agronomy

MDPI

Article

Sugarcane-YOLO: An Improved YOLOv8 Model for Accurate Identification of Sugarcane Seed Sprouts

Fujie Zhang, Defeng Dong, Xiaoyi Jia, Jiawen Guo * and Xiaoning Yu

Sugarcane Research Institute, Yunnan Academy of Agricultural Sciences, Kaiyuan 661699, China; 20030031@kust.edu.cn (F.Z.); 15239738900@sohu.com (D.D.); 15284417075@sohu.com (X.J.); 15138169516@sohu.com (X.Y.)

* Correspondence: gjw@yaas.org.cn; Tel.: +86-13769314690

Abstract: Sugarcane is a crop that propagates through seed sprouts on nodes. Accurate identification of sugarcane seed sprouts is crucial for sugarcane planting and the development of intelligent sprout-cutting equipment. This paper proposes a sugarcane seed sprout recognition method based on the YOLOv8s model by adding the simple attention mechanism (SimAM) module to the neck network of the YOLOv8s model and adding the spatial-depth convolution (SPD-Conv) to the tail convolution part. Meanwhile, the E-IoU loss function is chosen to increase the model's regression speed. Additionally, a small-object detection layer, P2, is incorporated into the feature pyramid network (FPN), and the large-object detection layer, P5, is eliminated to further improve the model's recognition accuracy and speed. Then, the improvement of each part is tested and analyzed, and the effectiveness of the improved modules is verified. Finally, the Sugarcane-YOLO model is obtained. On the sugarcane seed and sprout dataset, the Sugarcane-YOLO model performed better and was more balanced in accuracy and detection speed than other mainstream models, and it was the most suitable model for seed and sprout recognition by automatic sugarcane-cutting equipment. Experimental results showed that the Sugarcane-YOLO achieved a mAP50 value of 99.05%, a mAP72 value of 81.3%, a mAP50-95 value of 71.61%, a precision of 97.42%, and a recall rate of 98.63%.

Keywords: sugarcane sprouts; Sugarcane-YOLO; SimAM; SPD-Conv; E-IoU; small-object layer

check for updates

Citation: Zhang, F.; Dong, D.; Jia, X.; Guo, J.; Yu, X. Sugarcane-YOLO: An Improved YOLOv8 Model for Accurate Identification of Sugarcane Seed Sprouts. *Agronomy* **2024**, *14*, 2412. https://doi.org/10.3390/agronomy14102412

Academic Editor: Chengming Sun

Received: 19 September 2024
Revised: 13 October 2024
Accepted: 16 October 2024
Published: 18 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Driven by the need to enhance the accuracy of sugarcane seed sprout identification and considering the crucial role of intelligent sprout-cutting equipment in agricultural automation, this study proposes an improved YOLOv8 model aimed at precise sugarcane seed sprout recognition and efficient automation. As an essential economic crop, sugarcane increasingly demands accurate sprout identification for optimized planting and processing. As a kind of biological resource, sugarcane has shown rich raw material value in the fine processing of its various components. In addition to being a source of daily edible sugar, sugarcane can also be used in the preparation of a variety of products, such as pharmaceuticals, high-glucose syrup, and bioethanol. In addition, sugarcane bagasse and sugarcane mud also have a wide range of uses and can be used to produce a variety of fiberboard, paper, and other composite materials. Sugarcane planting is a labor-intensive production activity, and most areas still rely on artificial planting. At the same time, the development of mechanized sugarcane planting equipment in China is relatively backward, and the whole planting process includes multiple production processes, such as pre-seed-cutting, soil turning, seed arrangement, fertilization, and soil covering [1]. In the sugarcane planting process, in addition to a high degree of mechanization of sugarcane farming operations, other remaining sugarcane planting processes are still highly dependent on human labor, facing problems such as high labor intensity, high labor costs, and low production efficiency, which greatly hinder the progress of China's sugarcane planting industry. Promoting efficient sprout planting technology [2–7] and improving the sugarcane

planting mechanization level is not only an important step for agricultural modernization, but also a key measure to promote the process of agricultural modernization in our country.

In whole sugarcane, the weights of sugarcane sprouts and sugarcane stems account for about 30% and 70%, respectively. The traditional sugarcane planting method is to cut the whole sugarcane into several sections and plant the whole section of sugarcane containing three to four sugarcane sprouts into the soil. This planting method requires numerous seeds, extensive excavation of the planting area, and a large amount of sugarcane stems that are not needed for planting. It increases the cost of sugarcane cultivation and wastes the sugarcane stem that can be used for sugar extraction. The sugarcane sprouts and sugarcane stems are shown in Figure 1.

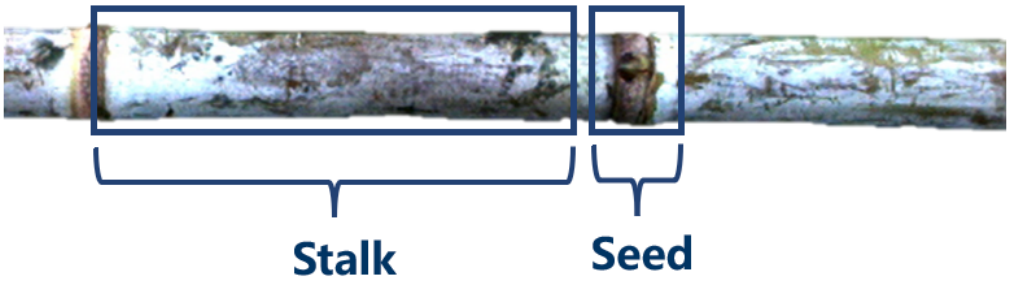


Figure 1. Sugarcane sprouts and sugarcane stems.

After the separation of sugarcane sprouts from sugarcane stems, sugarcane sprouts can be transported separately for farming, while sugarcane stems can be used for sugar extraction, which greatly reduces the transportation costs and the workload of farmers and increases the economic benefits of sugarcane. Meanwhile, the processes of sugarcane planting, sucrose production, and sugarcane sprout processing are separated, making it difficult to form a complete industrial chain of sugarcane planting and processing.

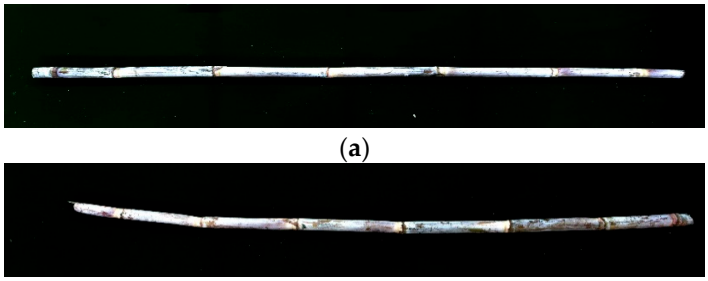
Machine vision technology is an important research topic in the field of artificial intelligence, and its core is machine learning recognition and detection algorithms. Machine learning has the ability to acquire knowledge and extract the original data model, which can give the computer the ability to act without explicit programming, and establish algorithms to identify data patterns and make predictions based on them [8]. Systems based on machine learning are applied in many fields, such as information analysis, agriculture, urban planning, and national defense [9–11]. Deep learning is one of the most widely used methods based on machine learning. An important feature of deep learning is its high abstraction and the ability to automatically learn the model presented in the image [12]. At present, the deep learning architecture with a convolutional neural network as the core is most widely used in image processing [13]. The first Iranian scholar, Huang et al. [14], conducted a preliminary study on the node recognition algorithm by using the method of gray image threshold segmentation, which includes convolution, thresholding, inversion, and table lookup operations. Then, Dong et al. [15] proposed a sugarcane node recognition method based on machine vision. They first extracted the R component of the RGB color space in the ROI region of the image and obtained the Sobel edge detection of the ROI region after median filtering, and they found that the gradient features were obvious near the sugarcane node. However, the features in the internode region were not obvious. According to these features, a rectangle detection operator was constructed. However, the recognition method was the image recognition method, the detection accuracy was greatly affected by environmental factors, the detection speed was slow, and the adaptability of migration application was poor. In 2021, Li et al. [16] proposed a field sugarcane node recognition method based on deep learning. YOLOv4 was used as the target detection algorithm. According to the influence of different light conditions, the recognition accuracy of sugarcane seed buds at different times of the day was studied. At present, it no longer has the advanced nature of the algorithm, and its recognition accuracy and recognition speed also have some limitations. Tang [17] and Yang et al. [18] used resistive strain gauges in

view of the bulge characteristics of sugarcane seed buds relative to the surface of sugarcane. The electrical signal generated by squeezing the strain gauge of seed buds was input to the microcontroller, and the single chip controlled the cutting knife for seed cutting. Sunkara and Luo [19] designed a sugarcane transportation and node detection device. The device transported sugarcane through the designed V-shaped roller, used the characteristics of the sugarcane node bulge, and used the contact displacement sensor to determine the node position. The accuracy rate of node detection was 95.4%. Du et al. [20] and Zhang et al. [21] designed a sugarcane node recognition and seed-cutting device based on the improved YOLOv3 recognition algorithm. The camera shot part was in the black box, the image features of the whole sugarcane node were located and identified in real time, and the PLC controlled the knife to complete the sugarcane node cutting and bud cutting.

2. Materials and Methods

2.1. Materials

As illustrated in Figure 2, in this study, Yun sugarcane 081609 in Honghe City, China and Guilin black sugar cane in Yunnan Province, China in 2023 were taken as the research objects. Considering the cost of the test and the amount of training required by the model, and to improve the universality of the model, 600 Yun sugarcane and 200 Guilin black sugar cane were selected in this study (sugarcane of 1.5 m-long after cutting was regarded as one). The water in sugarcane will be lost over time, and it is not easy to preserve for a long time. In the study, it was found that some sugarcane lost water due to the long transportation and storage time of materials, and this is a common phenomenon in the actual production of sugarcane. Therefore, in this section, different varieties and sugarcane samples with different storage times were selected for image collection. More comprehensive data support was provided to improve the universality of the model. It is noteworthy that the sugarcane material used in the production of this dataset was only used for the model study and could not be retained as the raw material for the equipment test, and the physical parameters of the test material were not measured.



(**a**)

(**b**)

Figure 2. Different kinds of test materials: (**a**) Yunnan sugarcane 081609 and (**b**) Guilin black sugar cane.

2.2. Image Acquisition of Sugarcane Seed Sprouts

In the dataset used in this study, the background of the sugarcane images was the actual background of the sugarcane on the laboratory prototype of the Modern Agricultural Engineering College of Kunming University of Science and Technology, which was the gray floor tile background at the interval between the surface of the synchronization belt sliding table and the sliding table of each synchronization belt. The material was obviously distinguished from the background. Using a model of MV-CA050-20GC, from China Hangzhou Hikvision Digital Technology Co., Ltd. (Hangzhou, China), Gigabit Ethernet industrial area array camera, and it had 5 megapixels, with a resolution of 2592 × 2048 pixels. The distance between the camera and the horizontal plane was 120 cm, and the height of the light source was set to 112 cm. The original image format of sugarcane images collected in this experiment was JPG. During image acquisition, each material was placed horizontally in the direction of the prototype conveyor channel, and image acquisition was performed

2–3 times for each material with different postures and different relative positions to obtain the required original images. After the elimination of the unqualified images due to blur, double shadow, and occlusion in the process of static photography, a total of 800 fresh sugarcane images (12,163 sprouts in total) after leaf stripping were obtained. The sugarcane material is shown in Figure 3 below.

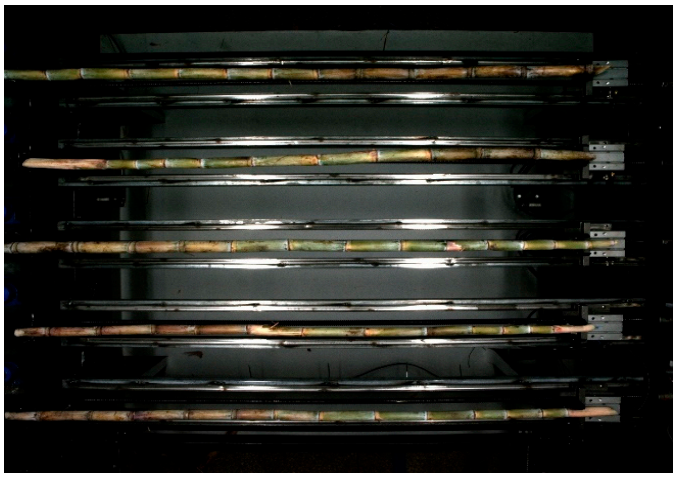


Figure 3. Sugarcane images in the dataset.

2.3. Making of Sugarcane Seed and Sprout Dataset

LabelImg software (1.8.6) was employed to label the sugarcane seed sprouts in each image, and the results were saved in XML format. The sugarcane dataset was divided into a training set (1280 images), a validation set (320 images), and a test set (320 images) at a ratio of 8:1:1. The original and enhanced images are shown in Figure 4 below.

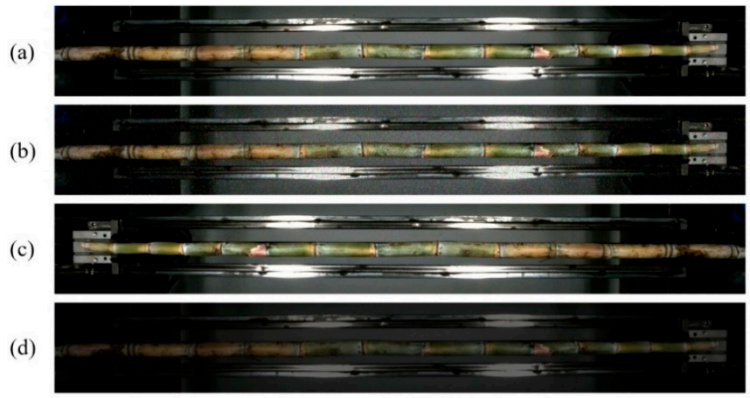


Figure 4. The original and enhanced images: (**a**) original image, (**b**) added noise, (**c**) mirror image, and (**d**) brightness transformation.

2.4. Establishment of the Sugarcane-YOLO Model for Sugarcane Seed and Sprout Identification

2.4.1. YOLOv8 Object Detection Network

Since its first release in 2015, the You Only Look Once (YOLO) family of computer vision models has been one of the most popular models in the field. It performs object detection and classification tasks simultaneously in a single neural network. YOLO uses a single forward pass for object detection, which can simultaneously predict all object categories and bounding boxes in an image in a single inference. YOLO has better real-time performance, fewer parameters, more scales of feature extraction, and a wide range of applications and community support due to its open-source nature, making YOLO one of the most popular object detection algorithms. The YOLO series target detection algorithm has been widely used in agricultural material recognition, so we chose the YOLO series

model with high recognition accuracy, high recall rate, fast recognition speed, and suitable for embedded systems. Specifically, the latest YOLOv8 model was used as the basic model for sugarcane seed buds' recognition model research. This choice aimed to ensure efficient and reliable identification of sugarcane seed buds in specific scenarios to meet practical application requirements.

YOLOv8 is the latest version of YOLO released by Ultralytics, the publisher of YOLOv5. It can be used for object detection, segmentation, classification tasks, as well as learning on large datasets, and can be executed on a variety of hardware, including cpus and Gpus. As a cutting-edge and most advanced (SOTA) model, YOLOv8 improves on the previous successful YOLO version, learns from the design advantages of YOLOv5, YOLOv5, YOLOX, and other models, and comprehensively improves the structure of the YOLOv5 model. While maintaining the advantages of YOLOv5's engineered simplicity and ease of use, we introduced new features to further improve the performance and flexibility, making it an excellent choice for object detection, image segmentation, and classification tasks.

Compared with YOLOv5, YOLOv8 has a new backbone network, a new anchor-free detection head, and a new loss function. It also supports previous versions of YOLO to facilitate switching between different versions and performance comparisons. In terms of the backbone network, the idea of CSP is still used, and the SPPF module is used, but the C3 module in YOLOv5 is replaced by the C2f module with a richer gradient flow, which realizes further lightweight, and adjusts the number of channels for different scale models. In the head part of the model, the previous coupling head structure was replaced by the current mainstream decoupling head structure, and the classification head and detection head were separated. In terms of the anchor box, YOLOv8 abandons the previous anchor-base and uses the idea of anchor-free. In terms of the loss function, YOLOv8 uses VFL loss as classification loss and DFL loss and C-IoU loss as regression loss. At the same time, YOLOv8 also supports previous versions of YOLO, which facilitates switching between different versions and performance comparisons. YOLOv8 has five pre-trained models with different model sizes: n, s, m, l, and x, which can be flexibly applied to various requirements of object detection and image classification tasks.

In the research and experimentation of the sugarcane seed and sprout model, it was found that the YOLOv8s model exhibited excellent recognition performance for sugarcane seeds and sprouts. However, it still produced false detections, and the detection accuracy needed improvement. To further improve the recall rate, accuracy, and efficiency of the sugarcane seed sprout recognition model, we proposed an improved network, Sugarcane-YOLO, based on the YOLOv8s framework. The specific improvement was to introduce the simple attention mechanism (SimAM) into the feature extraction network of the YOLOv8s model. The SimAM focuses on the features of the target region and improves the performance of the model with a small number of additional parameters. In the convolution module, the spatial-depth convolution (SPD-Conv) was used to replace part of the original convolution modules of the feature extraction layer. In terms of the loss function, the E-IoU (Efficient-IoU) loss function was utilized to replace the C-IoU (Complete-IoU) loss function in the YOLOv8s original model. Finally, the model's detection head was adjusted specifically for the small target layer to further improve its detection accuracy and precision.

2.4.2. SimAM Module

The SimAM [22] is a lightweight self-attention mechanism. It has a similar network structure as Transformer, but it uses linear layers instead of dot products to calculate attention weights. The network structure is as follows: input sequence > embedding layer > dropout layer > multi-layer SimAM layer > fully connected layer > SoftMax layer > output result. The SimAM layer consists of the following parts:

- Multi-head attention layer: The input sequence is divided into multiple heads through multiple linear maps, and each head calculates the attention weight.
- Residual connection layer: The output of the multi-head attention layer is added to the input sequence to avoid information loss.

- Forward-pass layer: The output of the residual connection layer is processed by linear transformation and activation functions and then added to the output of the residual connection layer.
- Normalization layer: The output of the forward-pass layer is normalized to speed up training and enhance model performance.

Current attention modules are usually integrated into each block to improve the output from the previous layer. This refinement step typically operates on the channel dimension (a) or the spatial dimension (b), generating 1D or 2D weights and treating neurons in each channel or spatial location equally. In channel-wise attention, 1D attention assigns different importance to different channels while treating all spatial positions equally. In spatial-wise attention, 2D attention assigns different importance to different spatial positions while treating all channels equally. However, this processing has limited ability to learn more discriminative cues. Hence, 3D weight attention (c) outperforms traditional 1D and 2D weight attention. The principle of the attention module is shown in Figure 5.

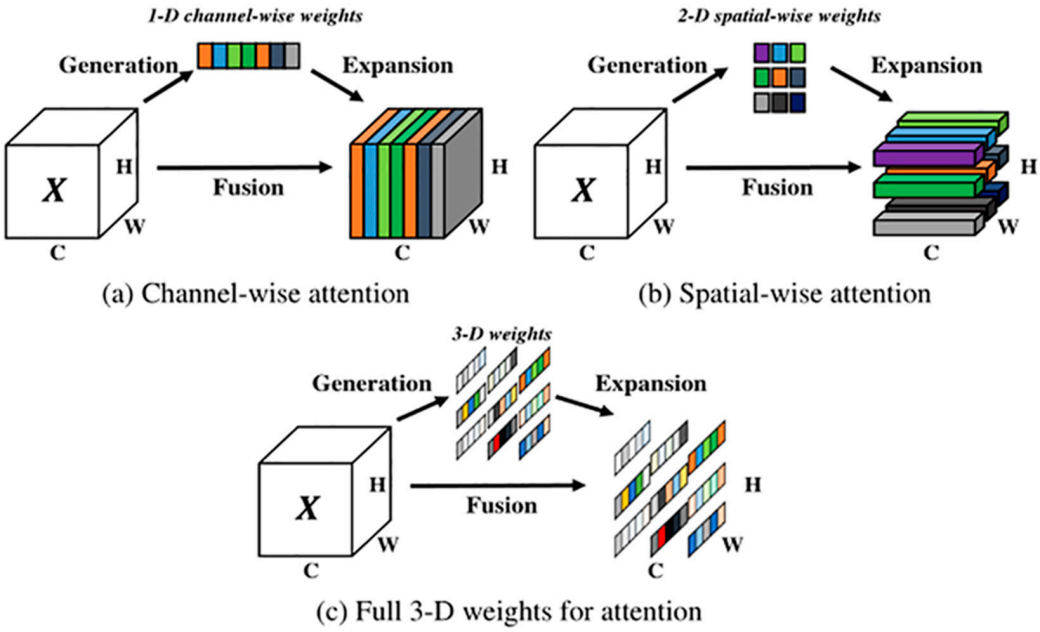


Figure 5. The principle of 3D weight attention.

The purpose here was to capture the attention distribution more comprehensively and accurately, providing the model with a more detailed focus of attention. The simplest method to identify important neurons is to measure the linear separability between neurons by defining an energy function, as follows:

$$et(wt, bt, \mathrm{y}, xi) = (yt - t^2)^2 + M - 11i = 1\sum M - 1(yo - x^i)^2 \quad (1)$$

Minimizing the above formulation is equivalent to improving the linear separability between neuron (t) and other neurons in the same channel. Since all neurons in each channel follow the same distribution, we computed the mean and variance of the input features across the two dimensions, (H) and (W), to avoid redundant computations:

$$et \times \frac{4(\sigma^2 + \lambda)}{(t - \mu)2 + 2\sigma^2 + 2\lambda} \quad (2)$$

The above formula means that the lower the energy, the more different the neuron t is from the surrounding neurons, and the higher the importance. Therefore, the importance

of a neuron can be obtained by: $1/e * 1/e\hat{}*1/e *$. The overall formula of the process is given as:

$$\widetilde{\chi} = sigmoid\left(\frac{1}{E}\right) \odot X \tag{3}$$

The SimAM's lightweight design effectively reduces computational complexity, while its 3D weight attention mechanism allows for more precise extraction of image details. Additionally, the energy function enhances the ability to distinguish target features, thereby significantly improving the recognition accuracy and model performance.

2.4.3. Space-Depth Convolution

The performance of popular convolutional neural network models usually decreases rapidly when the object is relatively small or has a low resolution. This is due to the continuous down-sampling of the feature map by the convolutional and pooling layers in the network structure. As the number of network layers increases, there is a risk of losing fine-grained information, which hampers effective feature extraction and learning. To solve this problem, we introduced the SPD-Conv [23] convolution module. SPD-Conv is composed of a spatial-depth (SPD) layer and a non-stride convolution layer, replacing the cross-convolution layer and pooling layer, respectively. It can be applied to most convolutional neural networks to assist in solving the problem of information loss in challenging object detection tasks, particularly those involving low resolution and small object sizes. Its specific principle is demonstrated in Figure 6.

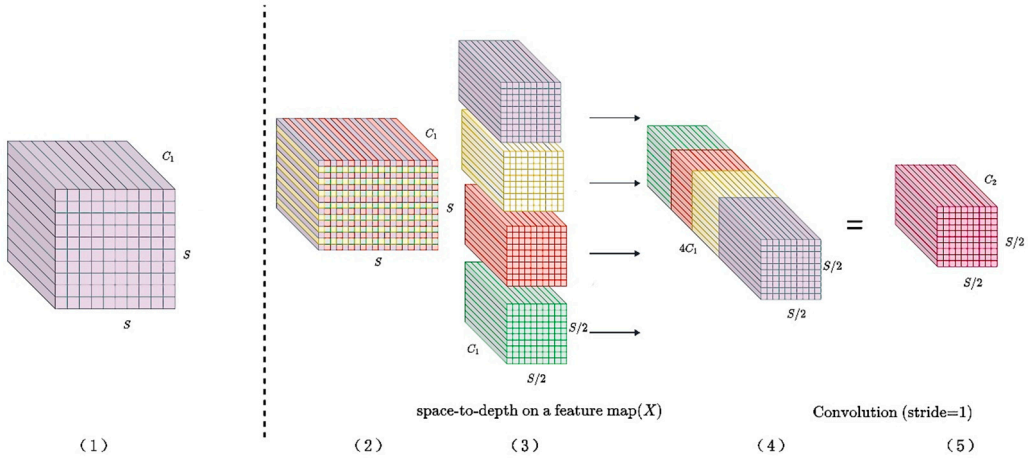


Figure 6. The principle of the SPD-Conv module. (1): Initial feature map, S × SS\times SS × S, C1C_1C1 channels. (2): Space-to-depth transformation, splitting into blocks. (3): Spatial size reduced to S/2S/2S/2, channels increased to 4C14C_14C1. (4): Concatenated feature map, S/2 × S/2S/2\times S/2S/2 × S/2, 4C14C_14C1 channels. (5): Convolution applied, channels reduced to C2C_2C2.

SPD-Conv is an effective architectural unit for dealing with small targets or low-resolution targets. It can retain the fine-grained information that was previously lost to the greatest extent and improve the learning ability of the model for small targets, such as sugarcane seed sprouts and their detailed representations, thereby improving the detection performance of the model. It is suitable for the identification and detection of sugarcane seed sprouts.

2.4.4. E-IoU Loss Function

An important index commonly used in object detection tasks is called intersection over union (IoU), also known as the intersection over union ratio. IoU is one of the key indicators to measure the performance of the target detection model. It describes the overlap between the bounding box predicted by the model and the real target bounding box. In the training

stage, IoU can be used as the basis for dividing positive and negative samples. It can also be used as a loss function. The IoU is calculated by dividing the intersection of the prediction box (A) and the real box (B) by the union of the two. The higher the value of IoU, the higher the degree of overlap between box A and box B, and the higher the accuracy of model prediction. On the contrary, the lower the IoU, the worse the prediction accuracy.

In terms of the loss function, the E-IoU (Efficient-IoU) [24] loss function was used to replace the C-IoU (Complete-IoU) [25] loss function used in the YOLOv8 original model. Compared with the Complete-IoU (C-IoU) loss function adopted in the original YOLOv8 model, the loss term of the aspect ratio was split into the difference between the predicted width and height and the width and height of the minimum external rectangle, respectively. This splitting sped up the convergence rate of the model and improved the regression accuracy, while the model could more accurately adjust the size of the bounding box to better match the shape and scale of the target. Focal loss was introduced to optimize sample imbalance in the boundary frame regression task; that is, to reduce the contribution of a large number of anchor frames with little overlap with the target frame to the optimization of B box regression, so that the model could focus more on samples that were difficult to classify, thus improving the model's attention to difficult samples and focusing the regression process on high-quality anchor frames.

The original YOLOv8 model uses the C-IoU loss function, which minimizes the distance between the predicted box's center point and the actual box to accelerate convergence. Also, the loss function considers the horizontal and vertical ratio of the anchor box, so it can effectively handle cases where the target anchor box has the same area but different aspect ratios. However, the aspect ratio difference reflected by C-IoU does not accurately represent the real differences in width and height and their confidence levels, and this may have a negative impact on the model's optimization similarity. To address these issues in C-IoU, scholars proposed the E-IoU loss function. The penalty term of E-IoU is based on that of C-IoU, but it separates the impact factors of the aspect ratio to independently calculate the length and width of the target box and the prediction box. Though the method for calculating C-IoU can be partially used for calculating the overlap loss and center distance loss, the width and height loss directly minimize the difference between the dimensions of the target box and the prediction box, resulting in faster convergence.

2.4.5. Small-Object Detection Layer

The YOLOv8 model uses the feature pyramid network (FPN) [26] structure to enable the network to better understand objects and structures at different scales in the image. Feature pyramid is a computer vision technique used to process information at different scales. In tasks such as object detection and image segmentation, features at different scales play a vital role in accurate detection and segmentation.

The FPN introduces multi-resolution feature maps into the network to capture different levels of semantic information in the image. These feature maps are obtained by performing convolution operations at different levels of the network. The low-level feature maps usually contain more detailed information, while the high-level feature maps have higher-level semantic information.

The main goal of the feature pyramid is to solve the problem that objects appear at different scales in the image. In object detection, smaller objects could be detected more easily in higher-level feature maps, while larger objects could be detected more easily in lower-level feature maps. The feature pyramid integrates information at multiple scales to improve the detection performance of the model for objects at different scales. The principle of the feature pyramid is shown in Figure 7.

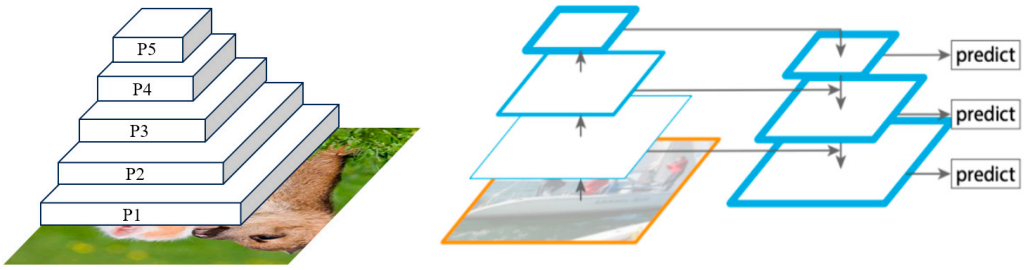


Figure 7. The principle of the feature pyramid.

The small-object detection layer refers to the specific network layer used to detect small-size objects in object detection tasks. Since small objects have a small size and a low resolution, they tend to be more difficult to detect and localize. The detection speed and accuracy of the YOLOv8 algorithm are balanced, but the detection effect of small targets is not good. The original YOLOv8 model uses three detection layers, P3, P4, and P5, for object detection by default. The size of the detection feature map corresponding to the P3 layer is 80×80 for detecting small objects, the size of the detection feature map corresponding to the P4 layer is 40×40 for detecting medium objects, and the size of the detection feature map corresponding to the P5 layer is 20×20 for detecting large objects. However, for smaller targets, fewer effective features can be extracted, which are easier to ignore or misjudge. To address this issue, a small target detection layer, P2, was introduced into YOLOv8 to make the model more effectively detect small target objects, thereby achieving higher precision and accuracy of small-object detection.

On the basis of adding the small-object detection layer P2, the model as a whole consists of four layers of detection output. The increase in model complexity leads to a decrease in model detection speed and an increase in the number of training parameters. Considering the relative image pixel size of the sugarcane seed sprouts in the sugarcane seed dataset, in this study, we removed the large-object detection layer, P5, and kept the three other detection layers, P2, P3, and P4. The YOLOv8 model was optimized into a detection model specifically for small and medium-sized targets, making it more suitable for detecting sugarcane seed sprouts. This optimization helped to improve the model's detection accuracy while maintaining the model's simplicity.

2.4.6. Sugarcane-YOLO Network Structure

In this study, to enhance the performance of the sugarcane seed sprout detection model, YOLOv8s was used as the foundational model, and several key improvements were made to its network structure:

- SIM-C2f module: The SimAM module was integrated into the C2f module of the YOLOv8s network, forming the SIM-C2f module to achieve a stronger image feature extraction capability.
- SPD-Conv module: The SPD-Conv module was used to replace the standard convolution in the backbone network of YOLOv8s, further enhancing the feature extraction of sugarcane seed sprouts.
- E-IoU loss function: The E-IoU loss function was used to replace the C-IoU loss function in the original YOLOv8s model, aiming to improve the accuracy of the prediction box regression.
- Detection layer adjustments: The small-object detection layer was added, and the large-object-layer output of the backbone network was removed. These adjustments refined the original model network structure to make it more suitable for the detection of small and medium-sized objects.

The improved Sugarcane-YOLO model, as illustrated in Figure 8, showcases these improvements, which collectively contributed to a higher detection accuracy without increasing the model's complexity.

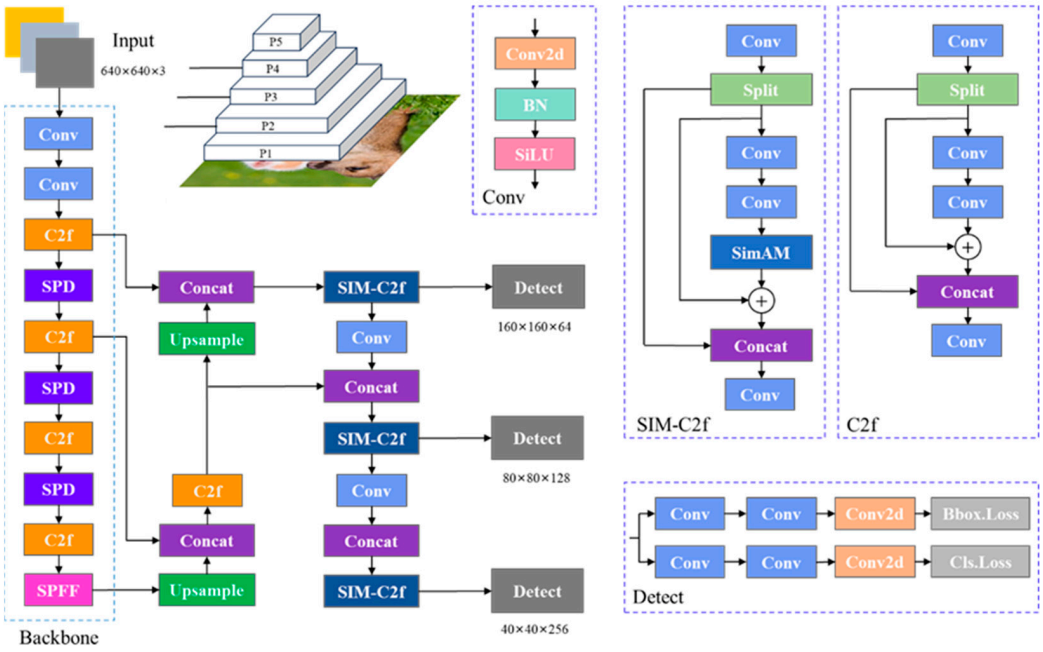


Figure 8. The structure of the Sugarcane-YOLO model.

2.5. Test Environment and Parameter Settings

All experiments are equipped with a core i7-9700F CPU (3.70 GHz) from the US Intel Corporation (Santa Clara, CA, USA) and a Quadro P5000 GPU (16 GB) from the US NVIDIA Corporation VRAM, Santa Clara, CA, USA, 2560 CUDA core) on the P2419h workstation from Dell Technologies Inc. a 64-bit Windows 10 system, PyTorch 1.8, and Python 3.8. The training settings included an input image size of 640 × 640 pixels, a batch size of 32, and training for 200 epochs, with weight files being saved every 10 epochs. The weight leading to the highest accuracy on the validation set was used for testing. Transfer learning was applied by using YOLOv8's features pre-trained on the VOC 2012 dataset. Additionally, the initial learning rate was 0.01, the weight decay was 0.0005, and the Adam optimizer was utilized to optimize the loss function.

2.6. Evaluation Index

To realize the identification and detection of sugarcane seed sprouts, the accuracy and real-time speed of the model should be considered. In this study, the size of the model network structure, the number of network parameters, and the number of real-time detection frames (FPS) were taken as the evaluation indicators of the model detection speed [27], where FPS represents the number of images detected by the model within 1 s. Recall (R), precision (P), and average precision (AP, %) were used as the evaluation indicators of model detection accuracy. The AP value is related to the precision and recall of the model, and it is used to measure the pros and cons of the model in all categories, while the mAP is the result of summing the AP values of all categories in the dataset and taking the average value. The detection problem considered in this study belonged to single-class target detection, and the m value was 1, i.e., AP and mAP had the same meaning in this study. The calculation of the evaluation indicators is given in Equations (4)–(7):

$$P = \frac{TP}{TP + FP} \times 100\% \tag{4}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{5}$$

$$AP = \int_0^1 P(R)dR \times 100\% \tag{6}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \times 100\% \quad (7)$$

where n represents the number of categories in the dataset, i denotes a specific category, TP refers to the number of true positives, FP stands for the number of false positives, and FN indicates the number of false negatives.

For image classification tasks, mAP aims at class recognition accuracy. In object detection tasks, there is a border regression task. The accuracy of anchor boxes is generally measured by the intersection over union (IoU), and the IoU threshold needs to be set in the task. Generally, the IoU threshold is set to 0.5 by default; that is, if the intersection over union ratio of the predicted box and the true box is greater than 0.5, the positive sample (TP) can be determined. The mAP value obtained under this condition was denoted as mAP50, and mAP75 was the mAP value obtained under the condition of the IoU threshold set to 0.75. mAP50-95 was the average value of the intersection ratio starting at 0.5 at 0.05 intervals until it reached 0.95. mAP75 and mAP50-95 are regarded as important evaluation indicators on many datasets, which emphasize higher detection accuracy and less false detection probability. In this study, the detection accuracy of sugarcane seed buds was an important factor affecting the accuracy of bud cutting. Therefore, the mAP50, mAP75, and MAP50-95 detection accuracy indicators were used as evaluation indicators for this model study, so as to conduct model comparison and analysis and study the optimal model for seed bud detection.

3. Results and Discussion

3.1. Comprehensive Comparison of Different Attention Mechanisms

To explore the influence of different attention mechanism modules on the performance of the YOLOv8 model and find the most suitable attention mechanism module for the sugarcane seed and sprout dataset, in this study, we introduced the current superior attention mechanism modules, including CBMA [28], SE [29], ECA [24], SimAM, and CA, into the FPN structure. Each attention mechanism module was placed at the same position in the YOLOv8 model, and the specific position was after the convolution module at the output position of the C2f module. Since AP and mAP had the same meaning in this study, the test results were expressed uniformly in AP. The trial results of different attention mechanisms are shown in Table 1.

Table 1. Experimental results of different attention modules.

Methods	Precision/%	Recall/%	AP50/%	AP75/%	AP50-95/%
YOLOv8s	96.55	98.07	98.69	78.99	70.29
YOLOv8s–CBAM	96.87	98.11	98.92	79.07	70.39
YOLOv8s–SE	96.79	98.2	98.9	79.09	70.36
YOLOv8s–ECA	96.94	98.21	98.79	78.31	70.01
YOLOv8s–SimAM	96.89	98.26	98.97	79.43	70.56
YOLOv8s–CA	96.84	98.26	98.92	78.32	70.08

Table 1 shows that the improved YOLOv8s based on the SimAM achieved the best effect. Compared to the original YOLOv8s model, the accuracy of the improved model incorporating the ECA attention mechanism module was increased by 0.49%, reaching 97.04%. This improvement was 0.17, 0.25, 0.15, and 0.1% higher than that of the CBAM, SE, ECA, and CA modules, respectively. In terms of the recall rate, adding the SimAM and CA modules led to the same improvement range, both of which were 98.26%, which was 0.19% higher than that of the original model. Compared with the AP50 of the original model, the addition of the five attention mechanism modules led to different degrees of improvement. Compared with other attention mechanism modules, the SimAM module had the most obvious improvement. Compared to the original YOLOv8s model, the accuracy of the improved model incorporating the ECA attention mechanism module was increased by 0.28%. The AP75 values for the models with ECA and CA modules were

slightly lower than that of the original model. However, adding the CBAM, SE, and SimAM modules improved the AP75 value by 0.08, 0.1, and 0.44%, respectively. Additionally, the addition of ECA and CA modules resulted in reduced values in AP50-95 indicators, indicating that these modules were not suitable for the model trained on this dataset. The SimAM demonstrated the most significant improvement in the AP50-95 indicators, reaching 70.56%. To sum up, besides the accuracy improvement with the ECA attention mechanism, the SimAM provided the largest improvement in recall rate, AP50, AP75, and AP50-95 indicators, demonstrating the best overall effect. This indicates that the SimAM module better distinguished between sugarcane seed sprouts and stems, thus improving the model's detection accuracy. Therefore, this paper introduced the SimAM to comprehensively enhance the YOLOv8s model.

3.2. Analysis of Ablation Test Results

The ablation test [25] is usually conducted in the design of complex neural networks to investigate the influence of some specific network substructures or training parameters on the training effect of the model, so it has important guiding significance in the design and improvement of complex neural network structures. In this paper, the SimAM module was introduced into the C2f structure of the YOLOv8s model. Meanwhile, SPD-Conv was incorporated into the feature extraction layer, and the E-IoU loss function was employed to further improve the model's performance. To verify the improvement effect of each module, in this paper, we conducted an ablation test, arranged and combined the improved modules in the improved YOLOv8s model one-by-one, and trained them. By using the control variable method, the influence of each module in model training was systematically studied to verify whether these improved modules had a positive significance in improving model performance. The results of the ablation tests are listed in Table 2.

Table 2. Ablation experiment results.

ID	Model	SimAM	SPD-Conv	E-IoU	P/%	R/%	AP50/%	AP75/%	AP50-95/%
1	YOLOv8s	×	×	×	96.55	98.07	98.69	78.99	70.29
2	YOLOv8s	√	×	×	96.89	98.26	98.97	79.43	70.56
3	YOLOv8s	×	√	×	96.81	98.15	98.85	78.58	70.29
4	YOLOv8s	×	×	√	97.05	98.35	98.93	79.12	70.37
5	YOLOv8s	√	√	×	97.37	98.63	99.03	79.69	70.85
6	YOLOv8s	√	×	√	97.13	98.39	98.98	80.08	70.56
7	YOLOv8s	×	√	√	96.89	98.26	98.97	79.43	70.56
8	YOLOv8s	√	√	√	97.19	98.89	99.03	80.38	71.32

Note: √ means that the structure was used in the current network structure, and × means that the structure was not used in the current network structure.

Table 2 shows that the model performances after introducing a single module and introducing partial combination modules were different. Among them, the first group was the original model, the second group was the introduction of a single SimAM module, the third group was the introduction of a single SPD-Conv module, and the fourth group was the change of the E-IoU loss function alone. The results demonstrated that the three improvement methods enhanced the model's accuracy and recall, as well as mAP50, mAP75, and mAP50-95 values. In contrast, the improvement of model detection accuracy and precision by adding the SPD-Conv module alone was not as obvious as that of adding the SimAM module and using the E-IoU loss function. The fifth group was the introduction of the SimAM module and SPD-Conv module at the same time. The precision, recall, as well as mAP50, mAP75, and mAP50-95 values were 0.48, 0.37, 0.06, 0.26, and 0.29% higher than those of the two groups adding the SimAM module alone. Compared with the three groups adding the SPD-Conv module alone, the proposed method improved the detection accuracy and precision by 0.56, 0.48, 0.18, 1.11, and 0.56%, respectively. Compared with adding one of the improvement modules alone, the combination of the SimAM module and SPD-Conv module improved the detection accuracy and precision. Among them,

the precision, recall, and AP75 had the most obvious improvements. The sixth group of experiments was the combination of the SimAM module and using the E-IoU loss function. The seventh group was the combination of the SPD-Conv module and using the E-IoU function. The eighth group was the combination of the SimAM module, the SPD-Conv module, and using the E-IoU function, and this was the group with the best improvement effect. Compared with the original YOLOv8s model, the accuracy, recall, and precision indicators, AP50, AP75, and AP50-95, were increased by 0.64, 0.82, 0.34, 1.39, and 1.03%, respectively. It can be seen from the results that each improvement module played a positive role to varying degrees in the model.

3.3. Small-Object Detection Layer Improvement

To investigate the influence of the output structure of different detection layers on the YOLOv8 model so that the model could detect and locate small targets more acutely, the YOLOv8s model was improved by incorporating the SimAM, using the E-IoU loss function, and substituting part of the convolution module with the SPD-Conv module. This improved model served as the basic model, with the addition of a small-object detection layer (P2) and the removal of the large-object detection layer (P5). Then, the three detection layers, P2, P3, and P4, were retained, and the Sugarcane-YOLO model was obtained by model training. The results of the model trials with the addition of the small-object detection layers are presented in Table 3.

Table 3. Addition of small-object detection layer test results.

Methods	Precision/%	Recall/%	mAP50/%	mAP75/%	mAP50-95/%	FPS
YOLOv8s	96.55	98.07	98.69	78.99	70.29	303
Improved YOLOv8s	97.19	98.89	99.03	80.38	71.32	263
Sugarcane-YOLO	97.42	98.63	99.05	81.3	71.61	294

As shown in Table 3, compared with the improved YOLOv8s model, the accuracy of the Sugarcane-YOLO model increased by 0.23 to 97.42%, and the recall slightly decreased to 98.63%. The three accuracy indicators of mAP50, mAP75, and MAP50-95 were improved, and the improvement in mAP75 and MAP50-95 was more obvious, which was 0.92 and 0.29%, respectively. Meanwhile, due to the increased model complexity, the FPS of the improved YOLOv8s model was lower than that of the original YOLOv8s model. Through the structural adjustment of the detection layer, the FPS of the Sugarcane-YOLO model was 31% higher than that of the improved YOLOv8s model. However, the detection speed was still slightly reduced compared to the original YOLOv8s model, but the reduction fell within an acceptable range. This demonstrates that the addition of the small-object detection layer could effectively differentiate sugarcane seed sprouts and stems, thereby enhancing the model's detection accuracy. Compared to the original YOLOv8s model, the accuracy, recall, and precision indicators (mAP50, mAP75, and mAP50-95) of the final Sugarcane-YOLO model were improved by 0.87, 0.56, 0.36, 2.31, and 1.32%, respectively. The addition of the small-object detection layer significantly improved the model's performance, particularly in terms of detection accuracy.

3.4. Comprehensive Comparison of Different Classical Models

To compare the detection performance of the improved model with that of commonly used classic models, the same datasets were used to train and test classic models, such as Faster R-CNN, VGG-SSD, YOLOv4, YOLOv5, and YOLOv6. The test results are shown in Table 4. This study had high requirements for the accuracy of cutting sugarcane seed sprouts, and low cutting accuracy would directly lead to an increase in the rate of damaged sprouts. However, the recognition speed of the model was affected by the operation of the equipment, the overall time consumption, and the one-time acquisition of all the seed sprout position information in the image, and then cutting one-by-one.

Table 4. Test results of different typical models.

Methods	Precision/%	Recall/%	mAP50/%	mAP75/%	mAP50-95/%	FPS	Size/M
Faster R-CNN	95.86	97.19	97.56	72.17	64.76	79	102
VGG-SSD	96.35	97.74	98.45	76.51	68.42	285	64
YOLOv5	96.63	97.57	98.46	77.19	69.17	277	18.2
YOLOv6	96.39	97.93	98.48	75.94	68.33	294	34.3
YOLOv8s	96.55	98.07	98.69	78.99	70.29	303	18.5
Sugarcane-YOLOv8s	97.42	98.63	99.05	81.3	71.61	294	22.4

As shown in Table 4, the performance indicators of the Faster R-CNN and VGG-SSD models were substantially lower than those of the YOLO series algorithms. This may be because the Faster R-CNN and VGG-SSD algorithms were proposed earlier, and their network structures and functional modules have limitations. YOLOv5 is an earlier algorithm proposed by the same author of YOLOv8. Compared with YOLOv5, YOLOv8 improved the overall structure and each module, so YOLOv8 achieved higher performance than YOLOv5. The original YOLOv5 model had the smallest size of only 18.5 M, but its FPS value was slightly lower than that of the YOLOv8s model, and its accuracy was significantly lower than that of the YOLOv8 model.

To sum up, compared with the current classical model, the Sugarcane-YOLO model proposed in this study had the best comprehensive effect, and its model size and detection speed were generally better and balanced. Its precision, recall, and precision indicators (mAP50, mAP75, and mAP50-95) were higher than those of other models. Therefore, it was the most suitable sugarcane seed sprout recognition model and provided effective and reliable theoretical and technical support for system research.

3.5. Real-Time Recognition Results

This section discusses a comprehensive analysis of the real-time recognition results using the improved Sugarcane-YOLO model. The objective was to evaluate the model's performance in accurately identifying sugarcane seed sprouts, which is a critical task for optimizing sugarcane planting and improving agricultural efficiency. The Sugarcane-YOLO model was employed to perform recognition and detection on the test set of the sugarcane seed and sprout dataset, and the detection results are shown in Figure 9.

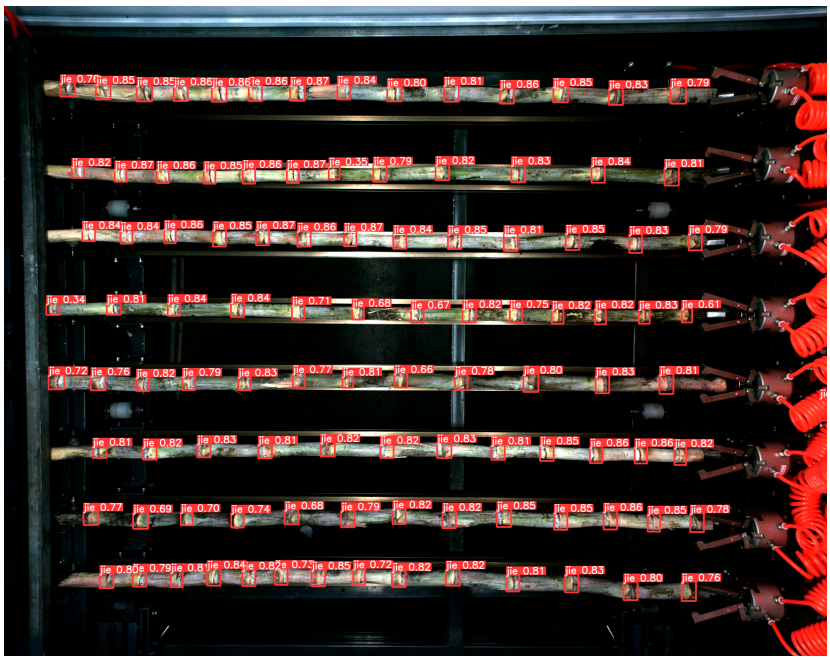


Figure 9. Model testing effect diagram.

It can be observed from Figure 9 that the model exhibited a strong recognition effect on sugarcane seed sprouts, with its detection frame closely fitting the primary parts of the sugarcane seed sprouts. This result provides a strong guarantee for the precise positioning and cutting of sugarcane seed sprouts, and growers can manage sugarcane seed sprouts more accurately, which provides important technical support for improving agricultural production efficiency and yield. The research results are highly significant for the development of the sugarcane planting industry and provide a reliable technical basis and referencing method for further research and practice in the field of sugarcane sprout cutting.

3.6. Discussion

The improved Sugarcane-YOLO model, through the integration of the SimAM and SPD-Conv modules, achieved significant improvements in both detection accuracy and speed. This not only enhanced the precision of sugarcane sprout detection but also reduced false detections, making the sugarcane planting process more efficient and automated. Compared to classic models, such as Faster R-CNN and YOLOv5, Sugarcane-YOLO excelled in small-object detection, achieving an mAP50 of 99.05% and an mAP50-95 of 71.61%. This indicated that the model could more accurately identify and cut sugarcane sprouts in practical applications, reducing resource waste.

The discussion should not be limited to the comparison and improvement of algorithm performance. The improved detection accuracy not only lays the foundation for the mechanization of sugarcane planting but also promotes the development of smart agriculture. The application of this technology can reduce the reliance on manual labor, improve production efficiency, and foster sustainable agricultural practices. By integrating existing literature, further discussions on the potential impacts of this technology, such as improving crop yield and quality, will help readers better understand and widely adopt this technology.

There are still some challenges in applying the model under different environmental conditions. Future research can focus on further optimizing the model to address these complex situations.

4. Conclusions

This study established a sugarcane seed sprout dataset and made several improvements to the basic model of YOLOv8s:

- SimAM: The SimAM module was added to the C2f module within the neck network, enhancing the model's ability to capture detailed image information.
- SPD-Conv module: The SPD-Conv module was integrated into the tail of the C2f module, boosting the efficiency of feature extraction.
- E-IoU loss function: The E-IoU loss function was utilized to speed up the model's regression process, improving training efficiency.
- Small-object detection layer: A small-object detection layer was added while removing the large-object detection layer, optimizing the model for better recognition accuracy of smaller targets.

The effectiveness of these improvements was validated through ablation experiments, which confirmed that incorporating the SimAM, replacing standard convolutions with SPD-Conv, and using the E-IoU loss function significantly enhanced the model's performance. The SimAM, in particular, outperformed other attention mechanisms, such as CBMA, SE, CA, and ECA. As a result, the improved Sugarcane-YOLO model achieved an accuracy of 97.42%, a recall rate of 98.63%, and mAP50, mAP75, and mAP50-95 values of 99.05, 81.3, and 71.61%, respectively. Compared to other mainstream models on the sugarcane seed sprout dataset, Sugarcane-YOLO showed a well-balanced combination of accuracy and detection speed, making it ideal for automatic sugarcane sprout-cutting applications.

To enhance the Sugarcane-YOLO model's performance, data augmentation can increase dataset diversity and robustness. Using transfer learning with larger pre-trained

models can improve feature learning and detection accuracy. Adding environmental data, such as soil moisture and temperature, provides valuable context and enhances precision.

Hyperparameter tuning can optimize the model for better accuracy and speed. It is also important to address variations in the sugarcane seed and sprout dataset, such as growth stages and lighting conditions, to ensure robustness in real-world applications. By focusing on these factors, the Sugarcane-YOLO model can achieve improved accuracy and adaptability in sugarcane cultivation.

Author Contributions: Conceptualization, F.Z. and D.D.; methodology, J.G.; software, J.G.; validation, J.G., X.J., and X.Y.; formal analysis, F.Z.; investigation, F.Z.; resources, F.Z.; data curation, D.D.; writing—original draft preparation, D.D.; writing—review and editing, D.D.; visualization, D.D.; supervision, J.G.; project administration, J.G.; funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Major Science and Technology Special Plan of Yunnan Province, research and development of intelligent agricultural machinery and equipment for mountain sugarcane, grant number 202402AE090013, and the APC was funded by Fujie Zhang.

Data Availability Statement: The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, Y. Study on the Current Situation and Development Strategy of Sugarcane Planting in Zhanjiang City. Master's Thesis, Guangxi University, Nanning, China, 2020.
2. Liu, H.; Qu, Y.; Zeng, Z. Investigation and analysis of sugarcane mechanization in Zhanjiang Agricultural Reclamation. *Mod. Agric. Equip.* **2013**, *6*, 28–33.
3. Huang, M. Study on Sugarcane production Mechanization in Guangxi. *Guangxi Agric. Mech.* **2014**, *4–7*, 10. [CrossRef]
4. Lv, W.; Zhang, X.; Wang, W. Agricultural machinery purchase subsidy, agricultural production efficiency and rural labor transfer. *China's Rural Econ.* **2015**, *8*, 22–32.
5. Liu, H. Focusing on improving the service capacity and supply level of technology extension to provide strong support for the development of agricultural mechanization during the "13th Five-Year Plan"—A speech at the 2016 national conference of agricultural machinery extension station owners (abstract). *Agric. Mach. Technol. Promot.* **2016**, *10*, 9–12. [CrossRef]
6. Luo, X.; Liao, J.; Hu, L.; Zang, Y.; Zhou, Z. Agricultural mechanization and sustainable development. *Trans. Chin. Soc. Agric. Eng.* **2016**, *32*, 1. [CrossRef]
7. Huang, M. Application of whole process mechanization technology in sugarcane Production. *Agric. South* **2017**, *11*, 16–18. [CrossRef]
8. Jiang, T.; Cui, H.; Cheng, X. A calibration strategy for vision-guided robot assembly system of large cabin. *Measurement* **2020**, *163*, 107991–108000. [CrossRef]
9. Chen, J.; Ma, B.; Ji, C.; Zhang, J.; Feng, Q.; Liu, X.; Li, Y. Apple inflorescence recognition of phenology stage in complex background based on improved YOLOv7. *Comput. Electron. Agric.* **2023**, *211*, 108048. [CrossRef]
10. Jing, J.; Zhang, S.; Sun, H.; Ren, R.; Cui, T. YOLO-PEM: A lightweight detection method for young "Okubo" peaches in complex orchard environments. *Agronomy* **2024**, *14*, 1757. [CrossRef]
11. Abdullah, A.; Amran, G.A.; Tahmid, S.A.; Alabrah, A.; AL-Bakhrani, A.A.; Ali, A. A deep-learning-based model for the detection of diseased tomato leaves. *Agronomy* **2024**, *14*, 1593. [CrossRef]
12. Márquez-Grajales, A.; Villegas-Vega, R.; Salas-Martínez, F.; Acosta-Mesa, H.-G.; Mezura-Montes, E. Characterizing drought prediction with deep learning: A literature review. *MethodsX* **2024**, *13*, 102800. [CrossRef] [PubMed]
13. Buayai, P.; Yok-In, K.; Inoue, D.; Nishizaki, H.; Makino, K.; Mao, X. Supporting table grape berry thinning with deep neural network and augmented reality technologies. *Comput. Electron. Agric.* **2023**, *213*, 108194. [CrossRef]
14. Huang, Y.; Wang, X.; Yin, K.; Huang, M. Design and experiment of cutting and preventing shoot of sugarcane seed based on induction counting. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 41–47. [CrossRef]
15. Dong, Z.; Shen, D.; Wei, J.; Meng, Y.; Ye, C. Design and research of sugarcane seed transport and sugarcane node detection device. *J. Guangxi Univ. (Nat. Sci. Edit.)* **2017**, *42*, 979–989. [CrossRef]
16. Li, S.; Li, X.; Zhang, K.; Li, K.; Yuan, H.; Huang, Z. Improved YOLOv3 network to improve the efficiency of real-time dynamic identification of sugarcane node. *Trans. Chin. Soc. Agric. Eng.* **2019**, *35*, 185–191. [CrossRef]
17. Tang, L. Research on Sugarcane Node Recognition and Cutting Based on Convolutional Neural Network. Master's Thesis, Agricultural University, Hefei, China, 2021.
18. Yang, L.; Zhang, R.; Li, L.; Xie, X. SimAM: A simple, parameter-free attention module for convolutional neural networks. In Proceedings of the Proceedings of the 38th International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 11863–11874.

19. Sunkara, R.; Luo, T. No more strided convolutions or pooling: A New CNN building block for low-resolution images and small objects. In Proceedings of the Machine Learning and Knowledge Discovery in Databases, Grenoble, France, 19–23 September 2023; pp. 443–459.
20. Du, S.; Zhang, B.; Zhang, P.; Xiang, P. An improved bounding box regression loss function based on CIOU loss for multi-scale object detection. In Proceedings of the 2021 IEEE 2nd International Conference on Pattern Recognition and Machine Learning (PRML), Chengdu, China, 16–18 July 2021; pp. 92–98.
21. Zhang, Y.; Ren, W.; Zhang, Z.; Jia, Z.; Wang, L.; Tan, T. Focal and efficient IOU loss for accurate bounding box regression. *Neurocomputing* **2022**, *506*, 146–157. [CrossRef]
22. Meng, Z.; Du, X.; Xia, J.; Ma, Z.; Zhang, T. Real-time statistical algorithm for cherry tomatoes with different ripeness based on depth information mapping. *Comput. Electron. Agric.* **2024**, *220*, 108900. [CrossRef]
23. Yu, C.; Feng, J.; Zheng, Z.; Guo, J.; Hu, Y. A lightweight SOD-YOLOv5n model-based winter jujube detection and counting method deployed on Android. *Comput. Electron. Agric.* **2024**, *218*, 108701. [CrossRef]
24. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 11531–11539.
25. Peng, S.; Jiang, W.B.; Pi, H.; Bao, H.; Zhou, X. Deep snake for real-time instance segmentation. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 24 June 2020; pp. 8530–8539.
26. Lin, T.Y.; Dollar, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 936–944.
27. Li, H.; Li, C.; Li, G.; Chen, L. A real-time table grape detection method based on improved YOLOv4-tiny network in complex background. *Biosyst. Eng.* **2021**, *212*, 347–359. [CrossRef]
28. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional block attention module. In Proceedings of the Computer Vision—ECCV 2018, Munich, Germany, 8–14 September 2018; pp. 3–19.
29. Hu, J.; Shen, L.; Sun, G.; Albanie, S. Squeeze-and-excitation networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *42*, 2011–2023. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.