Feature recognition algorithm in intelligent planting and preparation technology of single-bud segment sugarcane

Xinpeng Liu[1,6†], Xuehu Dong[2†], Mingxin Hou[3], Zhaojun Niu[4*], Zunxiang Li[1,6], Zhen Zhang[5], Li Huang[3*]

(1. *Institute of Subtropical Crops, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang 524088, Guangdong, China*;
2. *Hainan Agricultural Machinery Appraisal and Promotion Station, Haikou 570000, China*;
3. *Guangdong Ocean University, Zhanjiang 524088, Guangdong, China*;
4. *Institute of Agricultural Machinery, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang 524088, Guangdong, China*;
5. *Chinese Academy of Tropical Agricultural Sciences, Haikou 570100, China*;
6. *Key Laboratory of Tropical Fruit Biology, Ministry of Agriculture and Rural Affairs, Zhanjiang 524091, Guangdong, China*)

Abstract: In order to avoid the uneven phenomenon of sugarcane planting, such as seed missing and reseeding, the computer vision technology was applied to the intelligent identification of sugarcane varieties with single-bud segment, and the design idea of rapid detection of sugarcane planting distribution was proposed in this study. With sugarcane species with single-bud segment as the research object, the sugarcane species distribution image was acquired, and LabelImg was used for image annotation and format conversion to build the YOLOv5s target detection model. On the basis of depth-separable convolution, SE module is spliced to obtain the weights of extracted features and extract key features of input feature map. By adding regularization to constrain the BN layer coefficient, sparse regularization is carried out on the BN layer to reduce the network input size and improve the model training speed. On this basis, 600 rounds of iterative training were carried out to complete the target recognition of sugarcane species characteristics in single-bud segment. The results showed that the recognition accuracy, mAP value, and Recall value of YOLOv5s single-bud segment target detection model are 98.95%, 98.89%, and 98.69%, and the loss value converges in advance between 0-0.02. The results showed that YOLOv5s could effectively detect and identify sugarcane seeds with single-bud segment during field planting, which lays a foundation for promoting precise and intelligent sugarcane planting.

Keywords: sugarcane seed with single bud, object detection, YOLOv5s, intelligent identification
DOI: 10.25165/j.ijabe.20251804.8544

Citation: Liu X P, Dong X H, Hou M X, Niu Z J, Li Z X, Zhang Z, et al. Feature recognition algorithm in intelligent planting and preparation technology of single-bud segment sugarcane. Int J Agric & Biol Eng, 2025; 18(4): 275–281.

1 Introduction

Sugarcane originates from the tropics and subtropics and is a major cash crop in the southern hot zone of China and the main raw material for sugar production[1]. The total area of sugarcane planted in China is the third highest in the world, after Brazil and India[2]. According to the survey statistics of the national sugarcane production information monitoring department, the area of sugarcane planted in China in the 2020/21 crushing season was 1 476 180 hm^2, mainly distributed in the four main production areas, Guangdong, Guangxi, Hainan, and Yunnan. Traditional sugarcane planting methods are characterized by serious variety degradation, low planting efficiency, decreasing sugar content, and large consumption of seed stems for planting[3], while the healthy seedling technology of sugarcane has the advantages of variety purification and rejuvenation, improving sugar content and yield, and saving planted cane seeds[4,5]. However, the healthy seedling technology of sugarcane requires more reliance on machinery to be implemented more economically and effectively, and there is no mature dedicated machinery and equipment at home or abroad[6,7]. Combined with the technology of virus-free sugarcane seedling, cultivation technology, and the planting mode of wide row and thin planting, the research on the intelligent recognition planter of sugarcane stalk with single-bud segment was carried out[8,9].

Deep learning-based target detection is to detect specific classes of semantic objects with different poses in digital images and videos. The current successful target detection methods are based on deep learning model extensions[10-15]. In the process of target detection, RPN uses FPN to convolve the feature map of the detected target, and then adds it. After 1×1 convolution (horizontal connections), it is added to the feature map of the lower layer of the convolutional network to form one of the M feature layers. Following this operation, each of the M feature layers is constructed layer by layer from top to bottom, and each undergoes 3×3

Received date: 2023-09-25 **Accepted date:** 2025-03-20
Biographies: Xinpeng Liu, MS candidate, Assistant Research Fellow, research interest: tropical agricultural machinery, Email: liuxpchina@163.com; **Xuehu Dong**, BS, Research Fellow, research interest: tropical agricultural machinery, Email: 825346179@qq.com; **Mingxin Hou**, PhD, Associate Professor, research interest: machine Vision, Email: houmx@gdou.edu.cn; **Zunxiang Li**, BS, Assistant Research Fellow, research interest: tropical agricultural machinery, Email: 1192187600@qq.com; **Zhen Zhang**, MS candidate, Assistant Research Fellow, research interest: software engineering, Email: 1525848365@qq.com.
†The authors contributed equally to this work
***Corresponding author: Zhaojun Niu**, MS Candidate, Associate Researcher, research interest: tropical agricultural machinery. Institute of Agricultural Machinery, Chinese Academy of Tropical Agricultural Sciences, No 5, Shetan Road, Xiashan District, Zhanjiang 524013, Guangdong, China. Tel: +86-13030195117, Email: niuzhaojun0818@163.com; **Li Huang**, MS candidate, Lecturer, research interest: machine vision. Guangdong Ocean University, No 1, Haida Road, Mazhang District, Zhanjiang 524088, Guangdong, China. Tel: +86-18669509286, Email: 245041431@qq.com.

convolution to generate multiple scale region proposals. These multi-scale feature maps and regional proposals are fed into the ROI (Region of Interest) pooling layer to obtain the proposal feature maps. By inputting the proposed feature mapping into the fully connected layer for prediction, the target detection model is obtained[16]. Then, through mechanical vision, the pose of the object is provided to establish a model library, using tactile sensors to identify the target object[17].

Therefore, this study is inspired by the research on target detection and recognition. According to the physical characteristics of sugarcane seed such as color, length, and size, combined with the conditions of field operation, the iteratively updated YOLO algorithm was adopted with sugarcane planter as the carrier. Research on accurate identification of sugarcane seed with single-bud segment based on YOLOv5s was carried out to avoid the phenomenon of missing seed, multiple species, and too large a plant distance in the sugarcane planting process, and to accelerate the accurate and intelligent process of sugarcane seed with single-bud segment.

2 Algorithm principle

2.1 Structural principle of YOLOv5s

The YOLOv5s target detection algorithm inputs the single-bud sugarcane seed images into the convolutional neural network after data enhancement, and feature maps of different sizes are obtained by feature extraction. Then the feature maps are divided into rectangular grids, to obtain the prediction frame position and category after forward propagation. The loss function is applied to reduce the loss between the prediction frame and the real frame, the redundant windows are screened out by non-maximal value suppression, and the target detection results can be obtained[18-20].

YOLOv5s target detection mainly consists of four parts: input port, Backbone network, Neck, and prediction[21,22], as shown in Figure 1. The input port is mainly to complete the enhancement of the input single-bud sugarcane seed image data. The backbone network structure mainly consists of Focus, BottleneckCSP, and SPP, responsible for the feature extraction of the image, in which Focus slices the image, stitches it after the channel dimension, and the obtained feature map is subjected to convolution operation.

SPP performs the maximum pooling operation on the input features and performs channel stitching to get the features fused with multiple scales. The Neck structure adopts FPN+PAN network, which can achieve the fusion of features of different sizes and shapes, etc. to a certain extent, then sends them to the detection model for target prediction, and completes the prediction of the input target images.

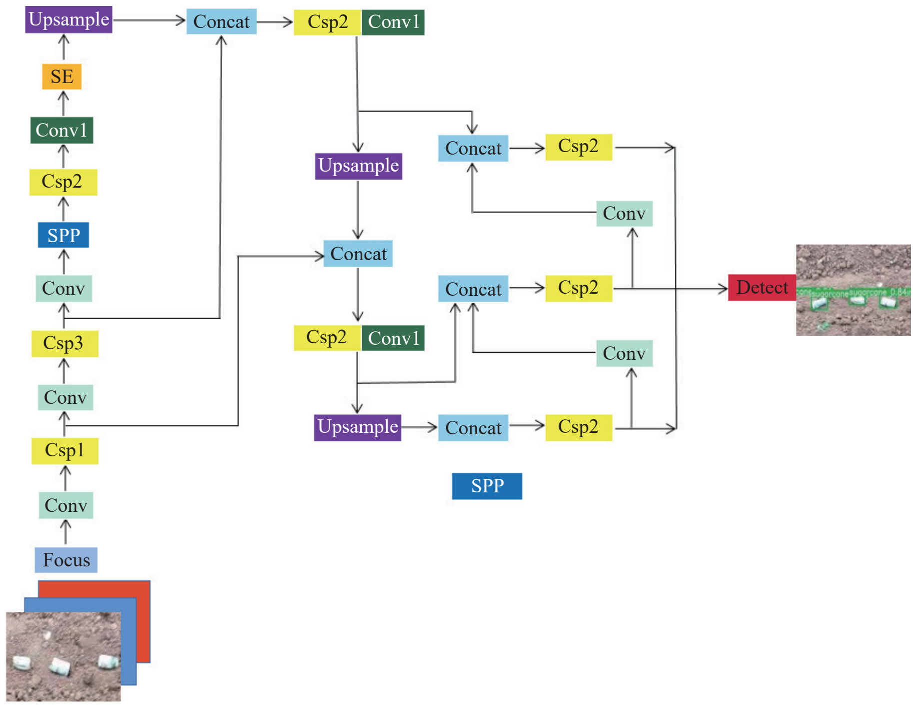

Figure 1 YOLOv5 structure

The basic block of the algorithm used in Figure 1 consists of the following units: (1) ConV layer, which mainly extracts the feature information of the input image; (2) BN batch layer, which mainly normalizes the input and suppresses the overfitting phenomenon and gradient dispersion problem; (3) LeakyRelu activation function, which introduces nonlinear factors so that the model can perform nonlinear tasks; (4) Concat layer, which mainly stitches the feature maps to achieve feature information fusion; (5) Slice, which is mainly used to slice the input image to ensure that the Focus module achieves two-fold downsampling; and (6) Maxpool layer, which mainly extracts the most important image features and improves the training speed to prevent the overfitting phenomenon.

2.2 Loss function

Based on the analysis of the target detection principle and algorithm structure, the sum of loss functions is obtained by performing target classification, target prediction frame regression, and target confidence regression operations on the target detectors before predicting the outcome model for single-bud sugarcane seeds[23].

Firstly, the prediction categories of the target classification loss function and the probability of the predicted outcome categories are calculated by differentiating the cane species in the input images of the single- bud sugarcane seeds that are input to the target detection model. According to the characteristics of IoU, GIoU, DIoU, and CIoU, combined with the target detection characteristics of single-bud sugarcane seeds, the aspects such as ratio and overlapping area of the prediction frame and the distance between the center points of the labeled frames are considered. The weight coefficient and the distance of the aspect ratio are added to the penalty term of DIoU to make the CIoU metric principle more consistent with the regression mechanism of the target prediction frame. So CIoU is chosen as the loss function for the regression task of the prediction frame position of single-bud sugarcane seeds. In the case that the overlap area between the prediction frame and the real frame is increasing, the distance measurement is added to the location confidence in IoU loss resolution target detection. The coincidence degree between the prediction box and the real box is thus effectively reflected. Therefore, the GIoU loss function is selected for the regression task of predicting target confidence for single-bud sugarcane seeds.

2.3 Dataset preparation

For the sugarcane seed image of the single-bud segment obtained, moderate intensive annotation method was used to label the sugarcane seed image of single-bud segment, and LabelImg was used to complete the image annotation. The .xml file is generated by annotating the external outline of sugarcane seed with single-bud segment, as shown in Figure 2. The database is provided for the target detection algorithm; the labeled single-bud segment sugarcane seed labeling box cannot be too large or too small, otherwise the deep learning model may not perform well in identifying sugarcane varieties. The dataset with single-bud sugarcane seed labels is input to the deep learning model, and the mapping relationship between the dataset input and the labels is obtained and applied to the test dataset for classification or regression purposes.

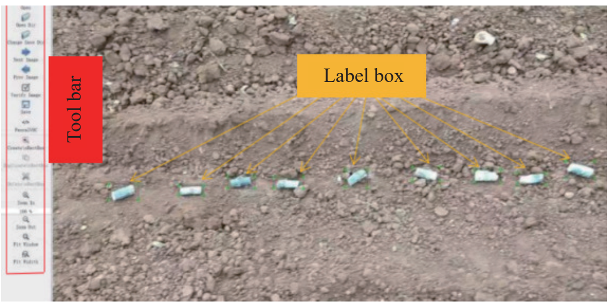

Figure 2 Labeling process of single-bud sugarcane seed

3 Network model construction

In this study, the single-bud sugarcane bud is selected as a single target and its coordinate information is mainly detected. Therefore, the shallow YOLOv5s network detection model is selected, which has fewer feature extraction operations, relatively few model parameters, and less computational amount of image semantic information extraction, which is conducive to improving the sugarcane species detection speed in single-bud segment and meeting the planting requirements of real-time detection in the process of sugarcane planting. In order to reduce the computation and speed up the training, the original images of the input network model are compressed to 640×640 pixels and trained with pre-training weights fine-tuned on the dataset. The structure of the network feature extraction module is shown in Figure 3, in which Conv1 network base module outputs the input feature map after convolutional normalization. The Bottleneck module superimposes the initial input feature map after two operations on Conv1, and outputs it by combining with different depth feature information. The CSP module improves the feature extraction capability; CSP1 and CSP3 modules are used as the backbone of feature extraction, which stack the input feature map with the output of the original feature map after Conv1, Bottleneck, and convolution operations, and then output after normalization, activation function, and Conv1 operations. CSP2 module is used for prediction, which differs from CSP1 and CSP3 in that the Bottleneck module is replaced by Conv1 module.

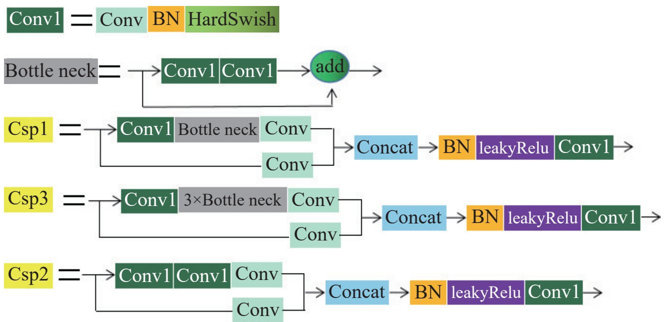

Figure 3 Structure of feature extraction module of detection target image

3.1 Network model optimization

Standard convolution is to use convolution kernels on all channels of input for feature extraction to enhance the network model training effect. In order to reduce the model parameters and reduce the number of operations, the depth-separable convolution is used. As shown in Figure 4, in the channel direction of each input channel, different sizes of convolution kernels are used for depth convolution operation. At the same time, 1×1 convolution kernel point-by-point convolution operation is conducted, and the number of channels is adjusted to improve the relevance of channel and spatial dimension information, so that the network is more flexible[24].

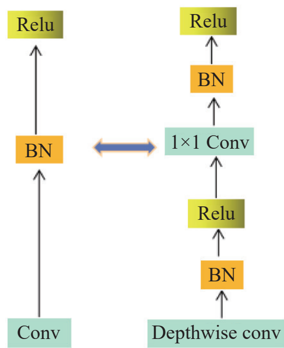

Figure 4 Standard convolution and depth-separable

After convolving a feature map with an input size of $W\times H\times C$ and an $S\times S\times C$ convolution kernel, the output feature map has a size of $W_1\times H_1\times N$, resulting in standard convolution and depth-separable convolution calculations, which are:

$$W\times H\times C\times S\times S\times C \tag{1}$$

$$S\times S\times C\times W\times H+C\times N\times W_1\times H_1 \tag{2}$$

$$\frac{S\times S\times C\times W\times H+C\times N\times W_1\times H_1}{W\times H\times C\times S\times S\times C}=\frac{1}{N}+\frac{1}{S^2} \tag{3}$$

where, W is the width of the input feature map; H is the height of the input feature map; W_1 is the width of the output feature map; H_1 is the height of the output feature map; C is the number of input network channels; N is the number of output network channels.

On the premise that the feature map does not change, the width and height W and H of the input feature map and W_1 and H_1 of the output feature map are ensured to be equal, as S is 3, and N is 256, 512, or 1024, etc. After the comparison of the two convolution parameters, we can obtain the depth-separable convolution operations account for $\frac{1}{9}-\frac{1}{8}$ of the standard convolution operations.

On the basis of deep-separable convolution, several feature extraction modules such as SE module and ReLU6 activation function are spliced, which can complete the operations of deep-separable convolution and key feature extraction of the input feature maps. Among them, SE module is a commonly used attention mechanism module in deep learning[25], which can strengthen the identification ability of sugarcane species in single-bud segment. It mainly obtains the weight of each feature channel through compression, pooling, excitation, and function activation of input feature map, filters key features, simulates the relationship between feature channels, and obtains the correlation between channels.

The input is $\boldsymbol{X} = [\chi 1, \chi 2, \ldots, \chi C]$ and $\boldsymbol{X} \in R^{H\times W\times C} R^{H^1\times W^1\times C^1}$, the convolution kernel is defined as: $\boldsymbol{V} = [v_1, v_2, \ldots, v_\iota]$, and the output after convolution is obtained as $\boldsymbol{Y} = [y_1, y_2, \ldots, y_\iota]$, $\boldsymbol{Y} \in \boldsymbol{R}^{H^1\times W^1\times C^1}$; where the ιth convolution kernel is: $\boldsymbol{v}_\iota = [v_\iota^1, v_\iota^2, \ldots, v_\iota^c]$ and the following relation is obtained:

$$\boldsymbol{y}_\iota = \boldsymbol{v}_\iota \times X = \sum_{i=1}^{c} v_\iota^i \otimes \chi^i \tag{4}$$

where, H and H^1 are the convolutional input and output feature channel lengths, respectively; W and W^1 are the convolutional input and output feature channel widths, respectively; C and C^1 are the numbers of convolutional input and output feature channels, respectively; $\otimes$ is the convolution operation; v_ι^i is the first two-dimensional convolution kernel. The feature channels given by y_ι are summed with the convolution kernel learning space relation to obtain the convolution output.

The average value of each feature map is calculated by F_{sq} to transform a single 2-D feature channel into a real number, and finally a 1-D vector is obtained to achieve the reduction in the number of computational parameters and the avoidance of overfitting. The weights of each channel are obtained by F_{ex}, and the original output is weighted to extract key features and suppress invalid features. The SE module is made to implement the weights of extracted features to change the degree of attention to different feature regions, as shown in Figure 5.

Figure 6a shows the prediction of single-bud sugarcane seeds in the single row planting area by YOLOv5s; the correct prediction of each single-bud sugarcane seed is obtained in the prediction area, and the confidence level of the model is between 69% and 76%. Figure 6b shows the prediction of single-bud sugarcane seeds in the single row planting area by SE-YOLOv5s. The correct prediction of each single-bud sugarcane seed was obtained in the prediction area. The confidence level of the model is between 77% and 84%, which is about 8% higher than the confidence level of YOLOv5s. This further demonstrates the effectiveness of the SE-YOLOv5s model in detecting single-bud sugarcane seeds within the planting area.

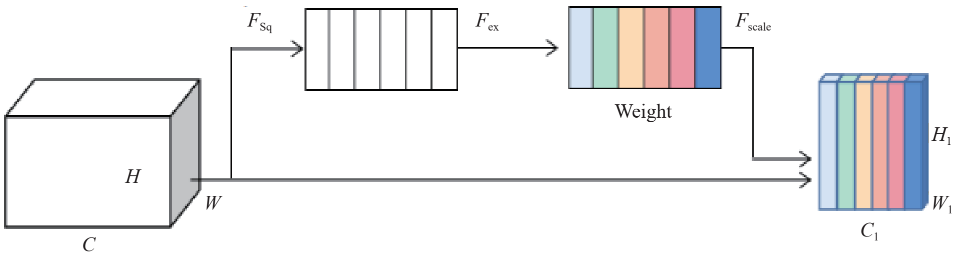

Figure 5 SE module

a. Sugarcane seeds predicted by YOLOv5s

b. Sugarcane seeds predicted by SE-YOLOv5s

Figure 6 Comparison of single-bud sugarcane seed model assays

3.2 Pruning comparison analysis

There are two trainable parameters in the BN layer, γ and β. When the training parameters γ and β are input to the BN layer, a normalized distribution is obtained; when γ and β are converging to 0, the input is equivalent to multiplying by 0. After threshold separation, the output is 0, and the input of the convolutional layer connected to it is 0. The principle is as follows:

Input: Values of $\mathcal{X}$ over a mini-batch: B={$\mathcal{X}_1$...};

Parameters to be learned: γ and β

Output: {$\mathcal{Y}_i$=BN$_{\gamma,\beta}$($\mathcal{X}_i$)}

To reduce the input size of the network, the BN layer coefficients are constrained by adding a regular, and after sparse regularization of the BN layer, all BN layer weights in the model are sorted statistically, and the sorted weight threshold (thres) is obtained. The mask of each layer is made (if weight > thres, the value is 1; if weight < thres, the value is 0), the unimportant channels are identified, the sparse smaller layers are cropped, and the specified reserved BN layer is obtained, namely, repeating the iteration, to build a new model structure according to the number of channels reserved on each layer, so as to obtain the mask non-zero index of

BN layer weight. Then the reset network assignment is conducted to adjust it slightly to achieve improvement in training speed.

Due to the great parameter redundancy in YOLOv5s, the comparison of model experimental effects after pruning cutting is clear, as listed in Table 1.

The model input size has been changed from 29 M to 1.3 M, and the speed accelerates from 12 ms to 4 ms. The average accuracy and recall rate are closer.

Table 1 Comparison of model indicators

Metric	Model size/M	Speed/ms	mAP	Recall
YOLOv5s	443	12	0.9563	0.9763
Pruned	15.3	4	0.9889	0.9869

4 Test

4.1 Test environment

The GPU is RTX3070TI, the CPU has 8 cores and 16 threads, the maximum frequency is 4.6 GHz, the memory is 16 GB, the hard disk is 512 G, and the system is Windows 10. The network model is implemented in Python 3.8 programming language, Pycharm software is the platform, and the PyTorch framework is adopted. Single-bud sugarcane seed dataset is converted from PASCAL VOC to YOLO, which is trained, verified, and tested through Pycharm.

4.2 Test results

In order to accelerate the training and convergence of the single-bud sugarcane seeds based on YOLOv5s, the hyper-parameters that are needed for the construction and training of the detection model are set, such as operating environment, pre-training weight, and dependent library. The model is calculated by GPU, as the initial model training is performed by using a smaller learning rate, and the learning rate strategy is adjusted by combining the gradient descent algorithm in cosine annealing to optimize the objective function[26-29]. The test image is 640×640, and the training Batch size is 16. The input data are processed by data enhancement techniques to generate more training samples to improve the accuracy.

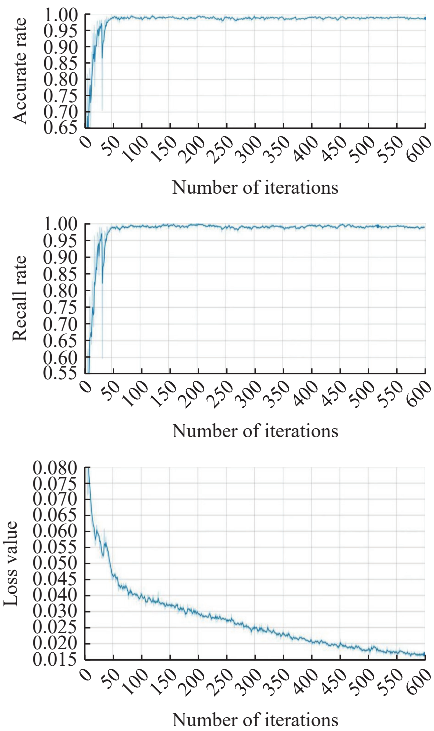

Figure 7 YOLOv5s model accuracy rate, recall rate, and loss value curve

After 600 rounds of model training, the operation process curves of precision rate and recall rate are obtained, as shown in Figure 7. The accuracy rate of the training set is 98.95% and the recall rate is 98.69%, and their convergence values are both high and closer to 1. This indicates that the target detection model for single-bud sugarcane seeds is not subject to overfitting or underfitting, and the target prediction results of the model are accurate and reliable. The target loss of the training set converges faster; its loss values converge between 0 and 0.02 in advance, indicating that the target detection model parameters are reasonable and the target detection is accurate.

4.3 Target detection evaluation

YOLOv5s adopts IoU and a matching strategy based on aspect ratio, which is to calculate the respective corresponding aspect and height ratios of each ground truth box and nine anchors, taking the most extreme ratio. When the ratio of ground truth box and anchor is less than the set ratio threshold, the anchor is responsible for predicting the ground truth box, which is called a positive sample, and all other prediction frames are negative samples. The mean accuracy (mAP), the precision P, and the recall R are used as evaluation indices to assess the detection effectiveness of the training model. The confidence is calculated by the Softmax function as an evaluation of the detection accuracy of single-bud sugarcane seeds, with P denoting the ratio of the number of single-bud sugarcane seeds among the identified targets and R denoting the ratio of recognized seeds among all single-bud sugarcane seeds. The P-R curves are obtained by the test, as shown in Figure 8; the P and R index values are 98.95% and 98.69%, close to 1, and the index value of mAP is 98.89%, which indicates that the model has good performance in detecting single-bud sugarcane seeds.

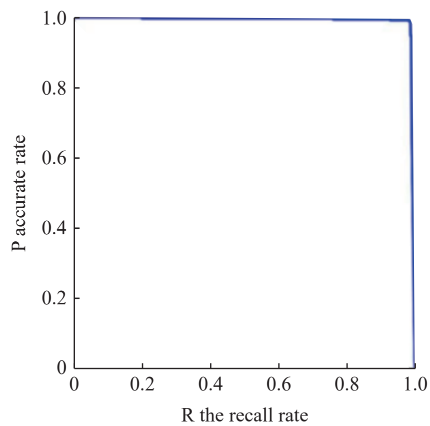

Figure 8 P-R curves

4.4 Visualization of target detection results of single-bud sugarcane seeds

To show the target detection results of YOLOv5s, for the original target images under different conditions such as field lighting, shaking, and irregular seeds, the YOLOv5s target detection model predicts the accurate target detection frame of single-bud sugarcane seeds and visualizes the single-bud sugarcane seed target detection results[30,31]. By calculating the weighting of the Grad CAM feature map with the corresponding category weights, the region most relevant to a specific category of single-bud sugarcane seeds can be displayed without changing the structure of the model and mapped to the original image, so that the heat map and the original image can be superimposed, as shown in Figure 9. The darker the

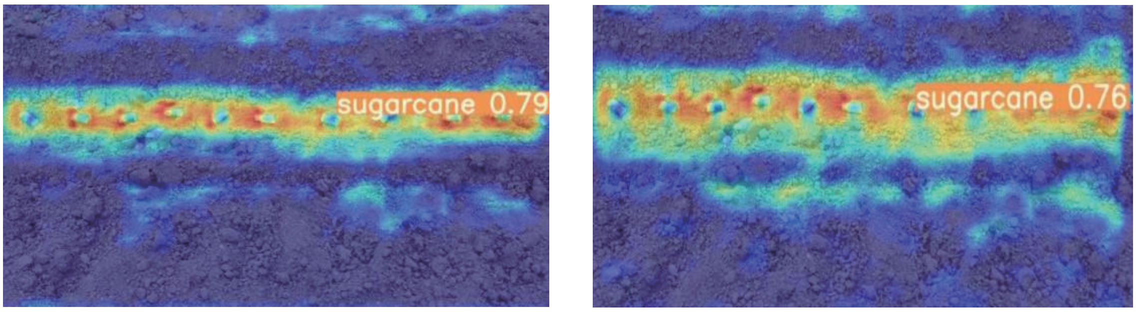

Figure 9 Detection effect of single-bud sugarcane seeds

color part of this visual output image, the larger the value, and the stronger the feature extraction ability of the model. In particular, the red part, which represents the active region, pays more attention to the shape and position of the single-bud species in this region. Highlighting its contour line can better reflect the location of the cane seeds in the original image, so that the feature extraction effect of the model can be better expressed. It reflects the good robustness and accuracy of the model, while the accurate prediction results of the model can provide a theoretical reference for target detection and intelligent recognition in actual field operations.

5 Conclusions

(1) Based on YOLOv5s target detection algorithm, a target detection model of single-bud sugarcane seeds is established. Through dataset annotation, parameter adjustment, training, verification, testing, and other measures, the accuracy of training and verification of single-bud sugarcane seeds in the model is close to each other and converges, respectively, avoiding overfitting or underfitting of the model, so as to complete the intelligent recognition of single-bud sugarcane seeds.

(2) Through parameter selection and debugging of the model, during YOLOv5s target detection of single- bud sugarcane seeds, the target detection accuracy rate of single-bud sugarcane seeds is 98.95%, the mAP value is 98.89%, and the Recall value is 98.69%, which indicates that the target detection model parameter selection is reasonable. The model training results are reliable, without overfitting, and it has good generalization ability and robustness. The requirements for real-time and accurate recognition performance during the field planting process of single-bud sugarcane seeds can be met. It has laid the foundation for promoting precise and intelligent planting of sugarcane.

Acknowledgements

We acknowledge that this work was financially supported by the Research Funds for the South Asian Tropical Crop Research Institute of the Chinese Academy of Tropical Agricultural Sciences (Grant No. 1630062025008), the Project of Guangxi Zhuang Autonomous Region Key Technologies R&D Program (Grant No. GK AB23026069 and GN AB241484034), Natural ScienceFoundation of Guangdong Province (2025A1515012901), Hot Zone Hilly Mountain Small Machinery Technology Innovation Team (Grant No. CATASCXTD202409), Hainan Provincial Natural Science Foundation (Grant No. 524QN332), and Fundamental Research Funds for Central Public Welfare Research Institutes (Grant No. 1630062024009).

[References]

[1] Vianna M d S, Nassif D S P, Carvalho K d S, Marin F R. Modelling the trash blanket effect on sugar-cane growth and water use. Computers and Electronics in Agriculture, 2020; 172(5): 105361.
[2] Elsheery N I, Sunoja V S J, Wen Y, Zhu J J, Muralidharan G, Cao K F. Foliar application of nanoparticles mitigates the chilling effect on photosynthesis and photoprotection in sugarcane. Plant Physiology and Biochemistry, 2020; 149(4): 50–60.
[3] Yang R, Li J, Liu Q, Huang W S, Yin K, Qiao X, et al. Gradient-based method for the identification of multi-nodes in sugarcane. Information Processing in Agriculture, 2020; 7(4): 491–499.
[4] Henry O, Patricia L, Raúl Mira, Angel D. Computer vision based food grain classification: A comprehensive survey. Computers and Electronics in Agriculture, 2021; 187(8): 1688–1699.
[5] He Q, Shen H, Jin Y. Research on novel lightweight object detection and recognition algorithms. International Conference on Information Systems Engineering (ICISE), 2023; pp.442–448. DOI: 10.1109/ICISE60366.2023.00100.
[6] Li B Z, Jiang W Z, Gu J J. Research on target detection algorithm based on deep learning technology. IEEE International Conference on Power Electronics, Computer Applications (ICPECA), Shenyang, China, 2021; pp.137–142. DOI: 10.1109/ICPECA51329.2021.9362714.
[7] He F G, Deng G R, Li M, Niu Z J, You X. Design of electromagnetic vibration-type single-bud sugarcane seed metering device. Sugar Tech, 2019; 21(5): 815–824.
[8] Fan Y L, Zhou D Q, Deng Q R, He F G, Wang M L. Design and experiment of sugarcane seed cutting equipment based on machine vision. Journal of Agricultural Mechanization Research, 2021; 43(5): 55–60.
[9] Liu X P, Niu Z J, Li M, Hou M X, Wei L J, Zhang Y, Huang L. Design and experimental research on disc-type seeding device for single-bud sugarcane seeds. Int J Agric & Biol Eng, 2023; 16(2): 115–124.
[10] Murthy C B, Hashmi M F, Keskar A G. Optimized MobileNetplus SSD: a real-time pedestrian detection on a low-end edge device. International Journal of Multimedia Information Retrieval, 2021; 10: 171-184.
[11] Velumani K, Madec S, de Solan B, Lopez-Lozano R, Gillet J, Labrosse J, et al. An automatic method based on daily in situ images and deep learning to date wheat heading stage. Field Crops Research, 2020; 252: 107793.
[12] Ji J T, Han Z H, Zhao K X, Li Q W, Du S C. Detection of the farmland plow areas using RGB-D images with an improved YOLOv5 model. Int J Agric & Biol Eng, 2024; 17(3): 156–165.
[13] Shi H K, Xiao W F, Zhu S P, Li L B, Zhang J F. CA-YOLOv5: Detection model for healthy and diseased silkworms in mixed conditions based on improved YOLOv5. Int J Agric & Biol Eng, 2023; 16(6): 236–245.
[14] Zou X G, Yin Z L, Li Y H, Gong F, Bai Y G, Zhao Z H, et al. Novel multiple object tracking method for yellow feather broilers in a flat breeding chamber based on improved YOLOv3 and deep SORT. Int J Agric & Biol Eng, 2023; 16(5): 44–55.
[15] Zhao L J, Jia Y F, Yin W K, Li Z H, Liu C D, Luo H, et al. Recognition and localization method of maize weeding robot based on improved YOLOv5. Int J Agric & Biol Eng, 2025; 18(2): 248–258.
[16] Deng R L, Tao M, Huang X N, Bangura K, Jiang Q, Jiang Y, et al. Automated counting grains on the rice panicle based on deep learning method. Sensors, 2021; 21(1): 281.
[17] Kraft D, Bierbaum A, Kjaergaard M, Ratkevicius J, Kjaer-Nielsen A, Ryberg C, et al. Tactile object exploration using cursor navigation sensors. WorldHaptics, 3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Salt Lake City, UT, USA, 2009; pp.296–301. DOI: 10.1109/WHC.2009.4810827.
[18] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, et al. SSD: Single shot multibox detector. Proceedings of the European Conference on Computer Vision. Cham: Springer, 2016; pp.21–37. DOI: 10.1007/978-3-

319-46448-0_2.

[19] Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017; 39(6): 1137–1149.

[20] He K M, Zhang X Y, Ren S Q, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015; 37(9): 1904–1916.

[21] Bi L Y, Deng L X, Lou H T, Zhang H Y, Lin S J, Liu X C, et al. URS-YOLOv5s: object detection algorithm for UAV remote sensing images. Physica Scripta, 2024; 99(8): 086005.

[22] Aytekin A, Rezaeitabar Y, Dogru S, Ulusoy İ. Railway fastener inspection by real-time machine vision. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2015; 45(7): 1101–1107.

[23] Zheng Z, Wang P, Ren D, Liu W, Ye R, Hu Q, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation. IEEE Transactions on Cybernetics, 2022; 52(8): 8574–8586.

[24] Nafea M M, Tan S Y, Jubair M A. A review of lightweight object detection algorithms for mobile augmented reality. International Journal of Advanced Computer Science and Applications (IJACSA), 2022; 13(11): 536–546.

[25] Tan M X, Le Q V. MixConv: Mixed depthwise convolutional kernels. Published online December 1, 2019. DOI: 10.48550/arXiv.1907.09595.

[26] Loshchilov I, Hutter F. Sgdr: Stochastic gradient descent with warm restarts. Published online May 3, 2017. DOI: 10.48550/arXiv.1608.03983.

[27] Chu Y, Wang J, Ma L, Wu C. LMSFA-YOLO: A lightweight target detection network in Remote sensing images based on Multiscale feature fusion. J. King Saud Univ. Comput. Inf. Sci., 2025; 37(63): 1–21.

[28] Song Z S, Liu Y, Zheng L, Tie J, Wang J. Identification of green citrus based on improved YOLOV3 in natural environment. Journal of Chinese Agricultural Mechanization, 2021; 42(11): 159–165.

[29] Liu J B, Huang Y P, Wang S C, Zhao X X, Zou Q, Zhang X Y. Rail fastener defect inspection method for multi railways based on machine vision. Railway Sciences, 2022; 1(2): 210–223.

[30] Zhou L, Zhang C, Taha M F, Wei X, He Y, Qiu Z, et al. Wheat kernel variety identification based on a large near-infrared spectral dataset and a novel deep learning-based feature selection method. Front Plant Sci, 2020; 11: 575810.

[31] Jia Z W, Yang X H, Wang Z, Yu R R, Wang R B. Automatic lameness detection in dairy cows based on machine vision. Int J Agric & Biol Eng, 2023; 16(3): 217–224.