Lightweight pineapple detection framework for agricultural robots via YOLO-v5sp

Jiehao Li[1,2], Chenglin Li[1], Xiwen Luo[1*], C. L. Philip Chen[2], Chenguang Yang[3]

(1. *College of Engineering, South China Agricultural University, Guangzhou 510642, China*;
2. *School of Computer Science and Engineering, South China University of Technology, Guangzhou 510641, China*;
3. *Department of Computer Science, University of Liverpool, Liverpool L693BX, United Kingdom*)

Abstract: Ensuring the accurate detection of pineapple fruits under the high planting density and serious homogenization represents a current and significant challenge. In this study, an enhanced lightweight detection framework, derived from the improved You Only Look Once version 5s (YOLOv5sp), is investigated in terms of the rapid and precise recognition of pineapple fruit for the agricultural robot. Three Convolutional Block Attention Module (CBAM) attention modules are considered the backbone network responsible for feature extraction, and the SIoU loss function is introduced to replace the CIoU loss function to handle the orientation angle and the penalization index. Eventually, the designed YOLOv5sp detection result of the mAP@0.5 value is 94.5%, which is 6.30% higher than YOLOv4, 1.83% higher than Faster R-CNN, and 6.90% higher than classical YOLOv5s. At the same time, compared with the models SHFP-YOLO and RGDP-YOLOv7-tiny in other pineapple detection literature, the mAP@0.5 of the designed model is 4.54% and 3.5% higher, respectively. Furthermore, when it comes to the agricultural robot operating in diverse natural situations, the YOLOv5sp algorithm can maintain a successful picking rate of 90% with an average time of 15 s, exhibiting the effectiveness of the visual component in engineering scenarios. These research results can accelerate the transition of pineapple harvesting from manual to automated operations.

Keywords: deep learning, target detection, lightweight networks, pineapple, YOLOv5s
DOI: 10.25165/j.ijabe.20251803.8984

Citation: Li J H, Li C L, Luo X W, Chen J L P, Yang C G. Lightweight pineapple detection framework for agricultural robots via YOLO-v5sp. Int J Agric & Biol Eng, 2025; 18(3): 204–214.

1 Introduction

Pineapple, a tropical fruit widely consumed both fresh and in processed forms such as juice, has been extensively studied for its economic significance and consumption trends[1-3]. Currently, pineapple harvesting largely depends on manual labor. However, challenges such as spiny foliage, the limited time window for optimal harvesting, and the physically strenuous nature of the task highlight the urgent need for mechanized harvesting solutions. In this context, automated localization and positioning technologies have emerged as key areas of focus for the future development of robotic harvesting systems[4,5]. Additionally, accurate detection of fruit targets is crucial for the effective operation of fruit-picking robots[6-9], as it directly affects both the success rate and overall efficiency of the harvesting process. However, pineapples are cultivated in environments with high planting density and complexity, which presents further challenges to accurate detection. Therefore, pineapple detection in robotic harvesting presents several critical challenges, including occlusion caused by branches or overlapping fruits within the robot's field of view[10], and decreased accuracy when identifying small, densely clustered fruits. These issues negatively impact detection speed and precision, hindering the ability to match the necessities of instant harvesting and practical deployment.

1.1 Prior work

Currently, the development of vision systems for fruit-picking robots primarily involves two main approaches: traditional image processing techniques and deep learning methods[11-15]. Conventional image processing algorithms typically focus on fruit characteristics such as color, shape, and texture for recognition purposes[16]. These techniques have been widely applied in agricultural engineering for an extended period. To illustrate, Zhuang et al.[17] presented a monocular vision-based method for citrus fruit detection, which combines multiple color domain features to extract candidate fruit regions, followed by the use of a support vector machine to analyze and filter texture characteristics. Similarly, Kim et al.[18] developed a detection technique that integrates various object attributes and encodes them into feature representations. In addition, Liu et al.[19] introduced a grapefruit maturity assessment approach based on an elliptical contour modeling framework. Li et al.[20] adopted texture and hierarchical contour features as the input of the integrated classifier-RUSBoost and achieved accurate detection of unripe citrus fruits under various occlusive conditions. Traditional image processing and machine learning techniques, such as color segmentation, thresholding, edge detection, and feature-based classifiers, have demonstrated effectiveness in controlled environments and for specific fruit types. These methods are relatively simple, fast, and easy to interpret. However, their generalizability is limited in unstructured agricultural environments, where challenges such as uneven lighting, occlusion, and variability in fruit appearance frequently occur. In contrast, deep learning

Received date: 2024-04-07 **Accepted date:** 2025-04-28
Biographies: **Jiehao Li**, PhD, research interest: intelligent agricultural equipment, Email: jiehao.li@ieee.org; **Chenglin Li**, Master candidate, research interest: intelligent agricultural equipment, Email: chenglin.li@ieee.org; **C. L. Philip Chen**, PhD, research interest: Intelligent Systems and Control, Email: Philip.Chen@ieee.org; **Chenguang Yang**, PhD, research interest: intelligent control, Email: cyang@ieee.org.
***Corresponding author: Xiwen Luo**, PhD, academician, research interest: intelligent agricultural equipment. The Key Laboratory of Key Technology on Agricultural Machine and Equipment, Ministry of Education, College of Engineering, South China Agricultural University, Guangzhou 510642, China. Tel: +86-10- 62737706, Email: xwluo@scau.edu.cn.

approaches exhibit greater robustness and adaptability in complex scenes by automatically learning hierarchical feature representations. Nonetheless, these advantages come at the expense of increased computational demands and the requirement for large volumes of annotated training data.

In recent years, the rapid advancement of deep learning has positioned it as the dominant approach for fruit detection[21-25]. Deep learning-based object detection algorithms can be broadly categorized into two types: one-stage and two-stage models[26]. Each approach has its distinct strengths and limitations. While two-stage models generally offer higher accuracy, they are typically slower in detection speed compared to one-stage models. In the context of fruit harvesting, it is crucial for the detection speed of the vision system to align with the robot's picking actions. Consequently, researchers are investigating deep learning algorithms[27–31] that balance both detection speed and accuracy. In terms of reducing model size, Zhu et al.[32] developed the YOLO-LM model by integrating the GSConv[33] into the Neck network to replace standard convolutions, thereby streamlining computational complexity. The optimized architecture achieved a mAP@0.5 of 93.18% with 10.17 million parameters and a model size of 19.82 MB. Yang et al.[34] developed the MFD-YOLO model, which reduced computational complexity by 28% through a lightweight backbone network and adaptive down sampling techniques. By applying a deep compression strategy, the model was streamlined to 3.58 MB, while achieving a mAP of 97.5%, outperforming the baseline model by 6.2% in accuracy. To solve the problem of small target detection, Wang et al.[35] proposed DSE-YOLO, which improved the detection accuracy of young strawberry fruit by combining point-by-point convolution and extended convolution. Li et al.[27] introduced BiFPN and dynamic label assignment strategy in pitaya detection, achieving a detection accuracy of 97.8%. Bai et al.[36] integrated Swin Transformer prediction head in YOLOv7 to build a lightweight strawberry detection model, achieving 45 FPS and 92.1% mAP. To solve the problem of uneven sample distribution and complex background interference, Ma et al.[37] used WIoU loss function to replace traditional CIoU loss function, and introduced the dynamic anchor frame quality adjustment strategy to optimize the bounding box regression performance of apple detection. However, the above detection methods generally fail to achieve a balance among recognition accuracy, real-time detection speed, and lightweight algorithm design, which are critical requirements for practical robotic harvesting systems that demand high precision, rapid response to dynamic environments, and deployment on resource-constrained embedded hardware.

1.2 Motivation and contributions

Based on the previous discussion, a YOLOv5sp lightweight network of pineapple detection for the agricultural robot is investigated in this article. The proposed model incorporates an updated Convolutional Block Attention Module (CBAM) within the Cross-Stage Partial Darknet 53 (CSP-Darknet53) layers of the classical YOLOv5s architecture, aiming to reduce irrelevant image information and enhance detection accuracy. To overcome the limitations of the Complete Intersection over Union (CIoU) loss function, particularly its inability to account for the directional alignment between predicted and ground-truth bounding boxes, which slows convergence, the Scylla Intersection over Union (SIoU) loss function is introduced. This improvement enhances angle optimization between predicted and ground-truth bounding boxes, thereby increasing the efficiency of model training and detection accuracy. In addition, in order to enhance the robustness and generalization capability of the model in challenging agricultural scenarios, image augmentation strategies were adopted to the dataset. The optimized network was effectively implemented on a resource-constrained microcontroller, demonstrating its competence in accurately and swiftly detecting pineapples in practical field environments.

The main achievements of this study are as below:

1) A high-accuracy, lightweight detection network is explored for pineapple recognition on agricultural robots. The CBAM attention module is integrated into the CSP-Darknet53 backbone network, and the SIoU loss function is utilized for enhanced detection speed.

2) Comprehensive experiments conducted on mixed pineapple datasets under various conditions, including different orientations, lighting, and backgrounds, demonstrate that the YOLOv5sp algorithm achieves outstanding performance, with an accuracy of 94.5%. This result significantly surpasses the performance of other object detection networks.

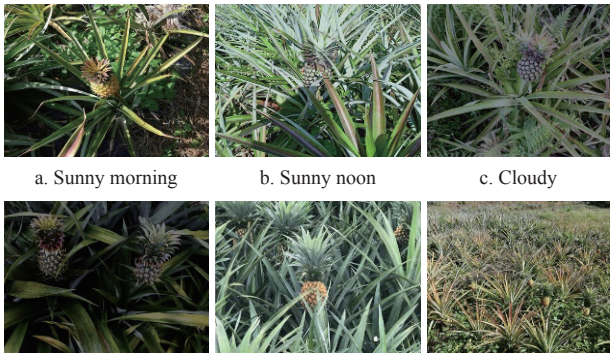
3) The YOLOv5sp model is implemented on the agricultural robot, achieving a harvest success rate of 90% and an average processing time of 15 seconds, demonstrating its exceptional performance in practical engineering applications.

The structure of the following chapters is as follows: Section 2 presents the establishment of a pineapple image dataset and the development of the pineapple detection network. Section 3 provides a comparison and analysis of the data collected from both simulations and experiments. Lastly, Section 4 concludes the paper and outlines directions for future research.

2 Materials and methods

2.1 Data acquisition

The image dataset used in this study comes from a pineapple garden located in Xuwen County, Zhanjiang City, Guangdong Province. The Intel RealSense D435i Camera serves as the image collection device, capturing images at a resolution of 1920×1080 from various angles, including 45° and 90°. In constructing the pineapple fruit dataset, environmental factors such as time of day and weather conditions during actual robotic harvesting operations were considered to enhance the model's generalization and robustness. Lighting conditions on sunny days fluctuate notably from morning to evening. As illustrated in Figure 1a, the light is relatively soft during the morning hours. Under strong sunlight conditions, the light becomes intense, creating dynamic shadow patterns, as shown in Figure 1b. In contrast, cloudy days provide relatively uniform and diffuse lighting, as depicted in Figure 1c. After sunset, light intensity decreases sharply, as shown in



a. Sunny morning b. Sunny noon c. Cloudy

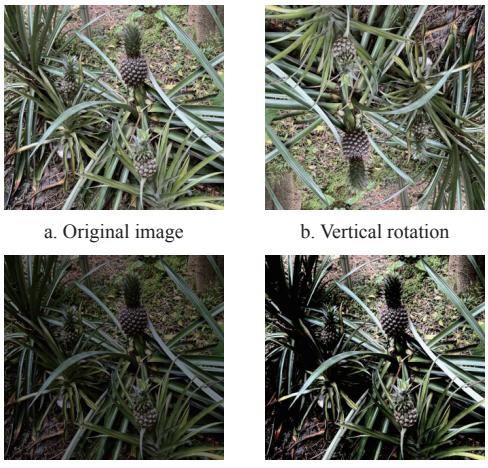
d. Evening e. Occlusion f. Small target

Figure 1 Images of pineapples captured under varying weather conditions and from multiple viewing angles

Figure 1d. In robotic vision scenarios, pineapple fruits are often heavily occluded by surrounding branches and leaves, as illustrated in Figure 1e. Additionally, when pineapples are viewed from a distance, the target appears small, as shown in Figure 1f. The dataset composition comprehensively accounts for a wide range of environmental interference factors, ensuring the robot can perform reliably across diverse natural conditions.

2.2 Image augmentation

Deep learning network for target recognition necessitates large datasets to strengthen the algorithm's ability to generalize and improve its robustness. Bailly et al.[38] found that a more comprehensive dataset can enhance the performance of the deep learning algorithm. Nevertheless, the datasets used in this study only comprise 1081 pineapple images. To solve the problem of a limited dataset, this work employs image enhancement methods such as geometric alterations, color space enhancement, filtering, and randomization. In addition, incorporating larger-scale training datasets has the ability to enhance the model's generalization ability and help mitigate the risk of overfitting[39]. Image enhancement is about improving the visual details of an image to detect objects[40] accurately. To augment the generalization and robustness of the algorithm in a complex agricultural background, this study employs the ImgAug library to augment the dataset. ImgAug provides various augmentation techniques, including the addition of random pixels, Gaussian noise, random rectangular occlusion, pixel dropout, motion blur, adaptive histogram equalization, horizontal and vertical flipping, HSV transformation, random brightness adjustment, and average pooling. It supports both single-image and batch processing, and automatically updates the corresponding label files with each enhancement. In this study, data augmentation was performed using image rotation, brightness adjustment, and adaptive histogram equalization. As shown in Figure 2, Figure 2a is the original image, Figure 2b is the image after vertical rotation, Figure 2c is the image after transform brightness, and Figure 2d is the image after histogram equalization. After image enhancement, the pineapple image dataset increased to 5190, as shown in Table 1. The dataset was randomly divided into training, validation, and test sets in an 8:2:1 ratio. Pineapple fruits were annotated using the open-source tool LabelImg, with the annotations saved in txt format and the label assigned as "pineapple".



a. Original image b. Vertical rotation

c. Brightness transformation d. Histogram equalization

Figure 2 Image enhancement

2.3 YOLOv5sp networks development

YOLOv5, developed by Glenn Jocher, achieves real-time detection speeds of up to 140 frames per second (FPS) while maintaining high detection accuracy and a compact model size[41]. These characteristics make it particularly well-suited for deployment in pineapple harvesting robots operating in real-world agricultural environments. The official versions of YOLOv5 include YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, all of which share the same underlying architecture. The primary differences among them lie in the depth and width multipliers specified in the "YAML" configuration files.

Table 1 Number of images produced through data enhancement approaches

Raw image	Equalization	Transform brightness	Image rotation	Total
1038	1038	1038	2076	5190

Given the need for efficient and lightweight models in complex orchard environments with uneven lighting and variable weather conditions, this study selects and further optimizes the YOLOv5s variant by proposing an improved network architecture. The enhancements include a backbone network with stronger feature extraction capabilities, an optimized neck for efficient feature aggregation, and a detection head incorporating a new loss function. Specifically, the backbone is based on an improved CSP-Darknet53[42], where the Convolutional Block Attention Module (CBAM) is embedded after each CSP Bottleneck with three convolutions (C3). This addition strengthens the network's ability to focus on pineapple-relevant features, assigning higher weights to semantically important regions and thus improving detection precision. The extracted features are output as Feature 1, Feature 2, and Feature 3 and forwarded to the neck for further processing. Moreover, the SIoU loss function replaces the CIoU loss function, enhancing bounding box regression accuracy by aligning predicted boxes more closely with ground truth and accelerating model convergence. The complete network architecture is illustrated in Figure 3.

2.3.1 CBAM attention module

Traditional detection models may have some limitations in obtaining and representing pineapple fruit characteristics. For example, they may not be effective at establishing relationships between channels or perform poorly when dealing with complex target structures and background noise. Therefore this study introduced the CBAM attention mechanism to improve this situation. CBAM is mainly composed of a feed-forward CNN with a one-way multi-layer structure, making it simple and efficient. It contains two key components: channel attention (CA) and spatial attention (SA). The relationship between the different channels on the feature map is established by CA, which can adjust their contributions by calculating the weight of each channel. This helps to highlight the target-related channels and suppress the impact of irrelevant information. SA is used to establish the relationship between different spatial locations in the feature map to extract more effective features. By combining CA and SA, the CBAM attention mechanism can adaptively adjust the weights of each channel and spatial position in the feature map to better capture the characteristics of pineapple fruit. Figure 4 describes the working process of this module. Initially, the module infers a one-dimensional CA map F_c through the channel sub-module and multiplies it item by item with the feature map F to obtain F_1. Next, F_1 is utilized as the input of the SA sub-module to infer the two-dimensional SA map F_s, multiplied by F_1 item by item, to obtain an adaptive refinement feature map F_2. Figure 5 describes the calculation process of the CA sub-module. The details of the two

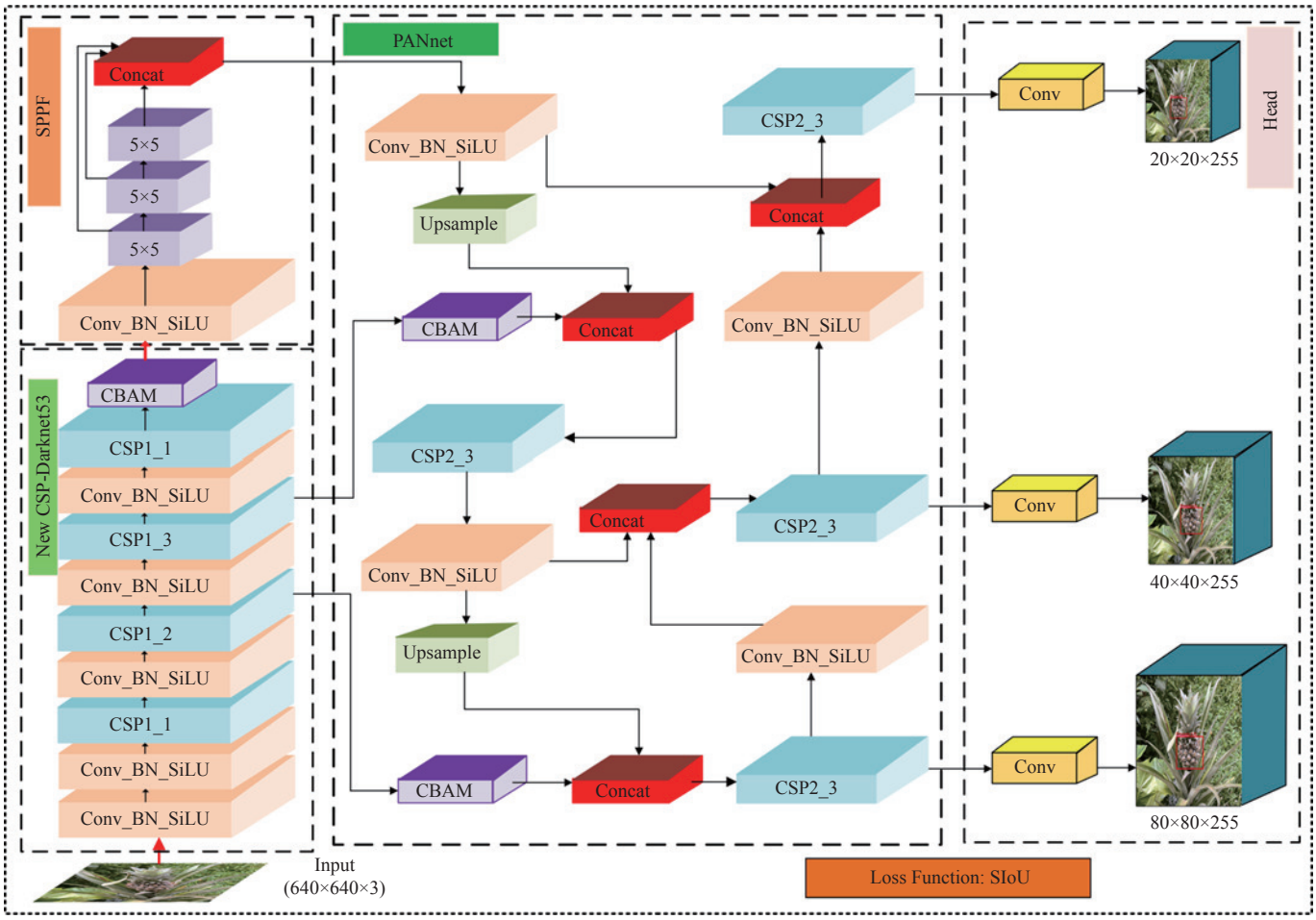


Figure 3 Network model structure of YOLOv5sp

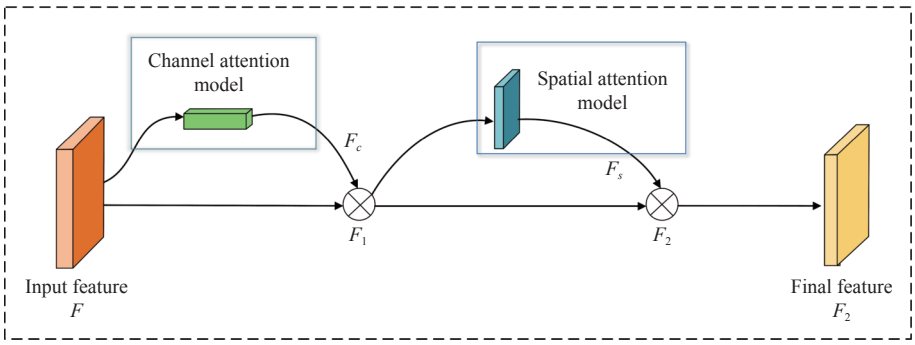


Figure 4 View of CBAM

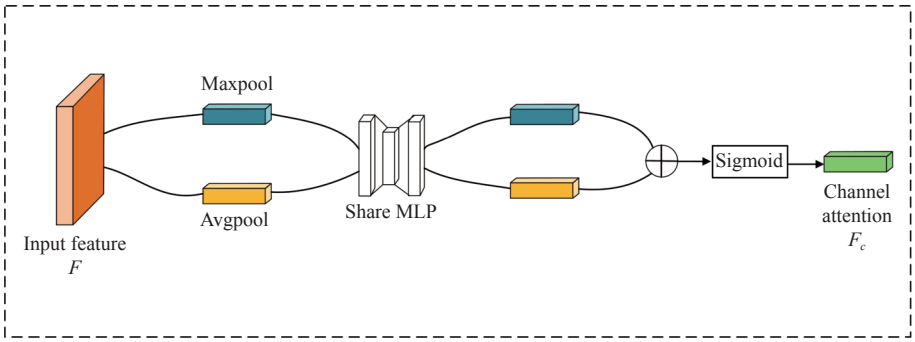


Figure 5 Schematic diagram of CA sub-module

sub-modules are described below.

The intermediate feature map F of an image is chosen as the CBAM input feature map. First, the CA sub-module of CBAM performs maximum pooling and average pooling on the F, respectively, which are to congregate the spatial information and output two various description symbols F_{c1} and F_{c2}. Then they are sent to a shared network comprised of multi-layer perceptions (MLP). Finally, the feature vectors are combined using the item-by-item summation approach to produce a CA map F_c. F_1 is obtained by element-by-element multiplication of F and F_c. Equations (1)-

(2) depict the CA calculation procedure.

$$F_c = \beta(\mathrm{MLP}(AvgPool(F)) + \mathrm{MLP}(MaxPool(F))) \quad (1)$$

$$F_1 = F_c \times F \quad (2)$$

The spatial attention (SA) sub-module mainly generates a two-dimensional SA map based on the spatial correlation of features. It initially performs average and maximum pooling manipulations along the direction of the channel to generate two 2D maps: F_{s1} and F_{s2}. They represent the average and maximum pool features through the channel, respectively. Then the convolution layer joins and convolves these two features to generate a two-dimensional SA graph F_s. Finally, the final refined output F_2 can be obtained by multiplying F_s and F_1 element by element. Figure 6 describes the calculation process of the SA sub-module. The calculation of SA is shown in Equations (3)-(4).

$$F_s = \beta(C(\mathrm{MLP}(AvgPool(F_1)) + \mathrm{MLP}(MaxPool(F_1)))) \quad (3)$$

$$F_2 = F_s \times F_1 \quad (4)$$

2.3.2 Improvement of loss function

An accurate loss function is a fundamental component for evaluating and optimizing the performance of a detection model. In the original YOLOv5 algorithm, the position loss function is implemented using the CIoU loss function, which examines differences in length and width, implying that a prediction box closer to the ground truth corresponds to higher accuracy. However, the aspect ratio represents a more ambiguous metric, and CIoU does not address the varying challenges in detecting pineapple fruits. For the first time, the SIoU function, proposed by Zhora[43], considers the direction of the mismatch between the prediction box and the ground truth by redefining the penalty mechanism. Incorporating the vector angle between regressions enhances detection precision, as illustrated in Figure 7, with Equations (5) and (6) outlining the calculation formulas.

$$L_{\mathrm{SIoU}} = 1 - \mathrm{IoU} + \frac{\Delta + \Omega}{2} \quad (5)$$

$$\mathrm{IoU} = \frac{|B \cap B^{\mathrm{GT}}|}{|B \cup B^{\mathrm{GT}}|} \quad (6)$$

where, Δ denotes the distance-based loss component, while Ω signifies the loss associated with shape discrepancies. To reduce the number of uncertain variables associated with distance, this study introduced an angle-aware LF component into the loss function. Figure 7 illustrates the scheme for calculating the angle cost contribution, and Equations (7)-(10) define the angle cost formula in detail.

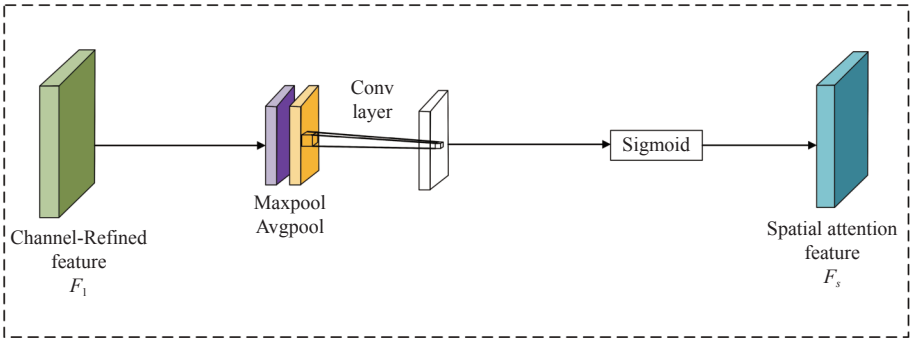


Figure 6 Schematic diagram of SA sub-module

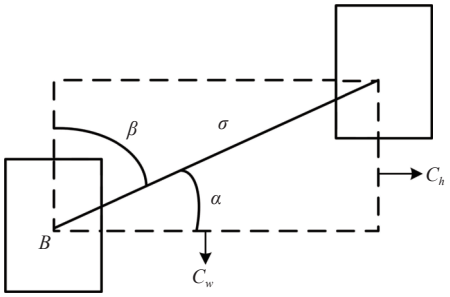


Figure 7 Schematic diagram of angle loss calculation

$$\Lambda = 1 - 2 \times \sin^2\left(\arctan(x) - \frac{\pi}{4}\right) \quad (7)$$

$$x = \frac{c_h}{\varepsilon} = \sin(\alpha) \quad (8)$$

$$\varepsilon = \sqrt{\left(b_{c_x}^{gt} - b_{c_x}\right)^2 + \left(b_{c_y}^{gt} - b_{c_y}\right)^2} \quad (9)$$

$$c_h = \max\left(b_{c_y}^{gt}, b_{c_y}\right) - \min\left(b_{c_y}^{gt}, b_{c_y}\right) \pm \quad (10)$$

where, $b_{c_x}^{gt}$ and $b_{c_y}^{gt}$ denote the x and y coordinates of the center point of the ground-truth bounding box, respectively. b_{c_x} and b_{c_y} represent the x and y coordinates of the center point of the predicted bounding box. The parameter ε indicates the minimum distance between the center points of the predicted and ground-truth boxes. c_h refers to the vertical distance, or height difference, between the center of the ground-truth box and that of the predicted box. Figure 8 illustrates the scheme for calculating the distance cost contribution, and Equations (11)-(14) depict the distance cost definition formula.

$$\Delta = \sum_{t=xy} (1 - e^{-\gamma\rho t}) \quad (11)$$

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w}\right)^2 \quad (12)$$

$$\rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h}\right)^2 \quad (13)$$

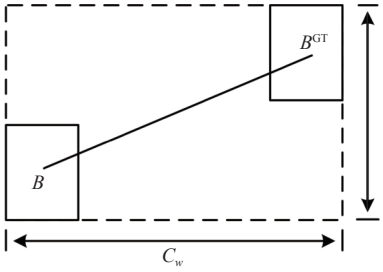


Figure 8 Schematic diagram of distance loss calculation

$$\gamma = 2 - \Lambda \tag{14}$$

where α approaches 0, the influence of the distance cost on the results becomes negligible. On the other hand, when α is closer to $\frac{\pi}{4}$, the effect of distance cost on the result is becoming significant. However, as the angle increases, the problem becomes more serious. As a result, when the angle grows, γ is a distance value given priority in time. The shape cost function is defined in Equations (15)-(17).

$$\delta = \sum_{t=h,w} (1 - e^{-w_t})^{\eta} \tag{15}$$

$$\omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \tag{16}$$

$$\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \tag{17}$$

where, h is used to denote the height of the prediction box, and w is used to represent the width of the prediction box. The model's attention to shape loss is defined by η, which is unique for each dataset.

By incorporating the direction problem between the prediction box and the authentic box in the loss function, SIoU significantly decreases the degree of freedom. Compared with the CIoU function, it can achieve faster convergence during the training phase and better inference performance. It is effective for improving the precision of the algorithm for real-time detection of pineapple fruit and reducing the time required for detection.

2.4 Model evaluation

In this study, five evaluation metrics are utilized to comprehensively assess the real-time detection performance of the enhanced pineapple fruit detection model. The first metric, precision (P), represents the proportion of correct positive predictions among all positive predictions. The second metric, recall (R), denotes the proportion of correctly predicted pineapple fruit instances relative to all actual pineapple fruit instances in the dataset. The third evaluation metric, average precision (AP), quantifies the area under the precision-recall (P-R) curve specifically for the pineapple category. The fourth metric, F1 score, represents the harmonic mean of precision and recall, providing a comprehensive measure that balances both aspects. The fifth and final metric, frames per second (FPS), indicates the model's processing speed by measuring how many images it can analyze per second. A model is generally considered capable of real-time detection when its FPS exceeds 30; otherwise, detection latency may result in information loss. These metrics comprehensively evaluate the network's detection performance and are calculated as shown in Equations (18)-(22).

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \tag{18}$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \tag{19}$$

$$\text{AP} = \int_0^1 P(R)\,\mathrm{d}R \tag{20}$$

$$\text{F1} = \frac{2P \times R}{P + R} \tag{21}$$

$$\text{FPS} = \frac{1}{\bar{y}} \tag{22}$$

where, TP (True Positive) refers to the accurate prediction of labeled pineapple fruits, while FP (False Positive) represents the number of images incorrectly predicted as containing pineapples. FN (False Negative) indicates the number of pineapple targets that were not detected, and TN (True Negative) refers to instances where pineapples are not present in the image and were correctly noted as not being present. The symbol $\bar{y}$ denotes the average detection time per image. Table 2 depicts the confusion matrix for categorization outcomes.

Table 2 Confusion matrix for categorization outcomes

Authentic	Predicted	Confusion Matrix
Negative	Positive	False Positive (FP)
Positive	Positive	True Positive (TP)
Negative	Negative	True Negative (TN)
Positive	Negative	False Negative (FN)

3 Results

3.1 Network training

The study is conducted on a Dellg3 3579 laptop. It is mainly equipped with an Intel (R) Core (TM) i5-8300H CPU with a central frequency of 2.30 GHz and a 4 GB GTX 1050 graphics card. Under the Windows 10 system, Anaconda3 is installed to build the PyTorch deep learning framework. During model training, libraries such as Opencv4.6, CUDA11.3, and Tensorboard2.10 are called. In addition, Visual Studio (VS) 2022 and Pycharm2022 are used in the experiment. To ensure the reliability of the experiment, the following experiments are carried out with the same training parameters and hyperparameters. This study set the batch size to 8, weight decay to 0.0005, momentum value to 0.937, learning rate to 0.01, and training epoch to 300.

3.2 Comparison experiments of different attention mechanisms in the backbone network

To enhance the capacity of YOLOv5s to obtain pineapple feature information, a CBAM attention module is inserted after the C3 module. In this experiment, the Squeeze-and-Excitation network (SEnet), CBAM, Coordinate Attention network (CAnet), and Efficient Channel Attention network (ECAnet) are added to the same position in the backbone network of YOLOv5s for comparative experiments.

According to Table 3, the number of Floating-point Operations Per Second (FLOPs) and parameters increased slightly after adding the attention mechanism, but the detection performance was enhanced to some degree. Among them, the CBAM attention mechanism had the most significant improvement on model performance, with mAP@0.5 value maintained at 91.28%, which was 4.35% higher than before. By comparison, adding SEnet, ECAnet, and CAnet modules resulted in 1.40%, 1.90%, and 2.39% increments of mAP@0.5 values, respectively. However, their FLOPs and Params were only slightly different. Obviously, CBAM performs better than other attention mechanisms. The experimental finding demonstrates that adding a CBAM attention module to the backbone network can significantly enhance the attention degree of the detection model to pineapple fruit and inhibit the interference of a sophisticated agricultural environment background, which can significantly increase the model's detection performance for pineapple fruit.

3.3 The influence of data expansion method

In this section, image enhancement techniques such as image rotation, brightness adjustment, and adaptive histogram equalization are employed to augment the dataset. A controlled variable method

is used to explore the impact of each enhancement approach on the performance of the YOLOv5s model. Initially, the model is trained using the fully enhanced dataset. Subsequently, each enhancement method is individually removed during training to evaluate its effect on precision and F1 score. The experiment also investigates the influence of multi-angle image acquisition, which simulates the varied perspectives encountered by the picking robot during operation. After data augmentation and multi-angle photography, the dataset size increased to 5190 images.

Table 3 Result comparison of performance across various attention mechanisms integrated into the backbone network

Model	Recall	Precision	mAP@0.5	FLOPs/G	Params/M
YOLOv5s	85.64%	87.37%	88.43%	15.94	6.69
Within CBAM	89.15%	91.43%	91.28%	16.71	6.75
Within SEnet	86.56%	88.49%	89.67%	16.63	7.21
Within ECAnet	87.61%	89.18%	90.11%	16.62	7.15
Within CAnet	88.04%	90.13%	89.97%	16.78	6.92

The experimental results, as summarized in Table 4, indicate that multi-angle image acquisition significantly improves model performance, as its removal causes the F1 score to decrease by 1.36% and the precision by 0.85%. Removing adaptive histogram equalization leads to a 1.48% reduction in precision, suggesting its importance in enhancing image contrast and detection accuracy. Similarly, brightness adjustment simulates illumination variability in natural environments and effectively improves model robustness and generalization, as its removal results in a decrease of 1.67% in F1 score and 1.06% in precision. In contrast, the vertical rotation method contributes relatively little to performance improvement, with only a 0.32% decrease in accuracy and a 0.94% drop in F1 score observed when it is excluded. These findings demonstrate the effectiveness of specific image enhancement techniques and acquisition strategies in improving the robustness and accuracy of the pineapple fruit detection model in complex field background.

Table 4 F1 score and mAP@0.5 value obtained by different image enhancement methods

Data enhancement method	F1 score	mAP@0.5
All dataset	95.7%	94.5%
Remove multi-angle viewing method	94.4%	93.7%
Remove vertical rotation method	94.8%	94.2%
Remove brightness transformation method	94.1%	93.4%
Remove adaptive histogram equalization method	93.7%	93.1%

3.4 Ablation experiments

To evaluate the effectiveness of the proposed pineapple detection algorithm, a series of ablation experiments were conducted to examine the individual contributions of the CBAM attention module, the SIoU loss function, and the image augmentation techniques. As listed in Table 5, the various enhancements to the model structure and algorithm strategy yield significant improvements. Specifically, the mAP@0.5 for YOLOv5sp increased by 6.9% compared to the original YOLOv5s, while the FLOPs and Params saw relatively modest increases of 4.83% and 0.9%, respectively. Incorporating an attention mechanism strengthens the network's feature extraction capabilities, enabling it to effectively reduce interference and enhance the model's detection accuracy in challenging environments. The mAP@0.5 improved by 4.35%, while FLOPs and Params increased by 4.83% and 0.90%, respectively. Simultaneously, by refining the loss function to account for the vector angle in the expected regression and redefining the penalty criteria, the overall degree of freedom in the loss is significantly reduced. This adjustment facilitates both the converging process and the detection accuracy during training. Without affecting the complexity of the model and the amount of computation, the mAP@0.5 further improves by 0.50%. Furthermore, the application of image enhancement techniques helps to expand the dataset, boosting the robustness of the detection model and leading to a further increase in pineapple detection accuracy by 1.93%. As demonstrated in Table 5, the mAP@0.5 of the network shows a gradual improvement as a result of the enhancements implemented using the aforementioned methods.

Table 5 Ablation experiments results

Tag	Basic Model	add CBAM	add SIoU	add Image enhancement	mAP@0.5	Flops/G	Params/M
1	✓	-	-	-	88.43%	15.94	6.69
2	✓	✓	-	-	92.28%	16.71	6.75
3	✓	✓	✓	-	92.74%	16.71	6.75
4	✓	✓	✓	✓	94.53%	16.71	6.75

3.5 Comparison of different algorithms

The whole pineapple dataset is utilized for training the model and assessing the performance of the improved YOLOv5s. Meanwhile, the proposed YOLOv5sp model is compared against YOLOv4, YOLOv5s, and Faster R-CNN to demonstrate its superiority. Table 6 shows that when combined with AP value, FPS, FLOPs, and Params, YOLOv5sp performs most suitably. The AP value of the YOLOv5sp model is 6.30% and 6.90% higher than those of YOLOv4 and YOLOv5s, respectively, when compared to the one-stage object detection algorithm. Compared with YOLOv4, the FLOPs of the YOLOv5sp model decreased by 74.07%. The parameters of the YOLOv5sp model are 88.76% lower than YOLOv4. The parameters of YOLOv5s are lower than those of YOLOv5sp, but considering the AP value and the performance of FLOPs, YOLOv5sp is still optimal. This shows that YOLOv5sp performs best in the one-stage object detection algorithm.

Table 6 Comparison of experimental results of different models

Model	YOLOv4	Faster R-CNN	SHFP-YOLO[44]	RGDP-YOLOv7-tiny[45]	YOLOv5s	YOLOv5sp
mAP@0.5/%	88.9	92.8	90.4	91.3	88.4	94.5
FPS	26.7	2.4	38.7	47.4	41.6	40.2
FLOPs/G	64.4	370.2	6.3	4.5	15.9	16.7
Params/M	60.5	137.1	2.7	2.3	6.7	6.8

Faster R-CNN has excellent detection performance in the two-stage object detection algorithm. Its FLOPs and Params are much higher than the one-stage detection model, so it has good detection performance. However, its AP value is still 1.80% lower than that of YOLOv5sp. In addition, compared with the SHFP-YOLO[44] and RGDP-YOLOv7-tiny[45] models reported in previous pineapple detection studies, the proposed model achieves higher mAP@0.5 values, exceeding them by 4.54% and 3.50%, respectively. Although its detection speed is slightly lower than that of RGDP-YOLOv7-tiny, the proposed model demonstrates superior detection accuracy. In summary, the improved model is more suitable for the pineapple-picking robot to identify pineapples. Figure 9 depicts the loss values of six models during training. It is evident that the loss value of the YOLOv5sp model during training decreases faster and

is lower than that of other detection models. Figure 10 illustrates the mAP@0.5 value curve of the six models during training. The mAP@0.5 values of all models rise rapidly in the beginning epochs and slow down gradually with the increase of epochs. The proposed YOLOv5sp model finally obtained 94.5% AP values, which exceeded other models.

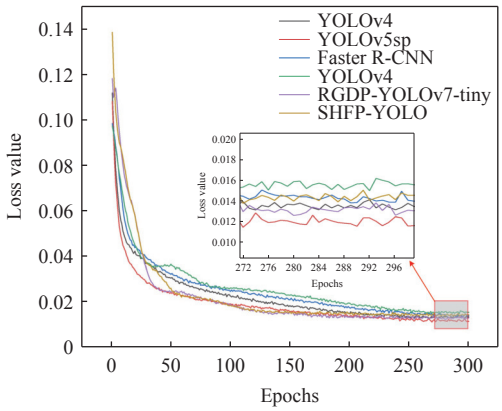


Figure 9 Loss value curves of six detection models

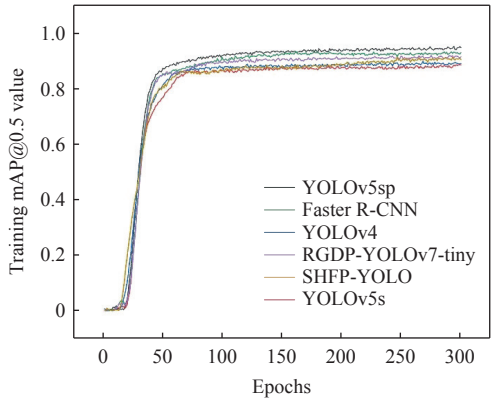


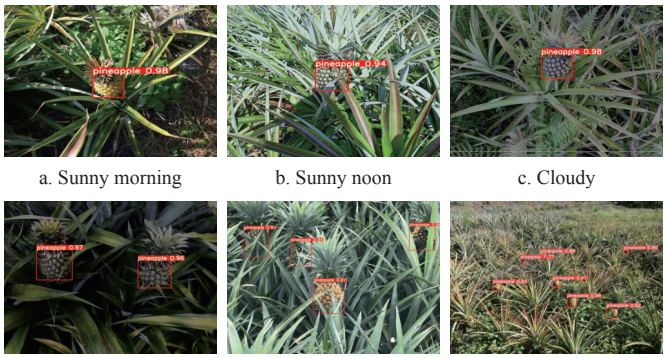
Figure 10 Training mAP@0.5 value curve of the six models

3.5.1 Model detection performance analysis

The growth environment of pineapple is complex. There are many problems, such as the fruit being blocked by branches and leaves, overlapping between fruits, too small pineapple fruit, and so on. The YOLOv5sp adds the CBAM attention modules to enhance the capacity to obtain features of pineapple. Figures 11 and 12 show that even when the fruit is occluded, overlapped, or the fruit target is too tiny, YOLOv5sp can detect the occluded pineapple, while the original YOLOv5s has missed detection. By comparing e. and f. of Figure 13 and Figure 14, it can be seen that YOLOv4 and Faster R-CNN failed to detect pineapple when the pineapples were obscured or the pineapple targets were too small. YOLOv5sp pineapple detection under different light conditions and angles is the most accurate and comprehensive. As a result, compared with other algorithms, the detection performance of YOLOv5sp for occluded or overlapped pineapples has been optimized to a certain extent, and real-time detection of pineapples in complex agricultural environments can be realized.

3.6 Model visualization analysis

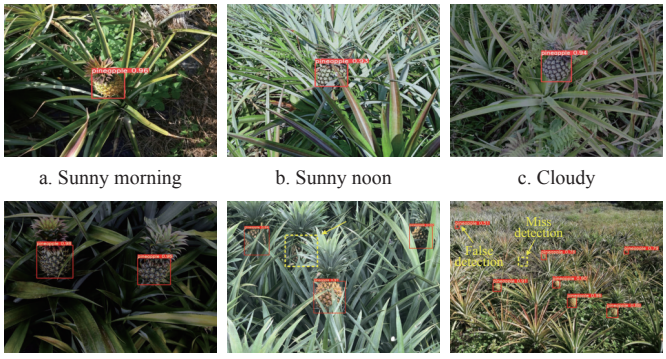
Convolutional neural networks only deal with object detection problems, but their interpretability in network processing is not very good. Therefore, in this study, the visualization effects of YOLOv5s and the improved YOLOv5sp model are compared with activation heat maps, and the features extracted from the final convolution layer are visualized to facilitate the detection of pineapple fruit. In the heat map, the intensity of the red area indicates the level of influence that specific location has on the model's final decision, with darker red areas representing greater influence. Since YOLOv5sp adds a CBAM attention mechanism, the model pays more attention to the pineapple fruit association region, which can effectively suppress the interference of the environment on its decision-making, as shown in Figure 15e and Figure 16e. Moreover, compared with the visualization results in Figures 15 and 16, the feature extraction capability of YOLOv5sp is typically superior to that of the original YOLOv5s across various conditions. Therefore, the proposed lightweight pineapple inspection model, YOLOv5sp, is better suited for integration into the vision module of the pineapple-picking robot for effective pineapple inspection.



a. Sunny morning b. Sunny noon c. Cloudy

d. Evening e. Occlusion f. Small target

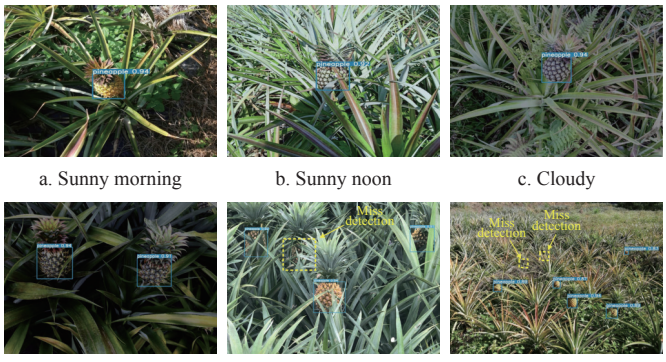
Figure 11 Pineapple detection results of YOLOv5sp under different conditions



a. Sunny morning b. Sunny noon c. Cloudy

d. Evening e. Occlusion f. Small target

Figure 12 Pineapple detection results of YOLOv5s under different conditions



a. Sunny morning b. Sunny noon c. Cloudy

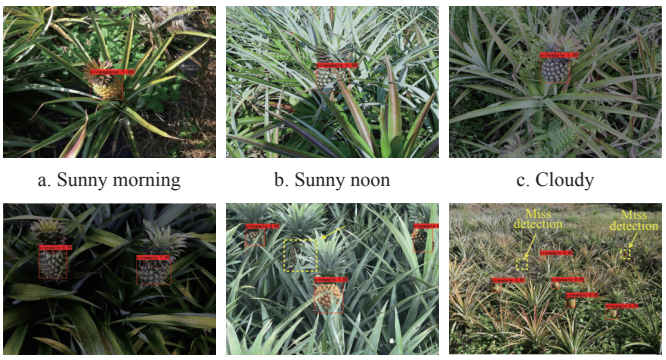
d. Evening e. Occlusion f. Small target

Figure 13 Pineapple detection results of YOLOv4 under different conditions

3.7 Real-world experimental demonstrations

In the laboratory environment, a picking experiment simulating

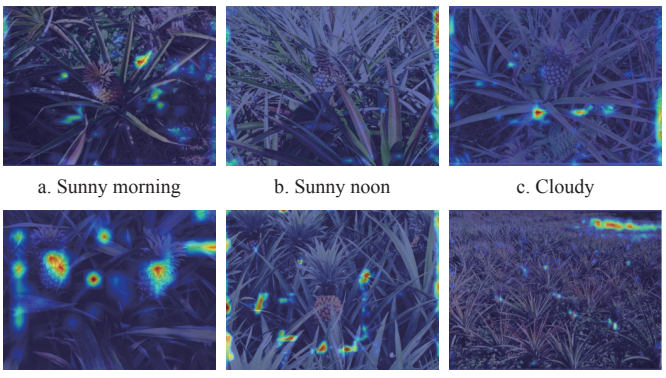
a pineapple-picking robot is carried out. This experiment aims to explore the precision of the proposed YOLOv5s algorithm in the actual picking process and to supplement the database simulation results. The experimental platform is composed of Intel RealSense Depth Camera D435i, 6-DOF Unitree manipulator, and end effector, as shown in Figure 17. The experimental steps are as follows: 1) Pineapple plants are placed at a distance of 60 cm from the platform. The camera collects pineapple images in real time and performs real-time detection through the YOLOv5sp algorithm to obtain the specific location of the pineapple. 2) The information about the pineapple's position and posture is transmitted to the robotic arm through the ROS system, the manipulator reaches the pineapple position autonomously through the path-planning algorithm, and robust adaptive control[46] and fuzzy sliding mode control[47] are used to complete the picking action, as shown in Figure 18.



a. Sunny morning b. Sunny noon c. Cloudy

d. Evening e. Occlusion f. Small target

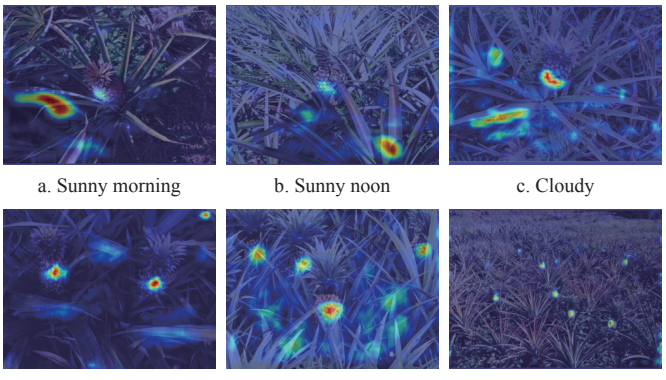
Figure 14 Pineapple detection results of Faster R-CNN under different conditions



a. Sunny morning b. Sunny noon c. Cloudy

d. Evening e. Occlusion f. Small target

Figure 15 YOLOv5s heat map



a. Sunny morning b. Sunny noon c. Cloudy

d. Evening e. Occlusion f. Small target

Figure 16 YOLOv5sp heat map

In the case of hilly areas, it is difficult for conventional crawler chassis machinery to operate in these areas[48]. Therefore, a demonstration using the Unitree robot attempts to evaluate the picking performance of two pineapples in the laboratory environment, as shown in Figure 19. Experiments demonstrate that by incorporating the CBAM attention module into the YOLOv5s model and substituting the CIoU function in the detection layer with the SIoU loss function, the achievement rate of pineapple fruit harvesting reaches 90%, and the average time consumption is about 15 s. In particular, when branches and leaves block the pineapple, the position of the pineapple fruit can also be well obtained. As a result, in robot engineering applications, the designed algorithm of YOLOv5s has achieved advanced performance with high-accuracy detection and strong anti-interference.

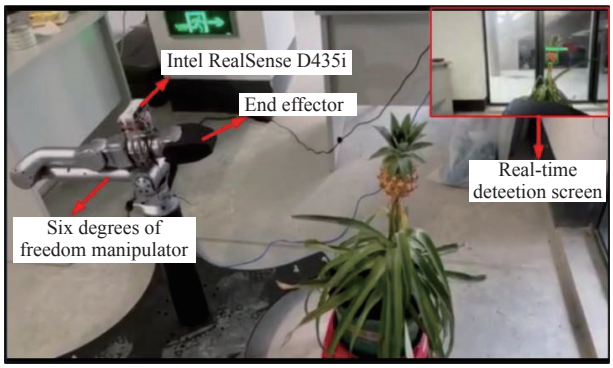


Figure 17 Pineapple picking experimental platform

Figure 18 Demonstration of pineapple picking process in robot system

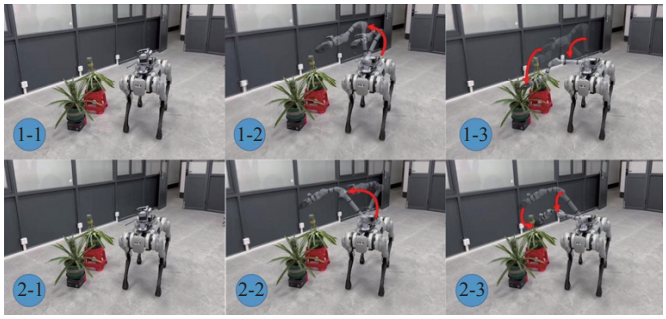


Figure 19 Pineapple picking process using the Unitree robot with manipulator's arm

4 Conclusion

This study proposes a real-time pineapple fruit detection approach based on an improved YOLOv5s model, designed to operate effectively in complex agricultural environments. The model integrates a CBAM after the C3 module to enhance feature extraction capabilities, particularly under conditions involving

occlusion, small targets, and dense fruit clusters, thereby improving detection accuracy. Furthermore, the CIoU loss function in the detection layer is replaced by the SIoU loss function to accelerate model convergence and enhance detection accuracy. The key findings are as follows:

1) Ablation studies indicate that the improved YOLOv5s model achieves excellent performance. Under the premise of increasing a small amount of calculation and parameters, the model's detection mAP@0.5 is increased to 94.5%, which is significantly higher than the original model (mAP@0.5=88.4%).

2) Evaluation on the custom pineapple fruit dataset shows that the YOLOv5sp model attains an mAP@0.5 of 94.5%, representing improvements of 6.30%, 1.83%, 4.54%, 3.50%, and 6.90% over YOLOv4, YOLOv5, Faster R-CNN, SHFP-YOLO, and RGDP-YOLOv7-tiny, respectively.

3) Laboratory-based picking experiments reveal a success rate of 90% and an average picking time of approximately 15 s, validating the practical effectiveness of the model.

In conclusion, the proposed YOLOv5sp model effectively meets the engineering requirements for robotic pineapple harvesting. However, this study focuses solely on fruit detection and does not address the classification of fruit maturity levels (e.g., immature, green-ripe, or fully ripe). Future research will aim to expand the model to include maturity level classification while also enhancing both detection accuracy and processing speed.

Acknowledgements

This work was supported by the 2024 Basic and Applied Research Project of Guangzhou Science and Technology Plan (Grant No. 2024A04J4140), the National Key Laboratory of Agricultural Equipment Technology Project (Grant No. NSKLAET-202408), the Key Laboratory of Spectroscopy Sensing Ministry of Agriculture and Rural Affairs (Grant No. 2024ZJUGP004), and the Young Talent Support Project of Guangzhou Association for Science and Technology (Grant No. QT2024-006).

[References]

[1] Hossain M. World pineapple production: An overview. African Journal of Food, Agriculture, Nutrition and Development, 2016; 16(4): 11443–11456.

[2] Li H F, Li G B, Yang B B, Chen G Q, Lin L, Yu Y Z. Depthwise nonlocal module for fast salient object detection using a single thread. IEEE Transactions on Cybernetics, 2020; 51(12): 6188–6199.

[3] Chen H, Li Y F, Su D. Discriminative cross-modal transfer learning and densely cross-level feedback fusion for RGB-D salient object detection. IEEE Transactions on Cybernetics, 2019; 50(11): 4808–4820.

[4] Zhang L Z, Tang S, Li P, Cui S, Guo H L, Wang F C. Structure design of a semi-automatic pineapple picking machine. IOP Conference Series: Materials Science and Engineering, 2018; 452: 042155.

[5] Lan R S, Sun L, Liu Z B, Lu H M, Pang C, Luo X N. Madnet: A fast and lightweight network for single-image super resolution. IEEE Transactions on Cybernetics, 2021; 51(3): 1443–1453.

[6] Tang M, Liao H C, Herrera-Viedma E, Chen C L P, Pedrycz W. A dynamic adaptive subgroup-to-subgroup compatibility-based conflict detection and resolution model for multicriteria large-scale group decision making. IEEE Transactions on Cybernetics, 2020; 51(10): 4784–4795.

[7] Li J, Qin H, Wang J Z, Li J H. Openstreetmap-based autonomous navigation for the four wheel-legged robot via 3d-lidar and CCD camera. IEEE Transactions on Industrial Electronics, 2022; 69(3): 2708–2717.

[8] Lu Y Z, Young S. A survey of public datasets for computer vision tasks in precision agriculture. Computers and Electronics in Agriculture, 2020; 178: 105760.

[9] Cong Y, Chen R H, Ma B T, Liu H S, Hou D D, Yang C G. A comprehensive study of 3-D vision-based robot manipulation. IEEE Transactions on Cybernetics, 2023; 53(3): 1682–1698.

[10] Anh N P T, Hoang S, Van Tai D, Le B, Quoc C. Developing robotic system for harvesting pineapples. In: 2020 International Conference on Advanced Mechatronic Systems (ICAMechS), Hanoi, Vietnam: IEEE, 2020; pp.39–44.

[11] Zhao Z Q, Zheng P, Xu S T, Wu X D. Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 2019; 30(11): 3212–3232.

[12] Liu Y F, Cheng L. Relentless false data injection attacks against Kalman-filter-based detection in smart grid. IEEE Transactions on Control of Network Systems, 2022; 9(3): 1238–1250.

[13] Li Y, Ma L F, Zhong Z L, Liu F, Chapman M A, Cao D P. Deep learning for lidar point clouds in autonomous driving: A review. IEEE Transactions on Neural Networks and Learning Systems, 2021; 32(8): 3412–3432.

[14] Yang C G, Chen C Z, He W, Cui R X, Li Z J. Robot learning system based on adaptive neural control and dynamic movement primitives. IEEE Transactions on Neural Networks and Learning Systems, 2019; 30(3): 777–787.

[15] Wang X, Chen H, Gan C X, Lin H J, Dou Q, Tsougenis E, Huang Q T, et al. Weakly supervised deep learning for whole slide lung cancer image analysis. IEEE Transactions on Cybernetics, 2020; 50(9): 3950–3962.

[16] Li J H, Dai Y P, Wang J Z, Su X H, Ma R J, Towards broad learning networks on unmanned mobile robot for semantic segmentation. In: 2022 International Conference on Robotics and Automation (ICRA). Philadelphia, PA, USA: IEEE, 2022; pp.9228–9234.

[17] Zhuang J J, Luo S M, Hou C J, Tang Y, He Y, Xue X Y. Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications. Computers and Electronics in Agriculture, 2018; 152: 64–73.

[18] Kim J-y, Vogl M, Kim S-D. A code based fruit recognition method via image convertion using multiple features. In: 2014 International Conference on IT Convergence and Security (ICITCS), Beijing, China: IEEE, 2014; pp.1–4.

[19] Liu T H, Ehsani R, Toudeshki A, Zou X J, Wang H J. Identifying immature and mature pomelo fruits in trees by elliptical model fitting in the Cr-Cb color space. Precision Agriculture, 2019; 20: 138–156.

[20] Li J, Li J H, Zhao X, Su X H, Wu W B. Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4. Computers and Electronics in Agriculture, 2023; 211: 107955.

[21] Tang Y C, Qiu J J, Zhang Y Q, Wu D X, Cao Y H, Zhao K X, et al. Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: A review. Precision Agriculture, 2023; 24(4): 1183–1219.

[22] Xiao F, Wang H B, Xu Y Q, Zhang R Q. Fruit detection and recognition based on deep learning for automatic harvesting: An overview and review. Agronomy, 2023; 13(6): 1625.

[23] Gupta S, Tripathi A K. Fruit and vegetable disease detection and classification: Recent trends, challenges, and future opportunities. Engineering Applications of Artificial Intelligence, 2024; 133: 108260.

[24] Li J H, Li C L, Zeng S, Luo X W, Chen C P, and Yang C G. A lightweight pineapple detection network based on YOLOv7-tiny for agricultural robot system. Computers and Electronics in Agriculture, 2025; 231: 109944. doi. org/10.1016/j. compag. 2025.109944.

[25] Li, J H, Zhang T, Luo Q F, Zeng S, Luo X W, Chen C P, et al. A lightweight palm fruit detection network for harvesting equipment integrates binocular depth matching. Computers and Electronics in Agriculture, 2025; 233:110061. doi.org/10.1016/j.compag.2025.110061.

[26] Zhu A, Zhang R R, Zhang L H, Yi T C, Wang L W, Zhang D Z, et al. YOLOv5s-CEDB: A robust and efficiency camellia oleifera fruit detection algorithm in complex natural scenes. Computers and Electronics in Agriculture, 2024; 221: 108984.

[27] Li H W, Gu Z N, He D Q, Wang X C, Huang J D, Mo Y M, et al. A lightweight improved YOLOv5s model and its deployment for detecting pitaya fruits in daytime and nighttime light-supplement environments. Computers and Electronics in Agriculture, 2024; 220: 108914.

[28] Wu M C, Yuan K, Shui Y Q, Wang Q, Zhao Z X. A lightweight method for ripeness detection and counting of chinese flowering cabbage in the natural environment. Agronomy, 2024; 14(8): 1835.

[29] Fu L S, Feng Y L, Wu J Z, Liu Z H, Gao F F, Majeed Y, et al. Fast and accurate detection of kiwifruit in orchard using improved yolov3-tiny model. Precision Agriculture, 2021; 22: 754–776.

[30] Zhang X, Gao Q M, Pan D, Cao P C, Huang D H. Research on spatial positioning system of fruits to be picked in field based on binocular vision and SSD model. Journal of Physics: Conference Series, 2021; 1748:

042011.

[31] Liu T H, Nie X N, Wu J M, Zhang D, Liu W, Cheng Y F, et al. Pineapple (*Ananas comosus*) fruit detection and localization in natural environment based on binocular stereo vision and improved YOLOv3 model. Precision Agriculture, 2023; 24: 139–160.

[32] Zhu X Y, Chen F J, Zheng Y L, Chen C, Peng X D. Detection of *Camellia oleifera* fruit maturity in orchards based on modified lightweight YOLO. Computers and Electronics in Agriculture, 2024; 226: 109471.

[33] Li H L, Li J, Wei H B, Liu Z, Zhan Z F, Ren Q L. Slim-neck by GSConv: A lightweight-design for real-time detector architectures. Journal of Real-Time Image Processing, 2024; 21(3): 62.

[34] Yang H Y, Yang L, Wu T, Yuan Y J, Li J C, Li P. MFD-YOLO: A fast and lightweight model for strawberry growth state detection. Computers and Electronics in Agriculture, 2025; 234: 110177.

[35] Wang Y, Yan G, Meng Q L, Yao T, Han J F, Zhang B. DSE-YOLO: detail semantics enhancement YOLO for multi-stage strawberry detection. Computers and Electronics in Agriculture, 2022; 198: 107057.

[36] Bai Y F, Yu J Z, Yang S Q, Ning J F. An improved YOLO algorithm for detecting flowers and fruits on strawberry seedlings. Biosystems Engineering, 2024; 237: 1–12.

[37] Ma B L, Hua Z X, Wen Y C, Deng H X, Zhao Y J, Pu L R, et al. Using an improved lightweight YOLOv8 model for real-time detection of multi-stage apple fruit in complex orchard environments. Artificial Intelligence in Agriculture, 2024; 11: 70–82.

[38] Bailly A, Blanc C, Francis É, Guillotin T, Jamal F, Wakim B, et al. Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. Computer Methods and Programs in Biomedicine, 2022; 213: 106504.

[39] Bargoti S, Underwood J P. Image segmentation for fruit detection and yield estimation in apple orchards. Journal of Field Robotics, 2017; 34(6): 1039–1060.

[40] Singh G, Mittal A. Various image enhancement techniques- A critical review. International Journal of Innovation and Scientific Research, 2014; 10: 267–274.

[41] Jocher G, Nishimura K, Mineeva T, Vilari R. no. yolov5. Code repository, 2020. Available: https://github.com/ultralytics/yolov5. Accessed on [2024-03-20].

[42] Wang C Y, Liao H M, Wu Y H, Chen P Y, Hsieh J-W, et al. CSPNet: A new backbone that can enhance learning capability of CNN. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA: IEEE, 2020; pp.390–391.

[43] Gevorgyan Z. SIoU loss: More powerful learning for bounding box regression. arXiv: 2205.12740, 2022; in Press. doi: 10.48550/arXiv.2205.12740.

[44] Yu G, Tao W, Guo G Q, Liu H C. SFHG-YOLO: A simple real-time small-object-detection method for estimating pineapple yield from unmanned aerial vehicles. Sensors, 2023; 23(22): 9242.

[45] Li J H, Liu Y W, Li C L, Luo Q F, Lu J H. Pineapple detection with YOLOV7-tiny network model improved via pruning and a lightweight backbone sub-network. Remote Sensing, 2024; 16(15): 2805.

[46] Peng G Z, Chen C L P, Yang C G. Robust admittance control of optimized robot-environment interaction using reference adaptation. IEEE Transactions on Neural Networks and Learning Systems, 2022; 34(9): 5804–5815.

[47] Dong S L, Chen C L P, Fang M, Wu Z G. Dissipativity-based asynchronous fuzzy sliding mode control for T–S fuzzy hidden markov jump systems. IEEE Transactions on Cybernetics, 2020; 50(9): 4020–4030.

[48] Li J H, Zeng D Z, Luo Q F, Luo X W, Chen C P, and Yang C G. Feature Assessment and Enhanced Vertical Constraint Lidar Odometry and Mapping on Quadruped Robot. IEEE Transactions on Instrumentation and Measurement, 2025; 74: 3533649.