinear expressions, which improves model generalization. The specific formulas for the LeakeyReLU and Mish activation functions as follows:

$$\text{LeakeyReLU}(x) = \begin{cases} x, x > 0 \\ \alpha x, x \le 0 \end{cases} \tag{14}$$

$$R_{\text{WIoU}} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right). \tag{15}$$

B. Abnormal Crops Location Mapping Method

Accurately locating the position of abnormal crops is the key to planning the robot path, which is related to whether the robot can successfully reach the crops for image acquisition operations. To realize the perception of the environment, the vision sensing system mainly acquires the image information of the environment by carrying a camera. For the localization of abnormal crops, it is necessary to meet the conditions of 3-D perception. At present, there are two main forms of stereo vision systems, the first is a binocular vision system based on optical geometry. The 3-D position of the target is obtained by traditional optical principles and optimization algorithms. The second is an RGB-D camera based on the time-of-flight method, which uses an infrared sensor to obtain the depth information of the target. However, the binocular vision system needs to go through a complicated calibration process, while stereo matching consumes a lot of computational resources, and the RGB-D camera is affected under strong illumination, leading to errors in the results.

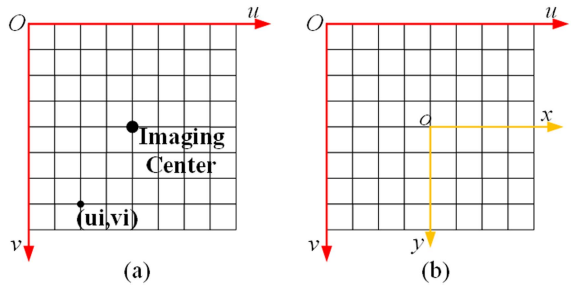


Fig. 7. Transformation of image coordinate system and pixel coordinate system. (a) uov in pixel coordinate system. (b) Image coordinate system xoy.

Based on this, this article proposes an abnormal crop location mapping method based on camera imaging principle and coordinate system transformation principle. First, a coordinate system for the orchard is established, then a reference is set in the orchard, and finally the camera imaging principle and the coordinate system transformation principle are combined to locate the abnormal crop. Since the positions of the reference and the camera are known, only the direction and distance of the abnormal crop relative to the reference in the image need to be calculated, and then its position can be mapped.

The camera model describes the process of mapping an object from 3-D space to a 2-D image plane, and the camera-internal reference describes the key parameters in the camera model. The relationship between the actual 3-D space points and the 2-D image plane is usually calculated using the principle of small-hole imaging. In practice, the image data are calculated in terms of pixel points, as shown in Fig. 7(a), and the pixel coordinates (u_i, v_i) are usually read in terms of the *uov* plane under the pixel coordinate system, which can directly correspond to the position in the image. To facilitate the calculation, the image coordinate system *xoy* is introduced, as shown in Fig. 7(b). Where the image center at o is the center of the image coordinate system, and the *x*- and *y*-axes are parallel to the *u*- and *v*-axes, respectively.

Each camera has a fixed image size for pixels, each pixel size length and width are *dx* and *dy*. Assuming that the pixel corresponding to the origin *o* of the image coordinate system is (u_o, v_o), its specific coordinates in the image coordinate system can be computed by the given specific pixel coordinates (*u*,*v*)

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} dx & 0 \\ 0 & dy \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} -u_0 dx \\ -v_0 dx \end{bmatrix}. \tag{16}$$

The mapping from a 2-D pixel coordinate system to a 2-D image coordinate system can be realized by the above formulas, and further the conversion from a 2-D image to 3-D coordinates in space can be realized by the spatial depth information.

The crop localization calculation model is shown in Fig. 8. Assuming that point *P* and point *I* in the Fig. 8 are the positions of the reference and the anomalous crop in 3-D space, respectively, point *p* and point *i* are the positions of both mapped in the image. $p(x_p, y_p)$ and $i(x_i, y_i)$ can be calculated by (16), $\text{P}(X_p, Y_p, Z_p)$ are the known coordinates of the reference relative to the camera, and f is the focal length of the camera. According to the

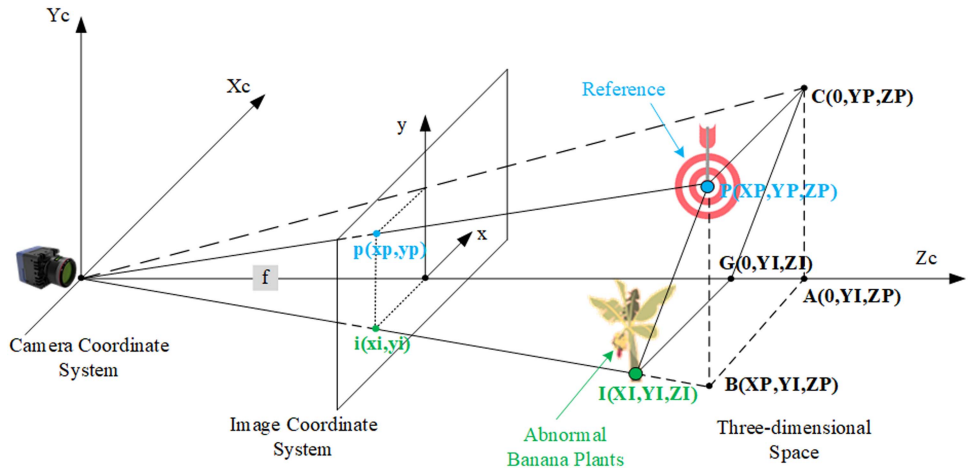


Fig. 8. Crop positioning model.

properties of similar triangles, there are

$$\frac{Z_P}{f} = \frac{Y_p - Y_I}{y_p - y_i} = \frac{X_p - X_I}{x_p - x_i} \tag{17}$$

$$\frac{X_p}{x_i} = \frac{Z_I}{f}. \tag{18}$$

The realistic 3-D coordinates of the anomalous crop I are solved for

$$\begin{cases} X_I = X_P - \frac{Z_P(x_p - x_i)}{f} \\ Y_I = Y_P - \frac{Z_P(y_p - y_i)}{f} \\ Z_I = \frac{X_I f}{x_i}. \end{cases} \tag{19}$$

The pseudocode shown in Algorithm 1 summarizes the detection and localization process of the proposed anomalous crop.

C. Improved Heuristic Path Optimization (IHO) Based on Abnormal Crop Severity

Path planning, as one of the most fundamental and important steps in the execution of image acquisition tasks by mobile robots, determines the efficiency of robots in performing image acquisition tasks. Ant colony optimization (ACO) [54] has been widely used in multiobjective path planning research due to its strong robustness and adaptability, but the algorithm has poor convergence and tends to fall into local optimum when searching for the optimal path. In addition, its path planning approach based on the traveler problem cannot meet the requirement of efficient image acquisition in this study. Therefore, this article proposes an IHO algorithm based on abnormal crop severity and ACO, which solves the optimal traversal order by combining CADI to obtain the optimal working path of the robot.

Environment modeling: This article studies the robot path planning problem in a known environment. In order to simplify the computational process, the grid method is used to abstract and discretize the robot working environment. First, the environment is divided into several identical grids according to its scale and requirements, and then the conditions of obstacles in the grid

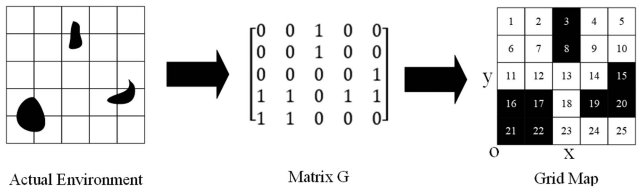


Fig. 9. Schematic diagram of the conversion between the actual environment and the raster map.

are set according to the actual environment, and the corresponding matrix G is constructed, which is finally transformed into a grid map. The 0 and 1 in the matrix denote the passable nodes and obstacle nodes, respectively, which are denoted as white grid and black grid in the grid map. The conversion scheme of the actual environment and the grid map is shown in Fig. 9.

Each grid in the grid map has a corresponding label and corresponding position coordinates, and the relationship between the grid label and the position coordinates is expressed as follows:

$$\begin{cases} x = \mathrm{mod}(\frac{N_i}{N}) + c \\ y = \mathrm{ceil}(\frac{N_i}{N}) - c \end{cases} \tag{20}$$

where x and y are the horizontal and vertical coordinates of the grid position, mod() is the remainder operation, ceil() is the rounding operation, N is the number of grids in each column, and N_i is the label of the ith grid node. $c = 0.5$, which represents the offset of the center of the grid with respect to the grid boundary.

Base transfer probability: The direction, in which ant $k(k = 1, 2, \ldots, m)$ transfers at moment t is determined by the pheromone concentration on each to-be-selected path. The probability that an ant selects the next position j from position i at moment t is determined as follows:

$$P_{ij}(t) = \begin{cases} \frac{[\tau_{ij}(t)]^{\alpha} \times [\eta_{ij}(t)]^{\beta}}{\sum_{j \in \text{allowed}k} [\tau_{ij}(t)]^{\alpha} \times [\eta_{ij}(t)]^{\beta}}, & if\ j \in \text{allowed}\ k \\ 0, & \text{else} \end{cases} \tag{21}$$

Algorithm 1: Detection of Abnormal Crops and Solution of Their Severity and Coordinates.

Input: Image A acquired by static node
Output: CADI and location coordinates of abnormal banana tree X_i
procedure INFORMATION EXTRACTION(A)
  YOLO-FDAC detects abnormal banana trees X_1, X_2, $\cdots$, X_i, $\cdots$, X_n and confidence W_1, W_2, $\cdots$, W_i, $\cdots$, W_n
  for $i = 1$ to n **do**
    if $W_i \geq 0.3$ **then**
      $S_i \leftarrow c_i \times (dx \times dy)$
      \\ Calculate the anomalous area S_i by Equation (2)
      $K_i \leftarrow P \times S_i \times W_i$
      \\ Calculate the anomaly index K_i of crop X_i by Equation (1) using W_i and S_i
      Read the camera's internal reference and reference P coordinates $P(X_P, Y_P, Z_P)$
      Get coordinates $x(x_i, y_i)$ and $p(x_p, y_p)$ by (16)
      \\ Calculate the coordinates of crop X_i and reference P in the image according to equation (16)
      $X_i \leftarrow X_P - \frac{Z_P(x_p - x_i)}{f}$
      $Y_i \leftarrow Y_P - \frac{Z_P(y_p - y_i)}{f}$
      $Z_i \leftarrow \frac{X_i f}{x_i}$
      \\ Calculate the position coordinates of crop X_i relative to the camera by equation (19)
    else
      break
    end if
  end for
  $K \leftarrow \{K_1, K_2, \cdots, K_i, \cdots, K_n\}$
  \\ K is the anomaly index of all banana trees
  $X \leftarrow \{X_1(X_1, Y_1, Z_1), X_2(X_2, Y_2, Z_2), \cdots, X_i(X_i, Y_i, Z_i), \cdots, X_n(X_n, Y_n, Z_n)\}$
  \\ X is the set of coordinates of all abnormal banana trees
  return K, X
end procedure

where $\tau_{ij}(t)$ is the pheromone on the path between position *i* and position *j* at moment *t*. $\eta_{ij}(t)$ is the heuristic function factor, which indicates the expected degree of the ant's transfer from position *i* to position *j* at moment *t*, and is usually taken to be $\eta_{ij}(t)=1/d_{ij}$. d_{ij} is the Euclidean distance between position *i* and position *j*. α is the information heuristic factor, β is the expectation heuristic factor, both α and β are constants, and allowed k denotes the ensemble of unvisited target points.

In this study, the path planning behavior of the ants is adjusted according to the severity of the crop, so that the ants are more likely to be attracted to the abnormal crop and can select the optimal image acquisition path in a more targeted way.

CADI-based pheromone initialization: Under the initial condition, the pheromone content on each path is equal, at this time, the ant colony is in the stage of blind search, with poor optimization seeking effect and low search efficiency. In this article, combining ACO and crop severity, CADI is introduced as a weight in the initial pheromone, which guides the ants to prioritize the crop with higher severity as the destination. The improved initial pheromone distribution is

$$\tau_{ij}(0) = K_j + f(j) \tag{22}$$

$$f(j) = \frac{K_j}{d_{ij} + 1} \tag{23}$$

where K_j is the CADI score of the targe crop j, i.e., the higher the severity of the crop, the larger the value of K, the higher the attraction of the target for the ants, and vice versa, the smaller it is. d_{ij} is the distance between the current grid and the target grid, the closer the current grid is to the target grid, the larger the initial pheromone on the route, and vice versa, the smaller it is. According to such positional relationship, the initial pheromone with uneven distribution is set to avoid blind searching of ants in the initial stage and to improve the searching efficiency of ant colony on abnormal crops in the initial stage.

Oriented heuristic function: The heuristic function of ACO generally takes the inverse of the distance of neighboring grid, so the ants tend to choose the optional grid closer to the current grid, but this situation will cause the ants to have a circuitous path or to get stuck when choosing, resulting in inefficient and ineffective search. In this article, we design a heuristic function oriented to abnormal crops, forcing the ants to be more inclined to choose the grid closer to the abnormal crops each time, so that the final path obtained will be closer to the shortest path, and the improved distance heuristic function formula is

$$\eta_{ij}(t) = \left(\frac{d_{ie}}{d_{ej}}\right) K_j^{\cos\theta} \tag{24}$$

where d_{ie} is the Euclidean distance from the current grid to the middle grid, d_{ej} is the Euclidean distance from the middle grid to the target grid, K_j is the CADI score of the crop j, and θ denotes the angle between $\overrightarrow{ie}$ and $\overrightarrow{ij}$.

Improved pheromone update method: Pheromone update is a crucial part of ant colony algorithm. It simulates the behavior of ants and guides the search process to the direction of better solutions. At the same time, the search diversity is maintained, so that the algorithm can effectively search in the solution space and find high-quality solutions. As the number of iterations increases, the pheromone differences on several optimal paths become less and less obvious, and the ant colony pays less and less attention to abnormal crops. In order to make the ant colony continue to pay attention to abnormal crops, combined with CADI, this article adopted a global pheromone update method to improve the pheromone update rule of the ant colony. After all ants complete one iteration, the pheromone on all paths is updated. The formula is as follows:

$$\tau_{ij}(t+1) = (1-) \times \tau_{ij}(t) + \Delta\tau_{ij}, 0 < \rho \leq 1 \tag{25}$$

$$\Delta\tau_{ij} = \sum_{k=1}^{m} \Delta\tau_{ij}^{k} \tag{26}$$

Algorithm 2: Robot Path Planning for Fine Image Data Collection.

Input: The value K of CADI and position coordinate X of all abnormal banana trees
Output: Ground robot working path

while receives K, X **do**
  procedure PATH PLANNING(K, X)
    Initialize the robot working environment, converting map coordinate X to grid coordinates by equation (20)
    for $i = 1$ to n **do**
      Calculate the initial pheromone concentration of abnormal banana tree i by equation (22)
      Calculate the heuristics of abnormal banana tree i by (24)
    end for
    for $e = 1$ to b **do**
  \\ b is the total number of ant colony iterations
      for $i = 1$ to n **do**
        Calculate the transition probability P_i of banana tree i by (21)
    end for
    $P \leftarrow \{P_1, P_2, \cdots, P_i, \cdots, P_n\}$
    Plan the robot acquisition path R_e according to P
    Update the pheromone by (25)
  end for
  From $\{R_1, R_2, \cdots, P_e, \cdots, R_b\}$ choose an optimal path as the final path R
end procedure
end while

$$\Delta\tau_{ij}^{k} = \frac{K_j}{L_k} \tag{27}$$

where ρ represents the evaporation rate of pheromone; $\Delta\tau_{ij}$ represents the sum of pheromone concentrations released by all ants along the path connecting location i to target j; $\Delta\tau_{ij}^{k}$ represents the pheromone concentration released by ant k on the path connecting location i to target j. L_k is the distance traversed by ant k from location i to target j.

The pseudocode of the proposed path planning method is summarized in Algorithm 2.

V. EXPERIMENTS AND ANALYSIS

This section focuses on the prerequisites and results of the experiments. First, the effectiveness of each improvement of YOLOv7-FDAC is evaluated and its performance is compared with other state-of-the-art models. Subsequently, the effectiveness of the position mapping method based on RGB images is verified. Then, the effectiveness of IHO in image efficient acquisition strategies is evaluated. Finally, the image-efficient acquisition strategy proposed in this article and the conventional image acquisition strategy are compared to evaluate the performance of the image-efficient acquisition strategy.

TABLE I
PARAMETERIZATION OF THE TRAINING PROCESS

Parameter	Value	Parameter	Value
Learning Rate	0.01	Batch Size	16
Image Size	640×640	Epochs	600
Weight Decay	5×10^{-4}	Momentum	0.937

TABLE II
IMAGE VALUES CORRESPONDING TO CROPS OF DIFFERENT SEVERITY LEVELS

Level of severity	I	II	III	IV	V	VI
$G(K_i)$	10	20	40	60	80	100

A. Experimental Conditions and Parameter Settings

The model training platform of YOLOv7-FDAC is built on Ubuntu 20.04 system with one Intel Xeon Gold 5218 CPU and two NVIDIA TITAN RTX GPUs, and the parameters of its training process are shown in Table I.

The performance analysis of YOLOv7-FDAC is conducted on an edge server equipped with one Intel Core i9-10920X CPU and NVIDIA GeForce RTX3090 GPU to simulate real-world application performance. The data set used in this article is the banana tree image dataset collected in the field, including more than 7000 images collected under different weather conditions, different brightness, backlight, fairing, and other conditions, and the data was cleaned, and the final dataset contains 1680 images. It is divided in a ratio of 7:2:1, in which 1176 images are the training set, 840 are the test set, and 168 are the verification set.

Based on the image acquisition framework proposed in this article, an experiment was designed to evaluate the image acquisition performance of the IHO-based robot. In the experiment, assuming that the robot moves and acquires images on a path at a speed of 0.5 m/s, each image of an abnormal banana tree takes 2 min to capture. To facilitate the calculation, this paper categorizes the severity into six levels, I–VI, based on the CADI scores, and defines the image values according to their severity levels, as shown in Table II. Among them, class I is normal and the highest severity is class VI.

B. Results and Analysis

1) Validation of the Improved YOLOv7-FDAC: In order to verify that each improvement proposed in this article is effective, a series of ablation experiments will be designed for comparative analysis, using the same parameters in the training process to ensure the accuracy of the experiments. For experiment A, the ShuffleNetv1 network is used as the new backbone network. In experiment B, GSConv module was introduced into the neck of the model for optimization. Experiment C replaces the LeakyRelu activation function at the neck of the model with the Mish activation function. Experiment D uses WIoUv3 as a loss function. To evaluate the effectiveness of ShuffleNetv1, GSConv, Mish function, and WIoUv3, the number of parameters, computational complexity, i.e., floating point operations (FLOPs), and AP0.5 are utilized as metrics to measure the performance of the models, where AP@0.5 is a widely used

TABLE III
COMPARISON OF ABLATION EXPERIMENT RESULTS

Model	Params	FLOPs	mAP@0.5
YOLOv7-tiny	6.1 M	13.7 G	84.54%
YOLOv7-tiny+A	5.3 M	11.3 G	84.23%
YOLOv7-tiny+B	6.0 M	13.0 G	85.01%
YOLOv7-tiny+C	6.1 M	13.7 G	84.88%
YOLOv7-tiny+D	6.1 M	13.7 G	84.94%
YOLOv7-tiny+A+B+C+D(Ours)	5.2 M	10.6 G	86.12%

TABLE IV
EXPERIMENTAL COMPARISON OF YOLOV7-FDAC WITH OTHER MODELS

Model	Params	FLOPs	mAP@0.5	mAP@0.5:0.95
YOLOv3-tiny	8.8 M	12.9 G	61.5%	28.1%
YOLOv4-tiny	5.7 M	16.1 G	63.4%	29.0%
YOLOv5s	7.1 M	16.4 G	70.6%	34.6%
YOLOv6s	18.6 M	45.4 G	76.9%	38.2%
YOLOX-s	8.1 M	23.7 G	84.7%	45.9%
YOLOv7-FDAC (Ours)	5.2 M	10.6 G	86.12%	47.8%

evaluation criterion in object detection, higher values indicate better model performance.

The results of the ablation experiments are shown in Table III, Experiment A verifies that ShuffleNetv1 can significantly reduce the number of parameters and computation amount of YOLOv7-tiny, but its accuracy also slightly decreases. Experiment B adopts the improvement of GSConv, which significantly reduces the computation amount of the model, and its accuracy is also basically the same as that of the original model, which proves the validity of GSConv. Experiments C and D, which do not increase the computation amount, the accuracy of the model is slightly improved, and the validity of the Mish function and WIoUv3 is verified. Finally, the combined improvement of YOLOv7-tiny with experiments A, B, C, and D reduces the amount of model parameters by 14.7%, the amount of computation is reduced by 22.6%, and the accuracy of the model is improved by 1.87%. Overall, the improved YOLOv7-FDAC network improves model accuracy while reducing model weight, effectively balancing accuracy, and weight, and providing feasibility for deployment in edge terminals.

To verify the effectiveness of the proposed model in this article, this section compares the proposed YOLOv7-FDAC model with other common lightweight target detection models. All algorithms use the same hardware devices, training parameters, and datasets to ensure the reliability and fairness of the experimental results. The Params, FLOPs, mAP@0.5%, mAP@0.5%:0.95%, and the execution speed are taken as the evaluation indexes, and the specific experimental results are shown in Table IV, Figs. 10 and 11.

As can be seen from Table IV and Fig. 9, YOLOv7-FDAC also reduces the number of parameters by 8.7% compared to the smallest YOLOv4-tiny model, while the computation is reduced by 37.2%, the detection speed is improved by 114.8%, and the mAP@0.5 and mAP@0.5:0.95 accuracies are improved by 40%

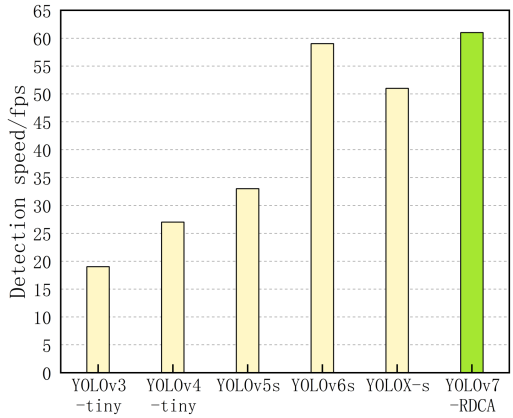


Fig. 10. Execution speed comparison of YOLOv7-FDAC with other models.

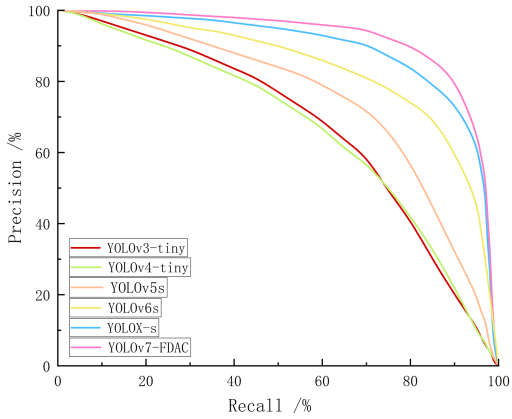


Fig. 11. Verify sets of P-R curves on different models.

TABLE V
EXPERIMENTAL COMPARISON OF YOLOV7-FDAC WITH OTHER MODELS

Distance /m	Successful solving of the position	Error from actual position /m
5	Y	0.71
10	Y	0.83
15	Y	0.85
20	Y	0.87
25	Y	0.92
30	N	N/A

In the table, Y is the successfully solved position and N indicates the failure to solve the position.

and 59.3%, respectively, when comparing YOLOv7-FDAC with the existing mainstream lightweight networks. Compared to YOLOX-s, YOLOv7-FDAC reduces the number of parameters by 35.8% and the computation by 55.2%, while the mAP@0.5 and AP@0.5:0.95 accuracies are slightly improved by 1.67% and 4.1%, respectively, and the detection speed is improved by 17.3%. Overall, YOLOv7-FDAC is not weaker than other models in the results of all five types of indicators, which indicates that the lightweight model proposed in this article even has a certain degree of accuracy improvement under the conditions of fewer number of parameters, smaller computational volume, and

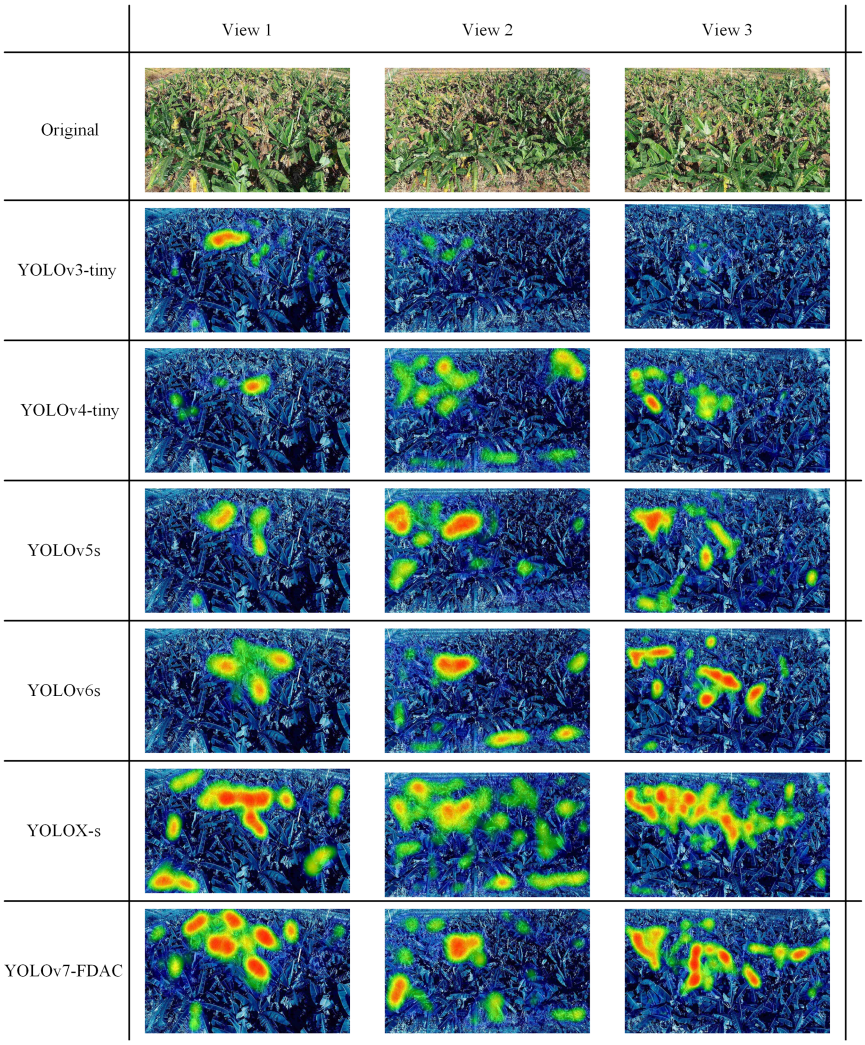


Fig. 12. Comparison of detection results for each model.

smaller model volume, which proves the validity and progress of the algorithm in this article.

To more intuitively compare the detection performance of the model in this article with other models, images from three different directions of angle in the same planting area are captured for detection. Some of the results are shown in Fig. 12. Due to poor generalization ability, YOLOv3-tiny and other models are unable to identify abnormal crops when faced with new complex scenarios. The improved YOLOv7-FDAC model greatly improves the model generalization ability, reduces the missed detection rate, and detects more targets. Although YOLOX-s can also detect abnormal banana plants relatively completely, there are more false detections in some heavily shielded areas. In summary, the detection accuracy of the improved YOLOv7-FDAC in this article is generally higher than that of other models, and the detection effect is improved on the whole, which further verifies the effectiveness of the improved model.

2) Calculation of Crop Location Based on RGB Images: In order to evaluate the reliability of the crop position solving method proposed in this article, we manually placed and recorded the position of a simulated tree and a reference for

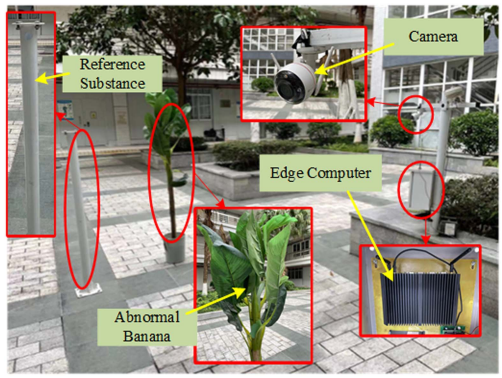


Fig. 13. Comparison of detection results for each model.

modeling the position of an abnormal banana tree in an agricultural field, which is schematically shown in Fig. 13. One image was taken at each location of 5, 10, 15, 20, 25, and 30 m from the simulated tree at an angle of 60° from the top view angle, and its coordinates were solved by the crop position mapping

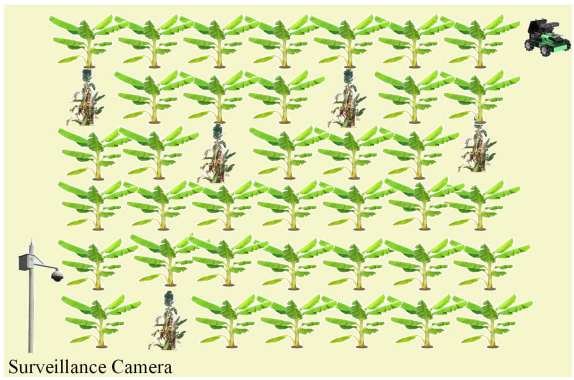


Fig. 14. Simulated robot working environment. (a) Robot working environment simulation. (b) Simplified simulation of a robot's working environment. Note: The darker the red, the higher the severity.

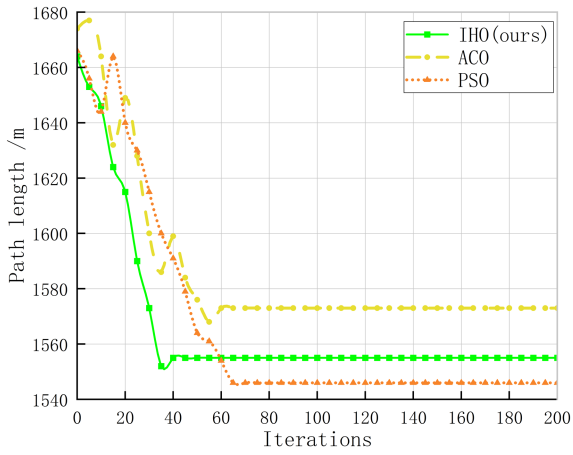


Fig. 15. Comparison of detection results for each model.

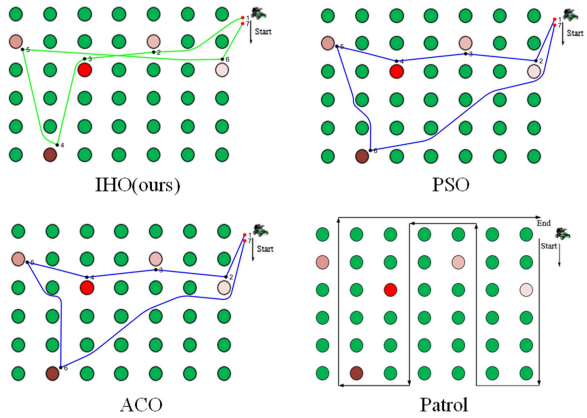
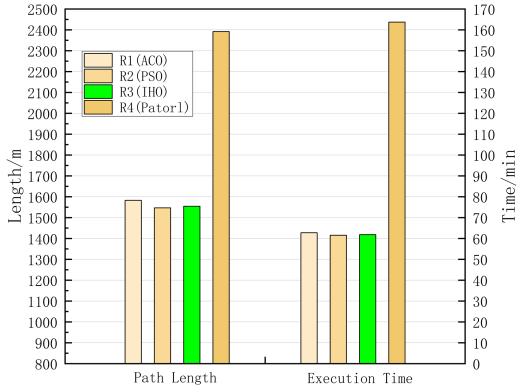


Fig. 16. Comparison of detection results for each model.



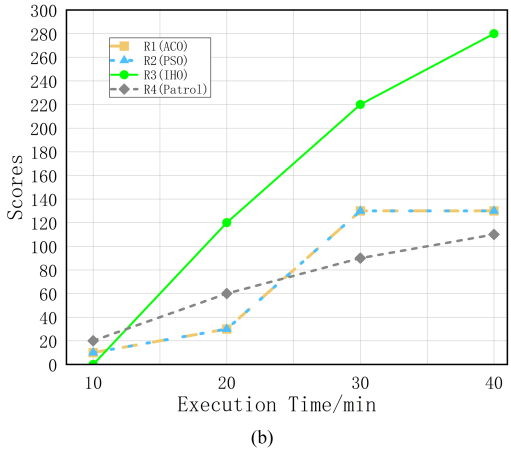


Fig. 17. Simulated robot working environment. (a) Distance and time required for robots to complete tasks using different path planning algorithms. (b) Total value of images captured by robots using different path planning algorithms.

method presented in this article and compared with the recorded position information.

The results as shown in the Table V, the position of the tree can be solved if the distance between the crop and the camera is between 5 and 25 m. Due to the large size of the tree, a clustering method is used to calculate its position, and the center point of the cluster is considered as the tree position, and the change of the center point position leads to the error. In addition, the systematic error of the camera is also partly responsible for the error. However, the robot needs to be at a certain distance from the target to capture the complete image of the target, so a reasonable error is considered acceptable. If the distance between the tree and the camera is greater than 25 m, the tree location is in the infinite focus range, and it is not possible to calculate the tree location from the image and the camera's internal reference.

3) Simulation Test of Robot Image Acquisition Based on IHO: To verify the effectiveness of the IHO proposed in this article, we

TABLE VI
EXPERIMENTAL COMPARISON OF YOLOV7-FDAC WITH OTHER MODELS

Methods	Distance Traveled	Time-consuming	Number of Images	Data Volume	Image Value	IVAD
Patrolling	2392 m	2.72h	42	94.5 MB	670	2.5
Ours	1556 m	1.03h	5	13.3 MB	300	21.93

simulated the robot working environment as shown in Fig. 14(a) with reference to the planting situation of banana plants in real environment, and randomly generated five abnormal banana trees in the environment. The simulation experiment was conducted in MATLAB 2023a, and the environment shown in Fig. 14(a) was further simplified into Fig. 14(b), where green circles indicate normal banana trees, red circles indicate abnormal banana trees, and the color depth indicates its severity, and the darker the color, the higher the severity.

We used traditional ACO and traditional particle swarm optimization (PSO) as a comparison, and conducted ten trials under the same environment and configuration to select the optimum among them. The main process is that the robot starts from the starting point, goes through all the targets to collect images, and finally returns to the starting point. The results are shown in Figs. 15 and 16, the traditional ACO and PSO are more likely to fall into the local optimum solution, converge slowly, plan paths that are not smooth enough, have redundant turning points, and increase the path length and danger of the robot. The IHO proposed in this article converges to the optimal solution after 39 iterations, which significantly improves the convergence speed, while the path is smoother.

Although the path length planned by IHO is not the shortest, it considers the urgency of pest control in agriculture and prioritizes the acquisition of images of abnormal crops. For this reason, this section proposes a method of scoring the value of the collected data per unit of time, where the higher the value of the image data collected by the robot within a certain period of time, the higher the relevance of its target, which is more consistent with the high-quality image collection method proposed in this article. As shown in Fig. 14(b), the severity of abnormal trees is classified into six levels, I–VI, based on the CADI score. The value of images corresponding to abnormal crops was scored according to Table II.

We made the three robots R1, R2, and R3 use ACO, PSO, and IHO for path planning, respectively, and made robot R4 use patrolling for image acquisition tasks. Timing was done from the start of the robots' operation, and statistics were performed every 10 min to calculate the image value score, and the results are shown in Fig. 17.

The robot performed the image acquisition task with 1.1% less path length and 1.9% less elapsed time for IHO compared to the unimproved ACO. Because the IHO first selects distant targets and high crops rather than closer crops, its average path length increases by only 0.5% and elapsed time by only 0.04% compared to the PSO. Similarly, no anomalous crop images were captured by robot R3 during the first 10 min, giving it a score of 0. However, for the 20th, 30th, and 40th min of the statistics, robot R3 obtains image value scores of 120, 220, and 280, respectively, which are much higher than those of R1 and R2. In summary, it is shown that IHO is able to find a better path solution for solving a multiobjective problem, obtain a better traversal order, and thus plan a path that is more consistent with the strategy of high-quality image acquisition.

4) Comparison of Image Efficient Acquisition Strategy With Traditional Image Acquisition Methods: To evaluate the impact of the image acquisition method proposed in this article on the image acquisition efficiency and data volume, we compare it with a traditional patrol-type image acquisition robot. The patrol-type image acquisition robot cannot determine the location of the abnormal crop, so it acquires images of the crop regardless of whether it is normal or not. The experiment was conducted in an environment based on Fig. 14, and the operation routes of both are shown in Fig. 16. The total distance traveled to perform an acquisition task, the total time spent, the amount of image data, and the IVAD score were used as evaluation metrics, where the IVAD score was obtained by calculating (5). The results are shown in Table VI.

The results showed that although the total data value score obtained by the traditional image acquisition method was as high as 670, a large number of images of normal banana trees were acquired, resulting in data redundancy, and thus the IVAD score was only 2.5. In addition, the increased amount of data also put great pressure on network transmission. Compared with the traditional acquisition method, the image acquisition method proposed in this article saves 34.9% and 62.1% in the distance traveled and time consumed in performing the acquisition task, respectively, which greatly improves the acquisition efficiency, while the IVAD is improved by 7.76 times. In summary, the present method proposed in this paper can effectively reduce data redundancy, improve data value density, relieve network pressure, and improve image acquisition efficiency.

VI. CONCLUSION

This article focuses on the problem of how to efficiently manage crops in agricultural production, and focuses on a high-quality image acquisition method for crops to ensure that crop abnormalities can be detected in time. An abnormal crop image data acquisition method based on edge intelligence and motion–static cooperation is proposed, which integrates the respective advantages of static acquisition nodes and motion acquisition nodes to optimize the image acquisition process, and at the same time proposes a heuristic path optimization algorithm based on the severity of the abnormal crop to improve the operation efficiency of the robot. Compared with the traditional method, the method exhibits superior performance in terms of the time to perform the acquisition task, the cost of image data transmission, and the value of image data, providing potential benefits for precise and real-time agricultural monitoring and

management. The study helps to improve the level of intelligent and precise management of agriculture, reduce labor cost, and is of great significance in promoting the accelerated development of intelligent agriculture.

REFERENCES

[1] D. Birkhaeuser, R. E. Evenson, and G. Feder, "The economic impact of agricultural extension: A review," *Econ. Develop. Cultural Change*, vol. 39, no. 3, pp. 607–650, 1991.

[2] F. TongKe, "Smart agriculture based on cloud computing and IoT," *J. Convergence Inf. Technol.*, vol. 8, no. 2, pp. 210–216, 2013.

[3] X. Li, B. Hou, R. Zhang, and Y. Liu, "A review of RGB image-based internet of things in smart agriculture," *IEEE Sensors J.*, vol. 23, no. 20, pp. 24107–24122, Oct. 2023.

[4] H. M. Easlon and A. J. Bloom, "Easy leaf area: Automated digital image analysis for rapid and accurate measurement of leaf area," *Appl. Plant Sci.*, vol. 2, no. 7, 2014, Art. no. 1400033.

[5] W.-z. Liang, K. R. Kirk, and J. K. Greene, "Estimation of soybean leaf area, edge, and defoliation using color image analysis," *Comput. Electron. Agriculture*, vol. 150, pp. 41–51, 2018.

[6] L. M. Romualdo et al., "Use of artificial vision techniques for diagnostic of nitrogen nutritional status in maize plants," *Comput. Electron. Agriculture*, vol. 104, pp. 63–70, 2014.

[7] A. Wang, W. Zhang, and X. Wei, "A review on weed detection using ground-based machine vision and image processing techniques," *Comput. Electron. Agriculture*, vol. 158, pp. 226–240, 2019.

[8] L. Nanni, G. Maguolo, and F. Pancino, "Insect pest image detection and recognition based on bio-inspired methods," *Ecological Inform.*, vol. 57, 2020, Art. no. 101089.

[9] T. Sritarapipat, P. Rakwatin, and T. Kasetkasem, "Automatic rice crop height measurement using a field server and digital image processing," *Sensors*, vol. 14, no. 1, pp. 900–926, 2014.

[10] W.-S. Kim, D.-H. Lee, Y.-J. Kim, T. Kim, W.-S. Lee, and C.-H. Choi, "Stereo-vision-based crop height estimation for agricultural robots," *Comput. Electron. Agriculture*, vol. 181, 2021, Art. no. 105937.

[11] Z. Wang, K. B. Walsh, and B. Verma, "On-tree mango fruit size estimation using RGB-D images," *Sensors*, vol. 17, no. 12, 2017, Art. no. 2738.

[12] S. Hua, M. Xu, Z. Xu, and H. Ye, "Kinect-based real-time acquisition algorithm of crop growth depth images," *Math. Problems Eng.*, vol. 2021, 2021, Art. no. 3913575.

[13] S. Haug and J. Ostermann, "A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks," in *Proc. Eur. Conf. Comput. Vis.*, 2015, pp. 105–116.

[14] M. A. Beck, C.-Y. Liu, C. P. Bidinosti, C. J. Henry, C. M. Godee, and M. Ajmani, "Presenting an extensive lab-and field-image dataset of crops and weeds for computer vision tasks in agriculture," 2021, *arXiv:2108.05789*.

[15] K. Wang, S. Zhang, Z. Wang, Z. Liu, and F. Yang, "Mobile smart device-based vegetable disease and insect pest recognition method," *Intell. Automat. Soft Comput.*, vol. 19, no. 3, pp. 263–273, 2013.

[16] X. Y. Feng and X. Q. Huang, "The research of the real-time high-resolution image acquisition based on embedded system," *Appl. Mechanics Mater.*, vol. 631, pp. 508–511, 2014.

[17] P. Faria, T. Nogueira, A. Ferreira, C. Carlos, and L. Rosado, "Ai-powered mobile image acquisition of vineyard insect traps with automatic quality and adequacy assessment," *Agronomy*, vol. 11, no. 4, 2021, Art. no. 731.

[18] A.-J. Garcia-Sanchez, F. Garcia-Sanchez, and J. Garcia-Haro, "Wireless sensor network deployment for integrating video-surveillance and data-monitoring in precision agriculture over distributed crops," *Comput. Electron. Agriculture*, vol. 75, no. 2, pp. 288–303, 2011.

[19] J. Zhang, A. Li, J. Li, Q. Yang, and C. Gang, "Research of real-time image acquisition system based on ARM 7 for agricultural environmental monitoring," in *Proc. Int. Conf. Remote Sens., Environ. Transp. Eng.*, 2011, pp. 6216–6220.

[20] F. Junqian, X. Deqin, and D. Xiaohui, "Agricultural field environment high-quality image remote acquisition," in *Proc. Comput. Comput. Technol. Agriculture VII*, 2014, pp. 50–60.

[21] D.-H. Kim, K.-H. Lee, C.-H. Choi, T.-H. Choi, and Y.-J. Kim, "Development of real-time onion disease monitoring system using image acquisition," *Front. Agricultural Sci. Eng.*, vol. 5, no. 4, pp. 469–474, 2018.

[22] X. Boxiang, W. Chuanyu, G. Xinyu, and W. Sheng, "Image acquisition system for agricultural context-aware computing," *Int. J. Agricultural Biol. Eng.*, vol. 7, no. 4, pp. 75–80, 2014.

[23] V. Filipović, D. Stefanović, N. Pajević, Ž. Grbović, N. Djuric, and M. Pani, "Bush detection for vision-based UGV guidance in blueberry orchards: Data set and methods," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2023, pp. 3645–3654.

[24] I. Beloev, D. Kinaneva, G. Georgiev, G. Hristov, and P. Zahariev, "Artificial intelligence-driven autonomous robot for precision agriculture," *Acta Technologica Agriculturae*, vol. 24, no. 1, pp. 48–54, 2021.

[25] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, "Machine vision systems in precision agriculture for crop farming," *J. Imag.*, vol. 5, no. 12, 2019, Art. no. 89.

[26] O. Barrero and S. A. Perdomo, "RGB and multispectral UAV image fusion for gramineae weed detection in rice fields," *Precis. Agriculture*, vol. 19, pp. 809–822, 2018.

[27] B. T. Kitano, C. C. Mendes, A. R. Geus, H. C. Oliveira, and J. R. Souza, "Corn plant counting using deep learning and UAV images," *IEEE Geosci. Remote Sens. Lett.*, early access, Aug. 8, 2019, doi: 10.1109/LGRS.2019.2930549.

[28] J. Zheng et al., "Growing status observation for oil palm trees using unmanned aerial vehicle (UAV) images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 173, pp. 95–121, 2021.

[29] H. A. M. Tran, H. Q. T. Ngo, T. P. Nguyen, and H. Nguyen, "Design of green agriculture system using Internet of Things and image processing techniques," in *Proc. IEEE 4th Int. Conf. Green Technol. Sustain. Develop.*, 2018, pp. 28–32.

[30] A. Aquino, B. Millan, M.-P. Diago, and J. Tardaguila, "Automated early yield prediction in vineyards from on-the-go image acquisition," *Comput. Electron. Agriculture*, vol. 144, pp. 26–36, 2018.

[31] S. Su, L. Fu, F. Zhang, and Y. Cui, "Image acquisition method of kiwifruit picking robot," in *Proc. Third Int. Conf. Control, Automat. Syst. Eng.*, 2013, pp. 9–12.

[32] M. Hassanein and N. El-Sheimy, "An efficient weed detection procedure using low-cost UAV imagery system for precision agriculture applications," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 181–187, 2018.

[33] S. M. Leonard, H. Xin, T. Brown-Brandl, and B. C. Ramirez, "Development and application of an image acquisition system for characterizing sow behaviors in farrowing stalls," *Comput. Electron. Agriculture*, vol. 163, 2019, Art. no. 104866.

[34] C. Zhang and J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: A review," *Precis. Agriculture*, vol. 13, pp. 693–712, 2012.

[35] G. Karthik and N. Praburam, "Detection and prevention of banana leaf diseases from banana plant using embeeded linux board," in *Proc. IEEE Online Int. Conf. Green Eng. Technol.*, 2016, pp. 1–5.

[36] M. G. Selvaraj et al., "AI-powered banana diseases and pest detection," *Plant Methods*, vol. 15, pp. 1–11, 2019.

[37] R. Sangeetha, J. Logeshwaran, J. Rocher, and J. Lloret, "An improved agro deep learning model for detection of panama wilts disease in banana leaves," *AgriEngineering*, vol. 5, no. 2, pp. 660–679, 2023.

[38] A. Aruraj, A. Alex, M. Subathra, N. J. Sairamya, S. T. George, and S. V. Ewards, "Detection and classification of diseases of banana plant using local binary pattern and support vector machine," in *Proc. IEEE Int. Conf. Signal Process. Commun.*, 2019, pp. 231–235.

[39] M. G. Selvaraj et al., "Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR congo and republic of benin," *ISPRS J. Photogrammetry Remote Sens.*, vol. 169, pp. 110–124, 2020.

[40] P. Senthilraj and P. Parameswari, "An effectual multivariate SVM integrated with CNN for identification of diseases in banana tree," *J. Pharmaceut. Negative Results*, vol. 13, pp. 1707–1719, 2022.

[41] G. Stamatescu, C. Drăgana, I. Stamatescu, L. Ichim, and D. Popescu, "IoT-enabled distributed data processing for precision agriculture," in *Proc. IEEE Mediterranean Conf. Control Automat.*, 2019, pp. 286–291.

[42] J. Wei and S. Cao, "Application of edge intelligent computing in satellite internet of things," in *Proc. IEEE Int. Conf. Smart Internet Things*, 2019, pp. 85–91.

[43] K. Bhargava, S. Ivanov, W. Donnelly, and C. Kulatunga, "Using edge analytics to improve data collection in precision dairy farming," in *Proc. IEEE Conf. Local Comput. Netw. Workshops*, 2016, pp. 137–144.

[44] C. Long, Y. Cao, T. Jiang, and Q. Zhang, "Edge computing framework for cooperative video processing in multimedia IoT systems," *IEEE Trans. Multimedia*, vol. 20, no. 5, pp. 1126–1139, May 2018.

[45] J. Ren, Y. Guo, D. Zhang, Q. Liu, and Y. Zhang, "Distributed and efficient object detection in edge computing: Challenges and solutions," *IEEE Netw.*, vol. 32, no. 6, pp. 137–143, Nov./Dec. 2018.

[46] F. M. Ribeiro, R. Prati, R. Bianchi, and C. Kamienski, "A nearest neighbors based data filter for fog computing in IoT smart agriculture," in *Proc. IEEE Int. Workshop Metrology Agriculture Forestry*, 2020, pp. 63–67.

[47] R. Zhang and X. Li, "Edge computing driven data sensing strategy in the entire crop lifecycle for smart agriculture," *Sensors*, vol. 21, no. 22, 2021, Art. no. 7502.

[48] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 7464–7475.

[49] C.-Y. Wang, H.-Y. M. Liao, and I.-H. Yeh, "Designing network design strategies through gradient path analysis," *J. Inf. Sci. Eng.*, vol. 39, no. 4, pp. 975–995, 2023.

[50] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6848–6856.

[51] Y. Wang, X. Xu, Z. Wang, R. Li, Z. Hua, and H. Song, "ShuffleNet-triplet: A lightweight re-identification network for dairy cows in natural scenes," *Comput. Electron. Agriculture*, vol. 205, 2023, Art. no. 107632.

[52] H. Li, J. Li, H. Wei, Z. Liu, Z. Zhan, and Q. Ren, "Slim-neck by GSConv: A lightweight-design for real-time detector architectures," *J. Real-Time Image Process.*, vol. 21, no. 3, 2024, Art. no. 62.

[53] D. Misra, "MISH: A self regularized non-monotonic activation function," 2019, *arXiv:1908.08681*.

[54] T. Stützle and M. Dorigo, "Aco algorithms for the traveling salesman problem," *Evol. Algorithms Eng. Comput. Sci.*, vol. 4, pp. 163–183, 1999.