Pak. J. Bot., 58(9) September 2026.

ARTICLE | DOI: http://dx.doi.org/10.30848/PJB2026-9(9)

YOLOv11s-SEED PRECISION INTELLIGENT CONTROL PRODUCTION LINE FOR PLUG TRAY SEEDLING CULTIVATION

QUAN ZHANG[1,#,*], ZIYE LAI[2,#] AND HENG CHEN[1]

[1]School of Mechanical Engineering, Xijing University, Xi'an, 710123, China
[2]Chang'an University, Xi'an, 710018, China
[#]These authors contributed equally.

*Corresponding author's email: 15235259616@163.com

Abstract

To address the low automation and poor precision of traditional Plug Tray Seedling Cultivation (PTSC), this study developed a novel intelligent PTSC model integrating AI with mass production. The system incorporates a YOLOv11 large model, machine vision, deep learning, and the Internet of Things (IoT) into a unified production line. This integrated system automates the entire process, including soil spreading, hole pressing, sowing, watering, and visual inspection. Prototype testing and core functionality verification were conducted to evaluate the system's performance. Results demonstrate that the system achieves full automation, precise control, and real-time visualization throughout the PTSC workflow. The findings validate the feasibility of this technical solution, effectively enhancing cultivation efficiency and seedling quality. This research provides a robust framework for implementing intelligent, lean production in the commercial vegetable cultivation industry.

Key words: Plug tray seedling cultivation; Smart machinery; Prototype testing and evaluation; YOLOv11; Machine vision; Automation in agriculture

Introduction

With the increasing global demand for high-quality vegetables, the vegetable cultivation industry has become a cornerstone of modern agriculture. In China, vegetable sowing areas reached 23,029.78 thousand hectares in 2023 (Chen, 2023), with production exceeding 810 million metric tons. As the foundation of this industry, Plug Tray Seedling Cultivation (PTSC) has been widely adopted due to its potential for high efficiency and standardization (Chen *et al.*, 2024a). However, traditional PTSC primarily relies on manual labor or single-function machinery, facing significant challenges such as low automation levels, poor seeding precision, and high initial investment costs (Chen *et al.*, 2024b). These limitations hinder the large-scale promotion of high-efficiency seedling technology and increase the financial burden on farmers (Ding *et al.*, 2024; Han & Moraga, 1995).

Current research in agricultural mechanization increasingly focuses on integrating vision detection technology to improve the "seedling establishment rate. Traditional pellet seed number target detection algorithms usually include two stages: trait extraction and classification. First, images are fetched from the camera, and the destination area is detached. Then, trait extraction is conducted on the segmented area. Finally, the classifier classifies and recognizes these extracted traits.

Ma, *et al*. took advantage of the YOLOV8S model by introducing the concept of a shared convolutional layer to improve the YOLOv8 detection head (Hao *et al.*, 2023), reducing the parameter count, achieving a lightweight design, and enhancing the runtime speed. Secondly, a vision transformer with a deformable attention mechanism was incorporated into the C2f module of the backbone network to enhance the network's feature extraction capability and improve the detection accuracy.

Niu, *et al*. proposed a lightweight (Jocher *et al.*, 2023), small target corn seed quality detection model based on an improved YOLOV8: I-YOLOV8. First, an Efficient Multi-scale Attention (EMA) mechanism was introduced, which effectively preserved the information among multi-channels and reduced computation. Next, the Space-to-Depth Convolution (SPD-Conv) module was selected for low-resolution images and small objects and applied to the backbone, addressing the issues of fine-grained information loss and low efficiency in feature representation learning in YOLOV8. Finally, the large detection layers were reduced, enabling the network to better focus on the positioning, channel, and dimensionality information of smaller objects, and the loss function was replaced with WIoUv3.

Over the past decade, the application of deep learning in this realm has continued to deepen (Kang *et al.*, 2024). Li, *et al*. (Li *et al.*, 2024), based on the YOLOv7 framework, proposed the "Cotton-YOLO" algorithm, which integrates ConvNext and Swin-Transformer: the ConvNext module enhances the feature extraction capability of the backbone network, while the combination of Swin-Transformer and ConvNext in the detection head optimizes multi-scale feature fusion, significantly improving both detection accuracy and velocity. It should be noted that although multi-view decision fusion technology can improve classification accuracy, its dependence on multi-angle images limits its practical application scope.

Received: 14-10-2025 | Revised: 18-03-2026 | Accepted: 25-03-2026 | Online: 24-04-2026

While deep learning models like YOLOv8 and YOLOv7 have been explored for seed detection, they often struggle with low accuracy and limited recall when identifying pelleted seeds in complex environments, such as those with soil clods or dense distributions.

To overcome the inherent limitations of traditional detection algorithms in complex agricultural environments, this study develops YOLOv11-SEED, a specialized deep learning architecture optimized for intelligent PTSC production lines. The innovation of this model lies in its multi-stage feature enhancement strategy:

(1) Precision Enhancement: A Sparse Feature (SF) module is introduced to prioritize key target features, significantly reducing background interference.
(2) Contextual Integration: By incorporating a Dilated Reparam Block (DRB) and a Content-Guided Attention Fusion (CGAF) module (Li & Shi, 2024; Luo *et al.*, 2016), the model effectively expands the receptive field and implements multi-scale feature weighting, addressing the poor recall typical of conventional models.
(3) Robustness in Dense Scenarios: The integration of the SEAM Attention Mechanism and DySample Dynamic Up-sampling allows for precise differentiation between pelleted seeds and soil clumps, even in high-density seeding conditions.

Beyond algorithmic improvements, this system realizes a dual-optimization framework: it provides real-time seeding accuracy assessment while simultaneously validating the mechanical consistency of the production line. This synergy represents a significant advancement over existing standalone software solutions, offering a robust technical foundation for intelligent seedling cultivation.

Materials and Methodology

Dataset Foundation: To evaluate the efficacy of the proposed YOLOv11-SEED model, a dedicated image dataset was constructed, focusing on pelleted seeds within intricate soil environments.

Dataset collection: Data acquisition was conducted in June 2025 using a Hikvision EV-EM200M industrial camera and a high-resolution mobile camera to ensure device diversity. A total of 1,000 raw images were captured and stored in JPG and EMP formats. To simulate real-world complex seeding scenarios, the acquisition procedure incorporated several key variables:

(1) Hole Pressing Depth: Seeding holes with varying depths and shadows.
(2) Lighting Conditions: Variations in light intensity and angles to simulate different factory environments.
(3) Seeding Density: Different distributions of pelleted seeds within the plug tray.
(4) Perspective Diversity: Capturing seeds from multiple shooting angles.

Specifically, a monochromatic (black-and-white) imaging approach was adopted under varying exposure conditions to emphasize the morphological traits of the pelleted seeds against the soil background. This method significantly improved the precision of trait identification. The specific dataset is shown in Fig. 1.

Data preprocessing and augmentation: All captured images were manually annotated using Labeling, with the label "Seed" designated for each target. To enhance the model's generalization capacity and robustness against environmental noise (Ma *et al.*, 2024), the initial dataset was expanded from 1,000 to 3,000 images through several data augmentation techniques:

(1) Geometric Transformations: Horizontal flipping, vertical flipping, and center cropping were implemented to account for varying seed orientations.
(2) Color and Noise Adjustments: Random transformations of brightness, contrast, and saturation were applied, alongside the introduction of Gaussian noise (Niu *et al.*, 2024).
(3) Environmental Simulation: These operations accurately replicated the illumination fluctuations and complex background interference common in automated seedling production lines.

The augmented dataset provides a comprehensive basis for training the YOLOv11-SEED model, ensuring high reliability in dense and obscured seeding scenarios. The processed dataset is shown in Fig. 2.

YOLOv11-seed model: This study adopts the YOLOv11 model as the foundational architecture. The YOLO series is widely recognized for its efficient detection capabilities and end-to-end training characteristics. YOLOv11, the latest iteration in the series, was introduced by Rahima *et al*., incorporating architectural improvements to the YOLO framework. Figure 3 illustrates that YOLOv11 is derived from YOLOv8 (Wang *et al.*, 2024), retaining the SPPF module but incorporating the C2PSA and C3k2 modules.

YOLOv11 enhances object detection accuracy and robustness relative to earlier model generations by optimizing the network architecture and training algorithm. Under complex background interference, YOLOv11 exhibits exceptional capabilities in object localization and classification. The model demonstrates significant adaptability in detecting pelleted seeds within challenging soil environments, accurately identifying and quantifying pelleted seeds against dry soil backgrounds characterized by numerous soil clods, with notable anti-interference capabilities. Moreover, relative to other prevalent detection models, YOLOv11 demonstrates a notable advantage in computational efficiency, facilitating effective deployment on resource-limited edge devices while adequately satisfying stringent real-time detection demands. Figure 4 illustrates that the core architecture of YOLOv11 comprises four carefully structured modules that operate in unison: the input module, backbone network, neck network, and detection head. The collaborative tuning of these modules underpins the high performance of YOLOv11.

Despite advancements in object detection, YOLOv11 continues to encounter difficulties in effectively processing the main properties of small targets in Plug seeding images. Specifically, YOLOv11 is inadequate in identifying the essential characteristics of small objects, offering the requisite Receptive Field for accurate detection, and the feature extraction proficiency following Multi-scale Feature Fusion.

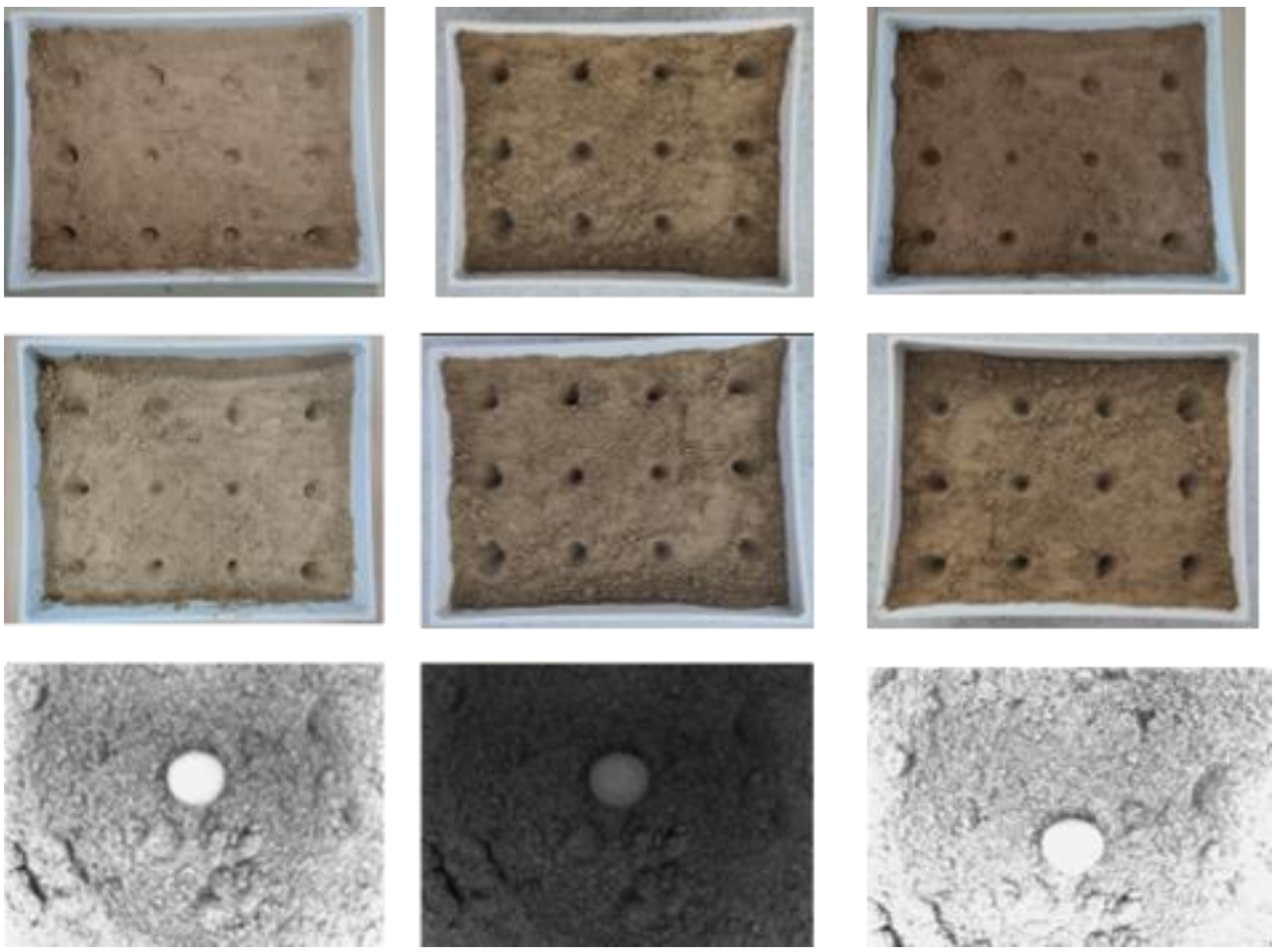

Fig. 1. Pelleted seed sowing image.

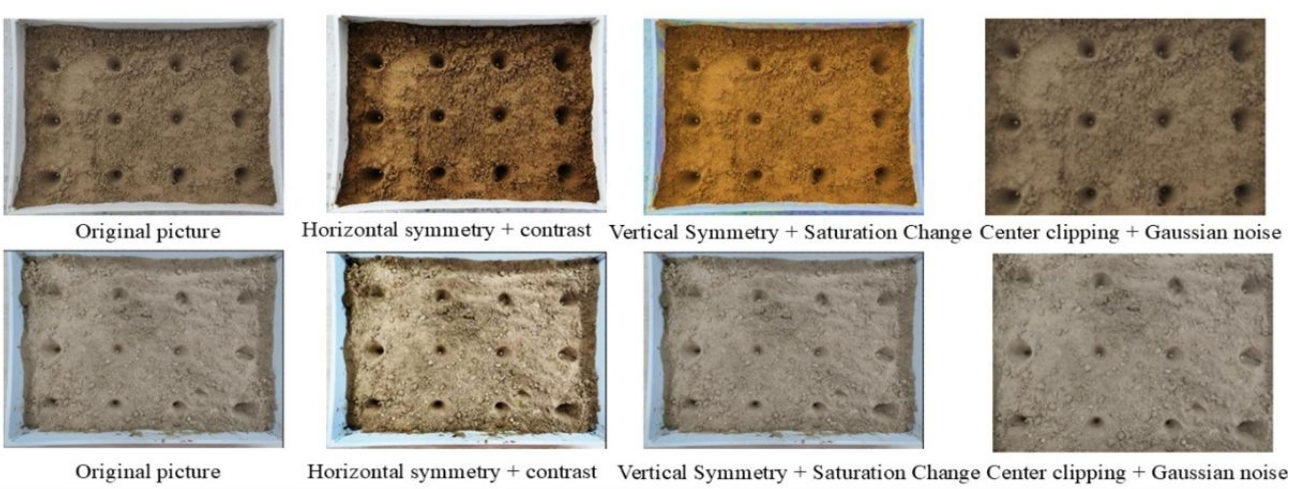

Fig. 2. Pelletized seed enhanced image.

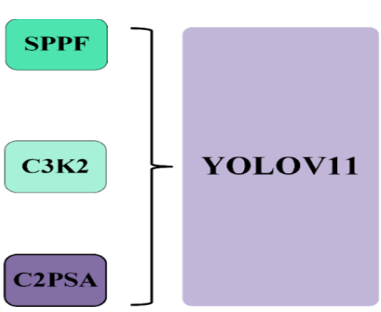

Fig. 3. Main modules of YOLOv11.

To get around these problems, a better version of the YOLOv11 model called YOLOv11-SEED was suggested. In YOLOv11-SEED, an innovative feature selection approach, SF, substitutes the SPPF module seen in YOLOv11. The SPPF module improves receptive field and feature representation using multi-scale pooling, however it does not explicitly tackle channel selection (Xin *et al.*, 2022). The SF module finds the channels that are most useful for distinguishing between different features during forward propagation. This cuts down on feature redundancy. Targeted channel selection is necessary because plug tray

pictures often include tiny target features and background noise. This allows the model to concentrate on key feature channels, improving feature contrast. The SF module enhances feature maps by preserving just the most relevant channels before to their processing in subsequent convolutional and detection layers. This feature sparsification method mitigates overfitting and improves the model's resilience in diverse maritime settings.

Within the C3k2 module, YOLOv11-SEED integrates the DRB module. By combining multi-scale dilated convolution with post-aggregation parameter reorganization, DRB merges several convolution branches into one large kernel convolution. This method broadens the receptive field, allowing the model to better capture fine target features inside the SSS picture.

Furthermore, YOLOv11-SEED incorporates the CGAF module. Traditional feature fusion approaches are noise-prone and fail to account for feature source significance fluctuation. Multi-attention mechanisms in the CGAF module adaptively weight channels, spatial locations, and pixels to emphasize target characteristics and minimize extraneous information.

At the same time, the C2PSS module is added to the original model to take the place of the C2PSA module. This module efficiently handles issues like inadequate sunlight and soil clods in pelleted seed identification by using feature extraction and processing via an attention method. To improve dense pelleted seed target identification, DySample replaces the up-sampler in the neck network. Figure 5 illustrates the architecture of the YOLOv11-SEED model.

C2PSS module: In tray seeding detection, the differing depths of the cells have emerged as a critical aspect influencing the model's accuracy. When seeds are planted at varying depths within the cells, the impact of light may cause YOLO11 to inaccurately identify the location or kind of the pelleted seed, resulting in detection inaccuracies. This article introduces an enhanced C2PSA module, termed C2PSS, which substitutes the attention mechanism in the PSABlock module with the Spatial Enhancement Attention Module (SEAM) (Xu *et al.*, 2020). Experimental findings indicate that the enhanced YOLO11 model proficiently addresses the issue of fluctuating cell pressing depths in intricate situations. Figure 6 illustrates the network architecture of the C2PSS module. Figure 7 illustrates the network architecture of the SEAM module.

A key element of the YOLO Model Backbone Network, the C2PSA Module significantly raises the model's detection accuracy and feature extraction capacity. The squeeze-and-excitation (SE) attention mechanism is the foundation of this module, which incorporates the pyramid slice attention (PSA) mechanism to improve feature extraction and representation. Pelletized seed detection occurs when the model struggles to distinguish pears and branches. The Spatial Enhancement Attention Module is used to fix this problem in place of the original attention method. SEAM blends Depth wise Separable Convolution with Residual Connection, reducing parameter count. To alleviate communication losses, the Conv module uses 1x1 Convolution to combine outputs from multiple depth convolutions. In order to effectively reduce information loss due to light, a two-tier Fully Connected Network is provided to explicitly enhance Inter-channel Relationship Modelling.

SF module: The SF module was created to address YOLOv11's failure to pay enough attention to tiny targets' important characteristics. By emphasising the most significant channels, the SF module improves the representation of crucial visual elements. This module makes the neural network better at focussing on targets by adding global attention, channel selection, and dimensionality reduction. Figure 8 depicts the design of the SF module, where H and W signify the dimensions of space of the input feature map and C denotes the number of input channels.

The SF module focusses on channel pruning by utilising global feature aggregation and assessing channel importance to enhance model performance. This module utilises Global Average Pooling (GAP) to consolidate the spatial information of each channel, resulting in a compact channel-level feature descriptor. The module further processes and adjusts these channel descriptors to accurately assess the importance weight of each channel, thereby enhancing the utilisation of the aggregated global features.

DySample sampler: Pelleted seeds depicted in the image generally display fewer characteristics, and conventional up sampling techniques may result in the loss of edge details of little items. Furthermore, the dense seeding pattern and the obstruction caused by small soil clods exacerbate the challenges in seed detection. This study presents a DySample sampler, grounded in point sampling, to improve the model's detection capability for small and densely clustered seeds.

By doing away with dynamic convolution, DySample considerably reduces the model's computing load. Unlike standard kernel-based upsampling approaches, it uses a simpler and more efficient point sampling strategy, resulting in fewer model parameters. Figure 9 depicts the network architecture of DySample. Panel 9(a) illustrates the sampling procedure with dynamic up sampling. The input feature map (x) creates a sampling set (S) using the sampling point generator, which is then resampled using the grid sampling function to produce the up sampled feature map.

The sample point generator in DySample uses both the static range factor and the dynamic range factor to generate sampling points, as shown in Fig. 9(b). The produced offset (O) and the starting sampling grid (G) make up the sample set (S) according to the static technique. The offset (O) is generated by a linear layer and pixel shuffle, and its range is influenced by both static and dynamic variables. The dynamic technique adds a linear layer, shuffles the pixels, and adds a dynamic range factor. This factor is first produced by a Sigmoid function and then used to adjust the offset (O).

The formulas used in these methods are as follows:

$$x' = grid_sample(x, S) \tag{1}$$

$$O_1 = 0.25[linear(x)] \tag{2}$$

$$O_2 = \text{linear1}(x) * 0.5\sigma[linear2(x)] \tag{3}$$

$$S = G + O \tag{4}$$

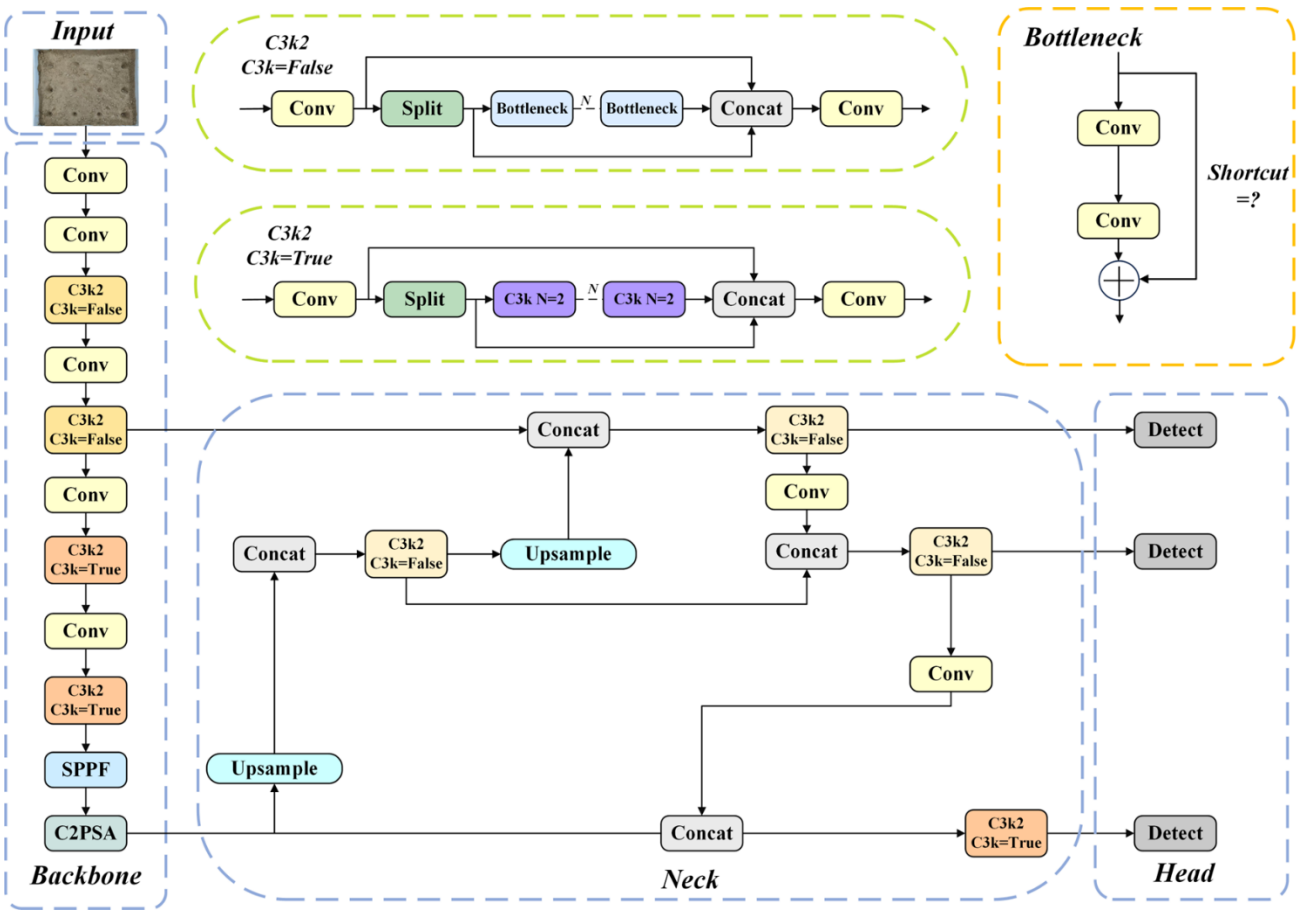

Fig. 4. YOLOv11 network architecture diagram.

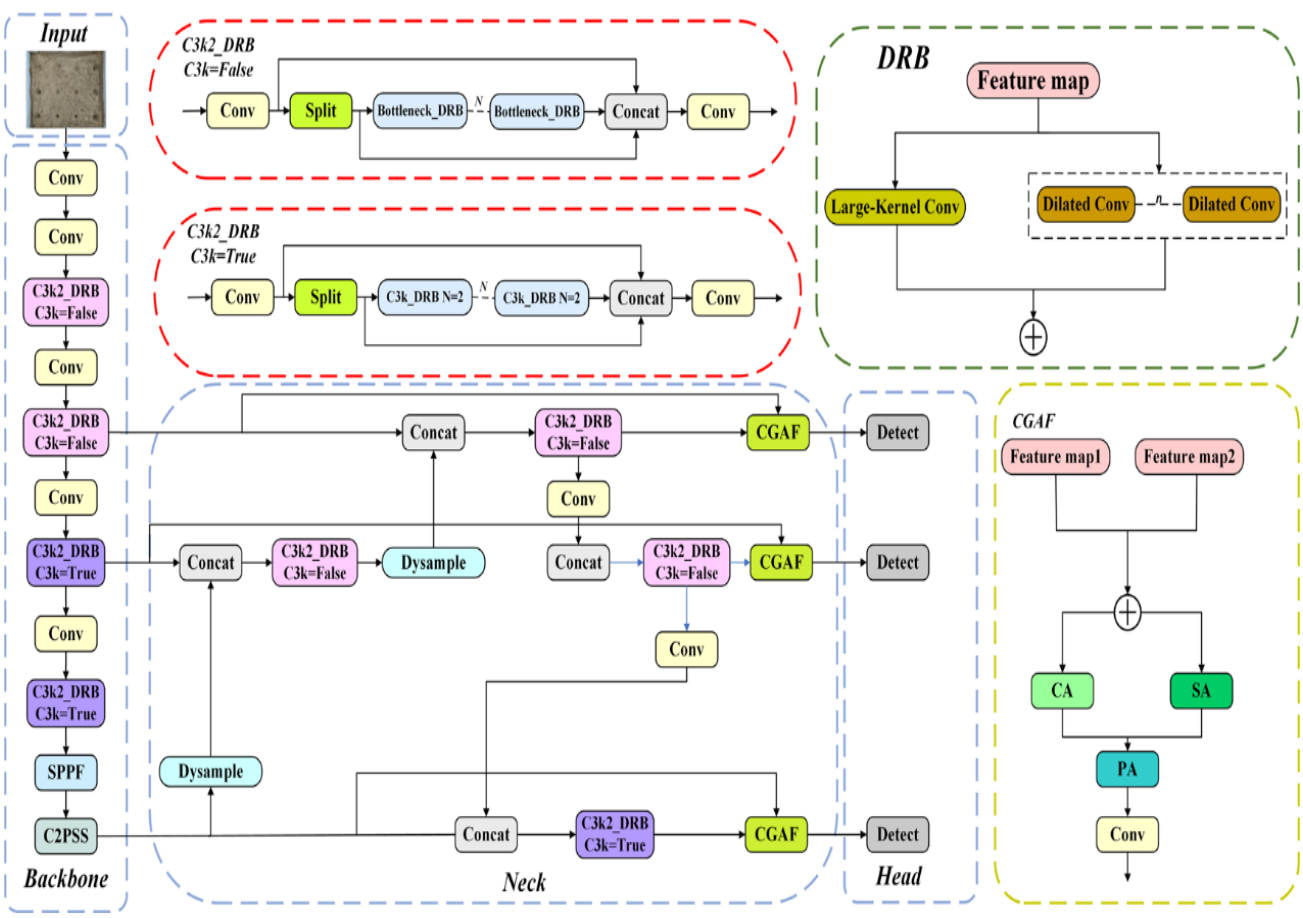

Fig. 5. YOLO11-seed model architecture.

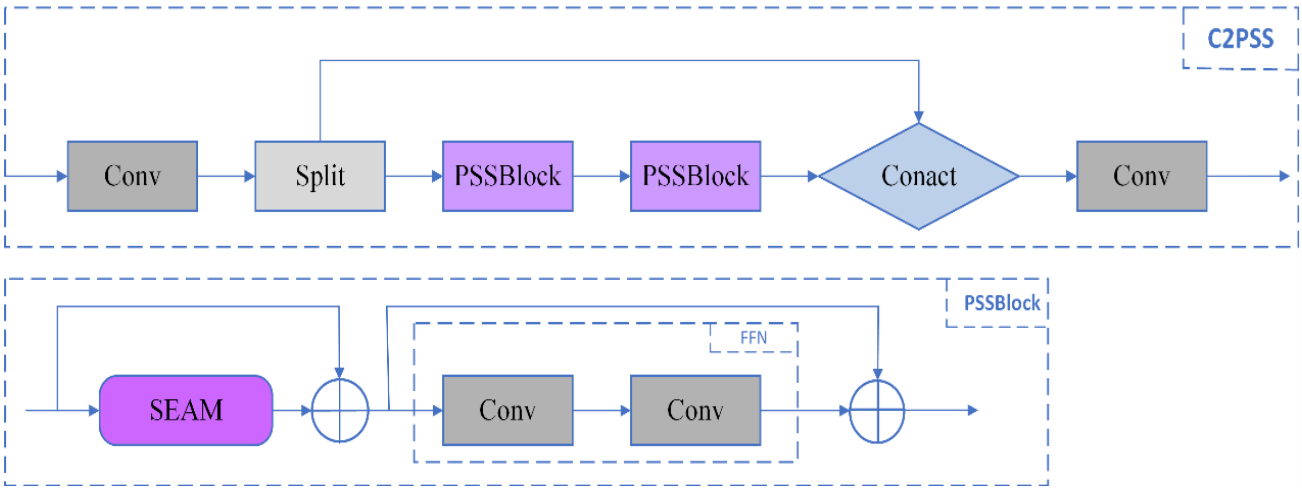

Fig. 6. Network architecture of the C2PSS module.

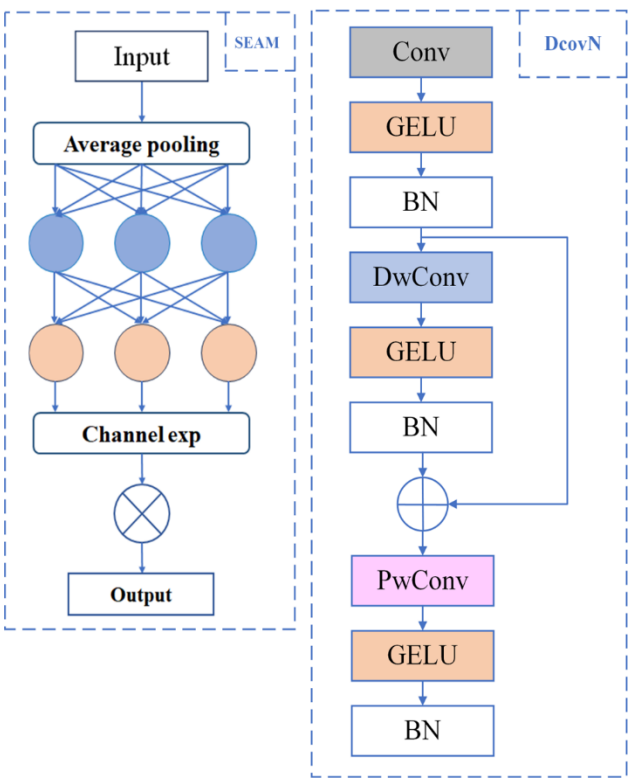

Fig. 7. Network architecture of the SEAM module.

In order to mitigate data imbalance and lower false positives and false negatives, DySample generates neighboring samples, which increases sample variety. It considerably enhances feature map resolution in pelletised seed identification, allowing tiny targets to be distinguished even under occlusion conditions. This makes it easier to find things and lowers the number of false positives and missed detections. Additionally, the point-based sampling technique and learning-based upsampling methods enhance model efficiency and performance with minimal computational cost by omitting dynamic convolution operations and additional sub-networks. Because of this, the module is especially good at finding dense pelletised seeds in places where there aren't many resources or computers.

DRB and CGAF module: The DRB module is introduced to address the limited receptive field issue of YOLOv11 in detecting small objects. This module integrates dilated convolution (Yang & Li, 2022) and reparameterization techniques to improve the model's feature representation capacity (Zhang *et al.*, 2024a). Dilated convolution expands the receptive field by introducing gaps between the elements of the convolution kernel, without increasing the number of parameters (Zhang *et al.*, 2024b).

Figure 10 demonstrates that Figure 10(a) utilizes a dilated convolution with a dilation rate of 1, yielding a 33 receptive field for each element. Figure 10(b) applies a convolution with a dilation rate of 2, resulting in a 7×7 receptive field for each element; Figure 10(c) uses a convolution with a dilation rate of 4, extending the receptive field of each element to 15×15. For a detailed derivation of the DRB formula, please refer to Supplementary1-6 in the appendix.

The reparameterization procedure in DRB mostly changes the complicated multi-fork architecture used during training into a simpler, more efficient single-fork structure for inference. This approach consists of a sequence of mathematical changes and parameter consolidation. For a detailed derivation of the CGAF formula, please refer to Supplementary 7-19 in the appendix.

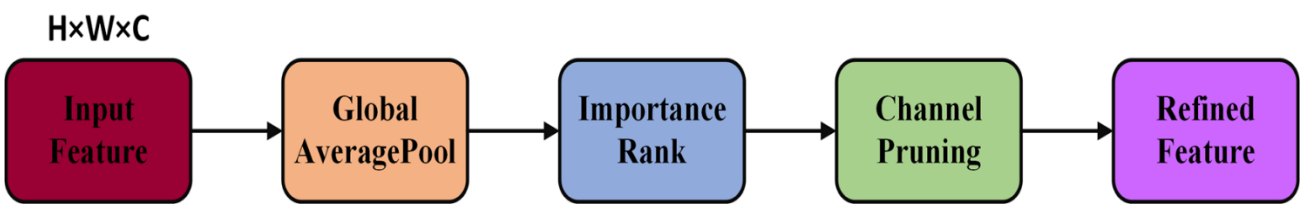

Fig. 8. Network architecture of the SF module.

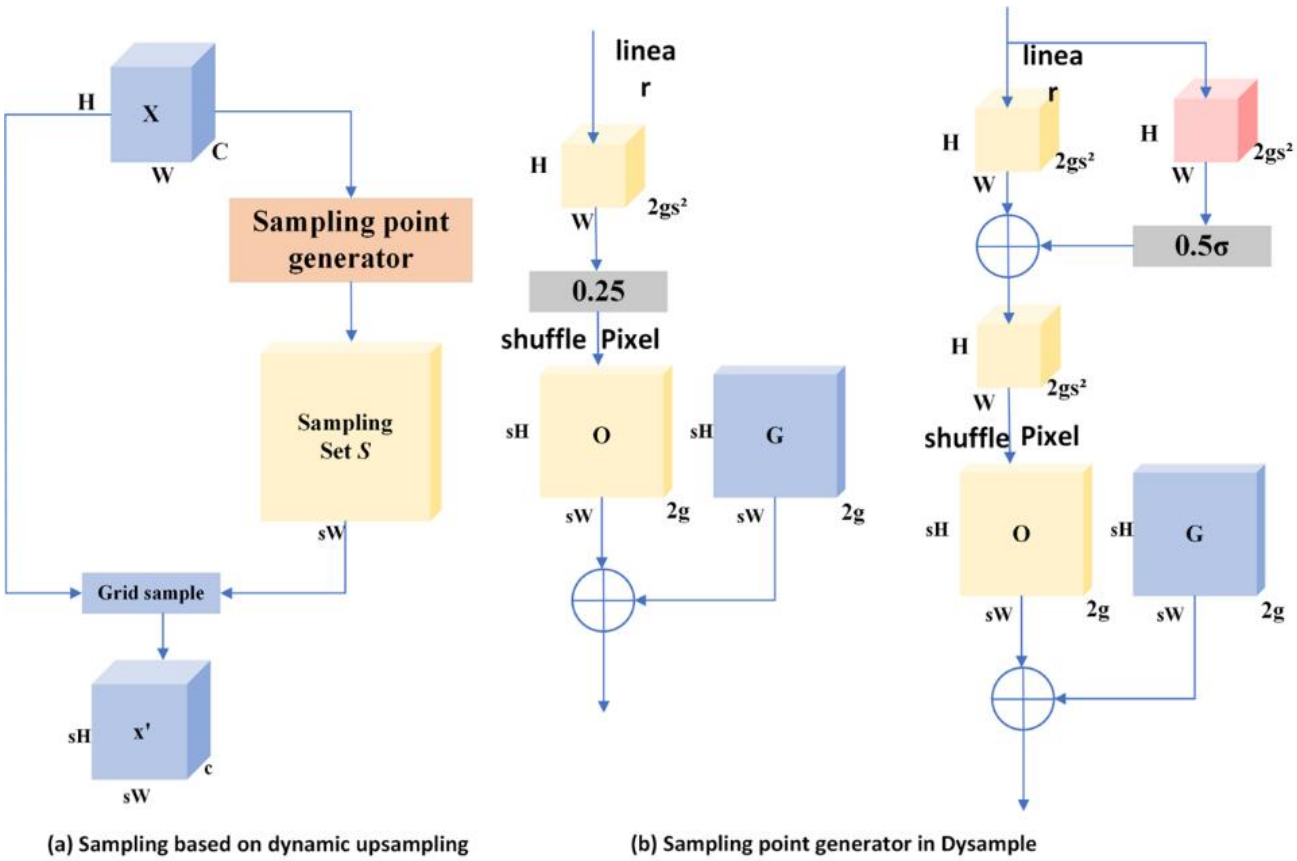

Fig. 9. Network architecture of the DySample module.

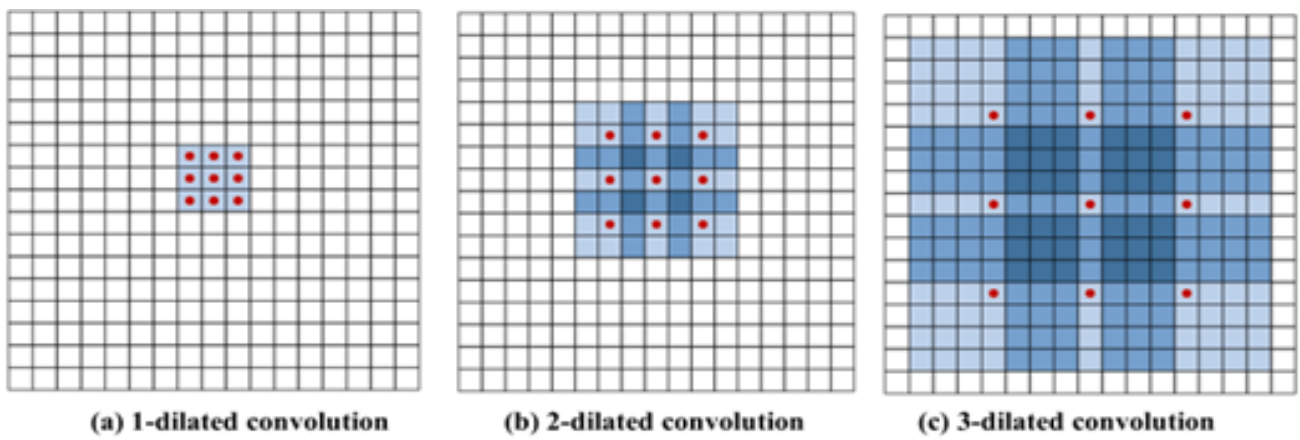

Fig. 10. Receptive field of dilated convolution.

To rectify the inadequate feature extraction proficiency of YOLOv11 after multi-scale feature fusion, we provide the CGAF module. Deriving attention weights from various scales and dimensions, then combining the two input feature maps to create the output feature map, is the fundamental concept of the CGAF module. The CGAF module has three parts: spatial attention (SA), channel attention (CA), and pixel attention (PA). and CGAF Module (Zhou *et al.*, 2024).

Experimental verification: This section introduces the assessment metrics, hardware, and parameters configuration used in the experiment. Meanwhile, the ablation experiment verifies the superiority of the proposed YOLOv11-SEED.

Model evaluation metrics: In the object detection task, common performance metrics such as Intersection over Union (IoU) precision, recall, and mean Average Precision (mAP) are typically used to assess the performance of a model on a given dataset. A series of bounding boxes, each with a matching confidence score, is predicted by the model for the input picture during the trial. During the object identification process, the accuracy of the predictions is assessed by comparing and matching the predicted bounding boxes with the ground truth bounding box. According to formula (24), the IoU metric is determined as the ratio of the overlap area of the predicted box and the ground truth box to their union area.

$$IOU=\frac{DetectionResult\cap GroudTruth}{DetectionResult\cup GroudTruth} \quad (24)$$

The ratio of True Positive (TP) predictions to all expected occurrences is known as precision. In the experiment, an IoU threshold of 0.25 was established, and the predicted bounding boxes were compared to the actual bounding box. If a predicted bounding box passes the IoU criterion and the projected category is right, it is dubbed a True Positive (TP) (Zou *et al.*, 2024). A False Positive (FP) prediction occurs when the model misclassifies an item or detects an entity that is absent from the real data. When the model fails to recognise an item existing in the actual data, it is referred to as a False Negative. Precision is articulated in Equation (25).

$$Precision=\frac{TP}{TP+FP} \tag{25}$$

The trade-off between accurately recognising positive instances and recalling all positive examples in detection is shown by the precision-recall (PR) curve. The curve illustrates precision on the Y-axis and recall on the X-axis. Average Precision (AP) (Chen *et al*., 2024b) quantifies the area beneath the Precision-Recall (PR) curve, encapsulating it as a singular scalar value. Equation (27) delineates the concept of AP.

$$AP=\int_0^1 Precision(Recall)dRecall \tag{27}$$

For tasks with multiple target categories, Average Precision (AP) is computed individually for each category. The mean Average Precision (mAP) is then obtained by averaging the AP values across all categories. The definition of mAP is given in Equation (28), where (N) represents the total number of categories.

$$\text{mAP}=\frac{1}{N}\sum_{i=1}^{N} AP_i \tag{28}$$

Experimental environment: PyCharm served as the development environment for the experiment, which was carried out on the Windows 11 operating system. We changed and improved the Ultralytics YOLOv11s framework to make it easier to find objects in plug tray pictures. This experiment used the following hardware setups: Table 1 shows the hardware setups, and Table 2 shows the program setups.

Table 1. Hardware configuration.

Name	Configuration
CPU	Intel(R) Core (TM) i9-14900HX
GPU	NVIDIA GeForce RTX 4070 Laptop GPU
RAM	64.0 GB

Table 2. Software configuration.

Name	Configuration
Python	3.10.14
Pytorch	2.2.2
CUDA	12.1

Model parameter settings: The hyperparameter used for training are listed in Table 3.

Table 3. Training parameters.

Parameters	Configuration
Number of training epochs	100
Batch size	64
Number of worker threads	10
Initial learning rate	0.01
Momentum	0.937
Input image size	0.0005
Weight decay	1280×1280

Figure 11 illustrates the effect of different training epochs on the performance of model. Experimental results confirmed that 100 epochs are the optimal number of training iterations.

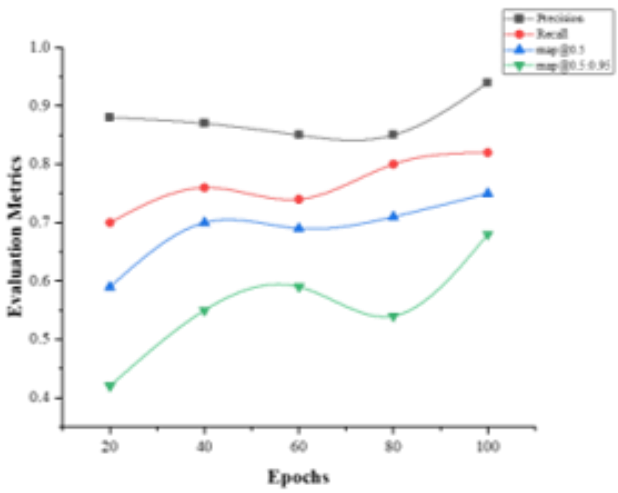

Fig. 11. Evaluation metrics at different training epochs.

Experimental results and analysis: A bespoke dataset was used to test the performance of the YOLOv11s-SEED model. Upon completion of 100 training epochs, the model achieved convergence. Figure 12 demonstrates that the box_loss measures how different the predicted bounding boxes are from the ground truth bounding boxes. Smaller values mean greater detection accuracy. The cls_loss parameter calculates the difference between anticipated and actual class labels, with lower values indicating better classification accuracy. Moreover, the dfl_loss converts continuous coordinate prediction into discrete probability distribution prediction, boosting coordinate localisation accuracy. Lower values signify greater predictive accuracy.

The P-R curve of YOLOv11-SEED shows accuracy and recall during pelletised seed identification training in Figure 13(a). The curve predominantly resides in the upper right quadrant of the figure with minimal variation, signifying the model's strong performance and generalisation capability. The mAP50 in detection tests for two pear fruit kinds was 83.3%, indicating the YOLOv11-Pear model's exceptional performance.

The YOLOv11-SEED F1 score was 82% in Figure 13(b). The model's capacity to identify distinct pear fruit varieties at different confidence levels is evaluated more thoroughly using the F1 score curve. The F1 score goes up quickly as confidence grows, then slowly goes down until confidence hits 0.8.

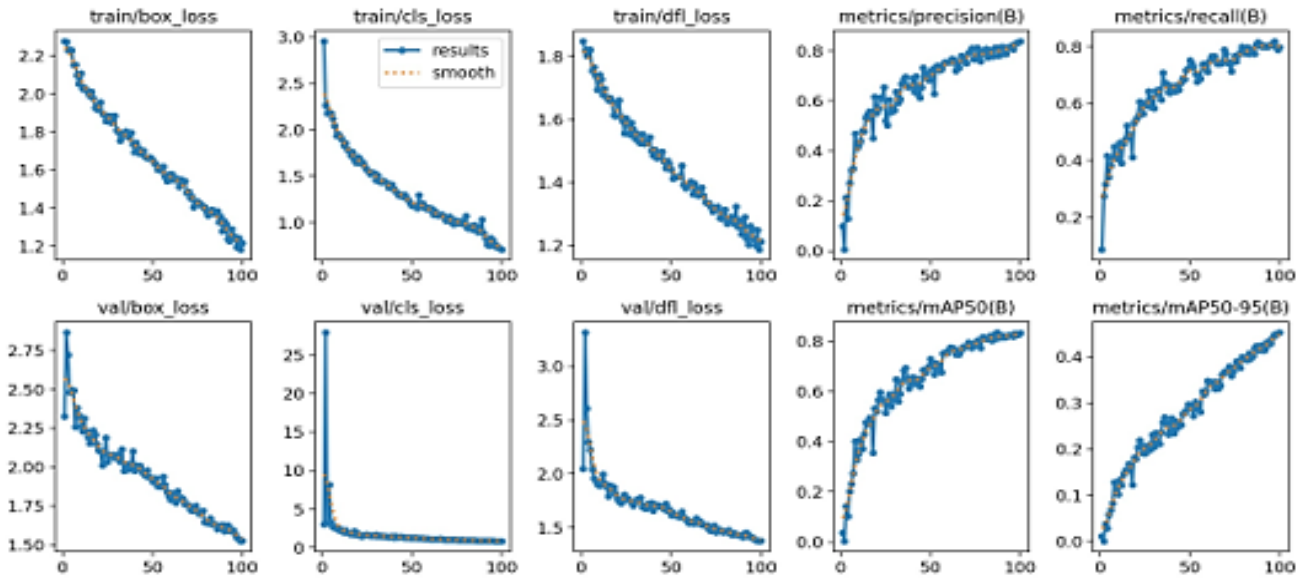

Fig. 12. Progress of training and validation loss and metrics.

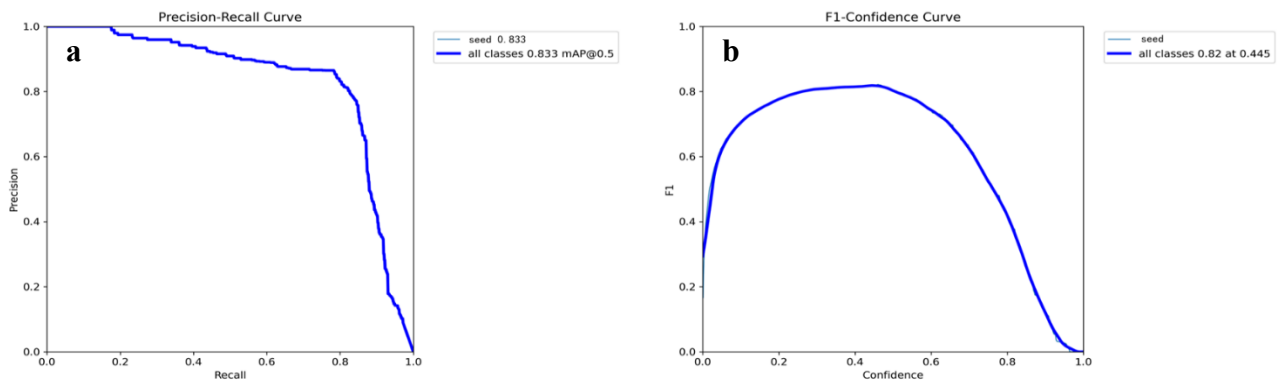

Fig. 13. F1 score curve and P-R curve of YOLO11-SEEDmodel.

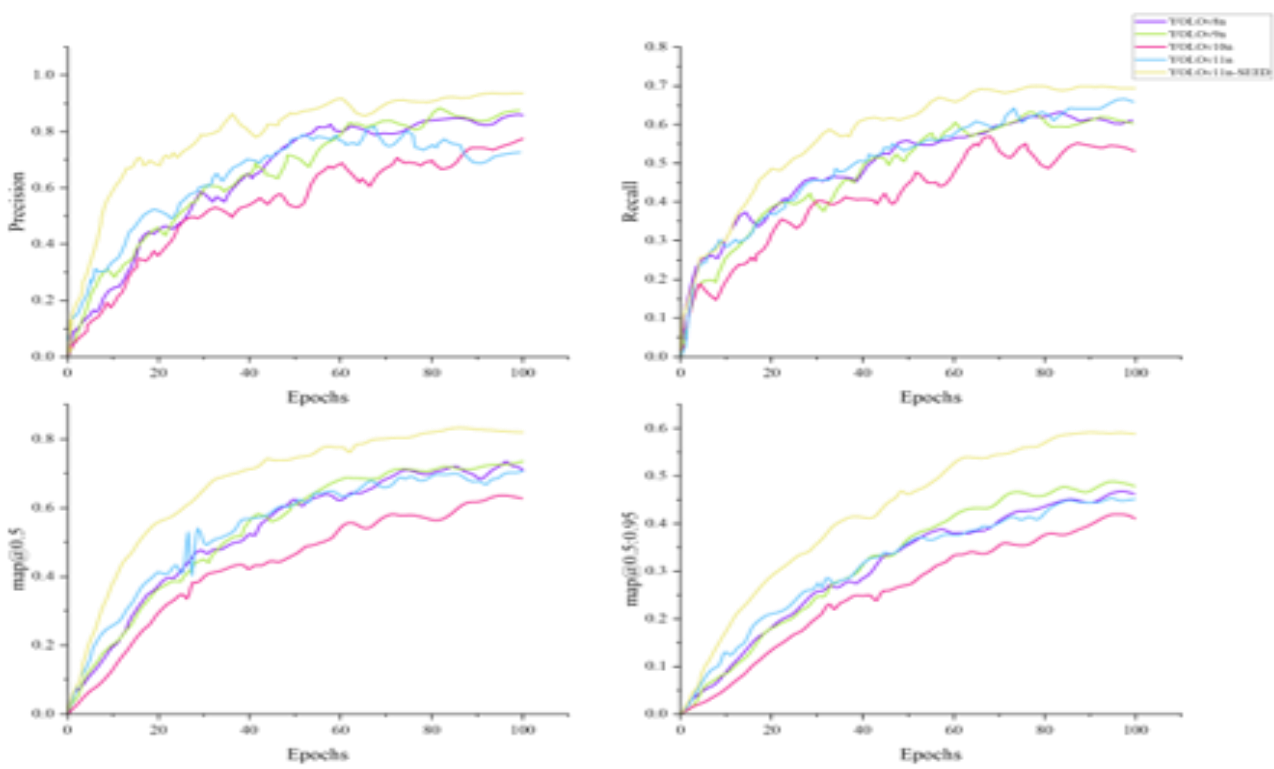

Fig. 14. Metrics curve assessment and comparison between YOLOv11n-SEED and baseline models.

This research also compared the suggested model to other well-known models, such as YOLOv8n, YOLOv9n, YOLOv10n, and YOLOv11n, and found that it was better in four areas.

At each training step, the experimental results, as shown in Figure 14, demonstrate that YOLOv11n-SEED performs better than all other models on all four measures. The performance is notably exceptional. The curves of YOLOv9s and YOLOv8s are similar, and in the latter phases of training, these two models provide equivalent measures following convergence. With the exception of recall, the other three metrics of these two models are only marginally lower than those of YOLOv11n-SEED. In this model, YOLOv11n has the lowest precision but the best recall, indicating an imbalance between accuracy and recall. YOLOv10n performs poorly across all four measures over the bulk of training stages, making it the least effective model in this experiment, with significant differences in recall and mAP@0.5 compared to other models.

Ablation study: Ablation tests were carried out in order to evaluate the performance gains that were obtained by integrating the SF, DRB, and CGAF modules into the YOLOv11s-SEED model. This research looked at how well the YOLOv11n, YOLOv11n-SF, YOLOv11n-DRB, YOLOv11n-CGAF, YOLOv11n-SF+DRB, YOLOv11n-SF+CGAF, YOLOv11n-DRB+CGAF, and YOLOv11n-SEED models trained based on four different measures.

In Figure 15, the individual integration of the SF, DRB, and CGAF modules improves performance across most evaluation measures, notably accuracy, compared to YOLOv11n. The only deviation is that YOLOv11n-CGAF's recall is somewhat worse than YOLOv11n's. YOLOv11n-SF, YOLOv11n-DRB, and YOLOv11n-CGAF demonstrate comparable performance in terms of precision and recall. On both the mAP@0.5 and mAP@0.5:0.95 measures, YOLOv11n-DRB performs quite well.

The integration of the SF, DRB, and CGAF modules results in the extension of YOLOv11n to YOLOv11n-SEED. Yulov11n-SEED's accuracy curve is greater than others during training and peaks at the end. In the advanced phase of training, it exhibits enhanced stability and improved convergence efficacy. In a similar vein, YOLOv11s-SEED achieves optimum convergence later on and shows greater recall than other models during the majority of the training process. When it comes to the mAP@0.5 measure, YOLOv11s-SEED consistently outperforms other models. The dominance of YOLOv11s-SEED is maintained even under the stricter mAP@0.5:0.95 measure, as its curve outperforms all other models, showing greater detection ability. Comparative analyses of YOLOv11s, YOLOv11s-SF+DRB, YOLOv11s-SF + CGAF, YOLOv11s-DRB + CGAF, and YOLOv11s-SEED were carried out in order to ascertain if the SF, DRB, and CGAF modules needed to be included.

Figure 16 shows that incorporating any two of the three modules into YOLOv11 improves all four-assessment metrics relative to the original YOLOv11 model. The YOLOv11-SEED model beats setups employing just two modules across all four measures by combining all three modules. The precision curve of YOLOv11n-SEED consistently surpasses that of other models during the majority of the training process, ultimately reaching the highest value. It exhibits better convergence performance and increased stability in subsequent training phases. The recall of YOLOv11n-SEED exceeds that of other models throughout the majority of the training process and exhibits superior convergence performance in the later stages. YOLOv11s-SEED always does better than other models when it comes to the mAP@0.5 statistic. Under the stringent mAP@0.5:0.95 metric, YOLOv11s-SEED demonstrates superior detection performance, as evidenced by its curve surpassing those of all other models.

YOLOv11n-SF and YOLOv11n-DRB perform better than other Baseline Models in terms of precision and recall, whereas YOLOv11n-CGAF performs somewhat worse than YOLOv9n, as seen in Table 4. YOLOv11n-SF and YOLOv11n-DRB demonstrate superior mAP@0.5 performance compared to other baseline models. At mAP@0.5:0.95, YOLOv11n-DRB and CGAF beat all baseline models, with the former being much better. The combination of SF, DRB, and the CGAF Module within YOLOv11n leads to a performance improvement that exceeds the enhancements achieved by utilising any two of these modules in isolation. SF, DRB, and the CGAF Module allow YOLOv11n-SEED to outperform all models in the experiment across all four Evaluation Metrics.

In contrast to YOLOv11n, YOLOv11n-SEED improved precision by 14.3%, recall by 4.1%, mAP@0.5 by 10.6%, and mAP@0.5:0.95 by 14%. The findings demonstrate the enhancements realised by the YOLOv11s-SEED model.

Figures 17 (a-h) respectively present the original images alongside the number of pelletized seeds detected under the same conditions by YOLOv9n, YOLOv10n, YOLOv11n, YOLOv11n-SF + DRB, YOLOv11n-SF + CGAF, YOLOv11n-DRB + CGAF, and YOLOv11n-SEED. The results indicate that YOLOv11n-SEED achieved the highest recognition accuracy.

Table 4. Comparison of evaluation metrics for YOLOv11n-SEED and other model.

Model	Precision	Recall	map@0.5	map@0.5:0.95
YOLOv8n	0.730	0.668	0.747	0.474
YOLOv9n	0.879	0.607	0.727	0.48
YOLOv10n	0.790	0.546	0.643	0.424
YOLOv11n	0.792	0.658	0.718	0.459
YOLOv11n-SF	0.883	0.686	0.728	0.479
YOLOv11n-DRB	0.887	0.677	0.763	0.562
YOLOv11n-CGAF	0.857	0.608	0.70	0.495
YOLOv11n-SF + DRB	0.91	0.695	0.760	0.582
YOLOv11n-SF + CGAF	0.827	0.695	0.757	0.519
YOLOv11n-DRB + CGAF	0.872	0.682	0.772	0.573
YOLOv11n-SEED	0.935	0.699	0.826	0.599

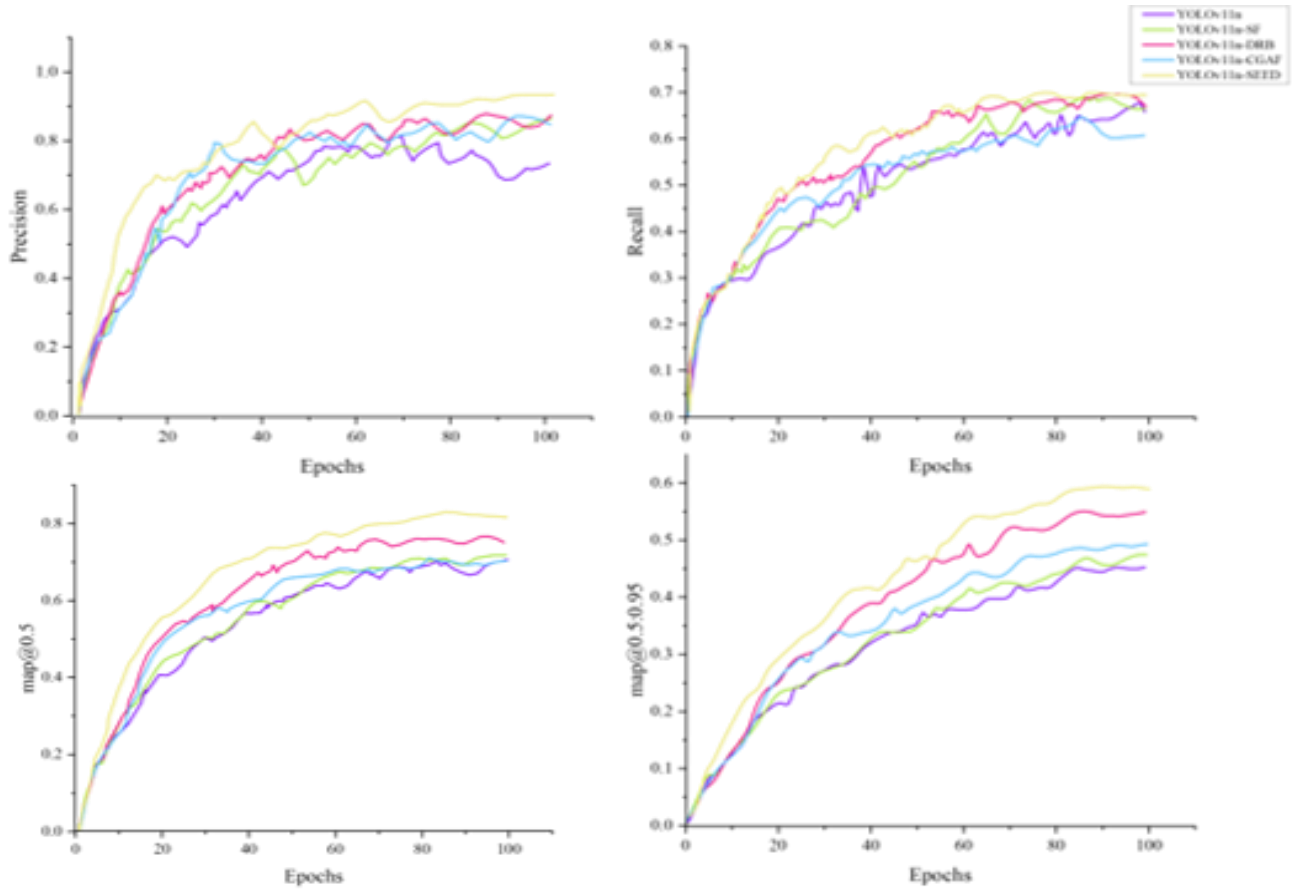

Fig. 15. Comparison curves of evaluation metrics between YOLOv11n-SEED and single model module.

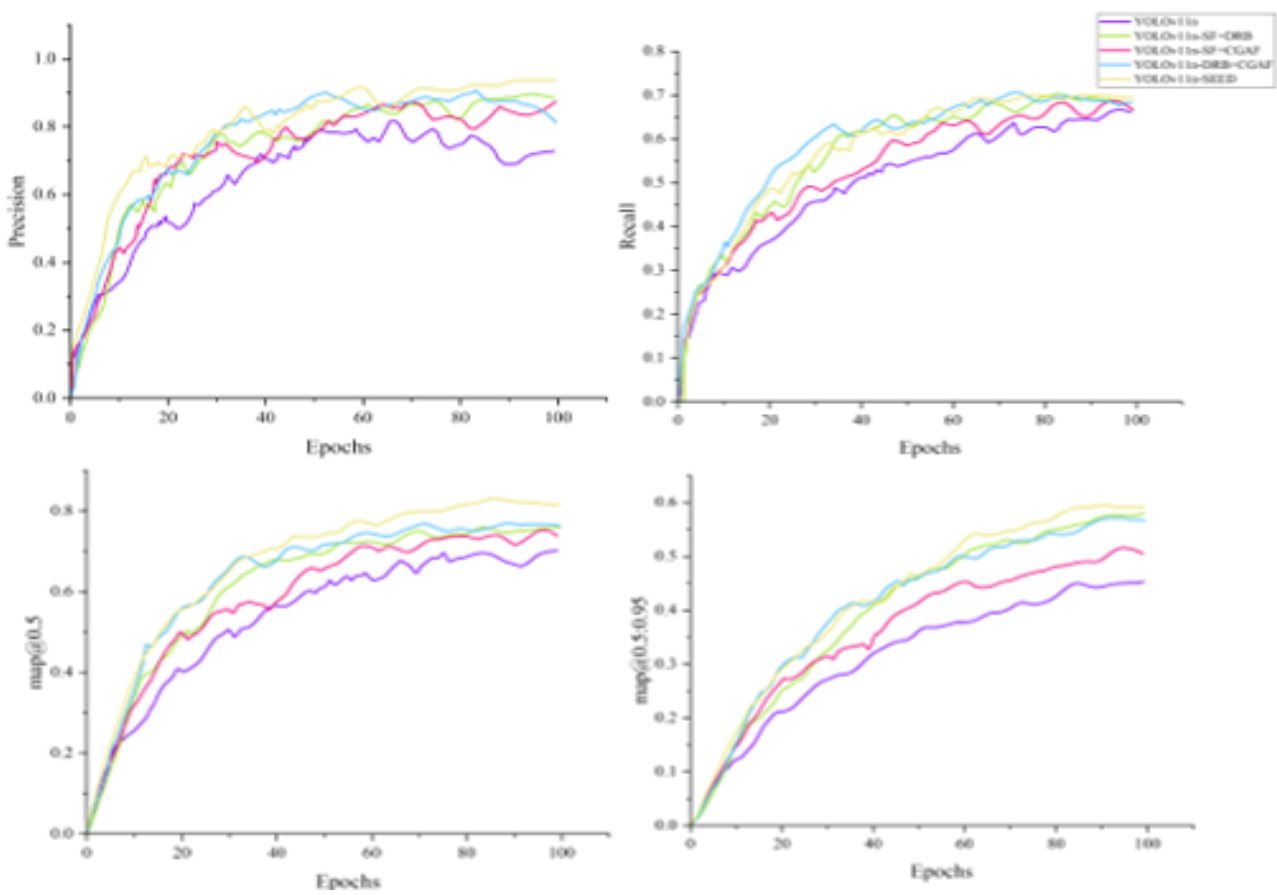

Fig. 16. Assessment and comparison of the metrics curves of YOLOv11n-SEED and the dual-module model.

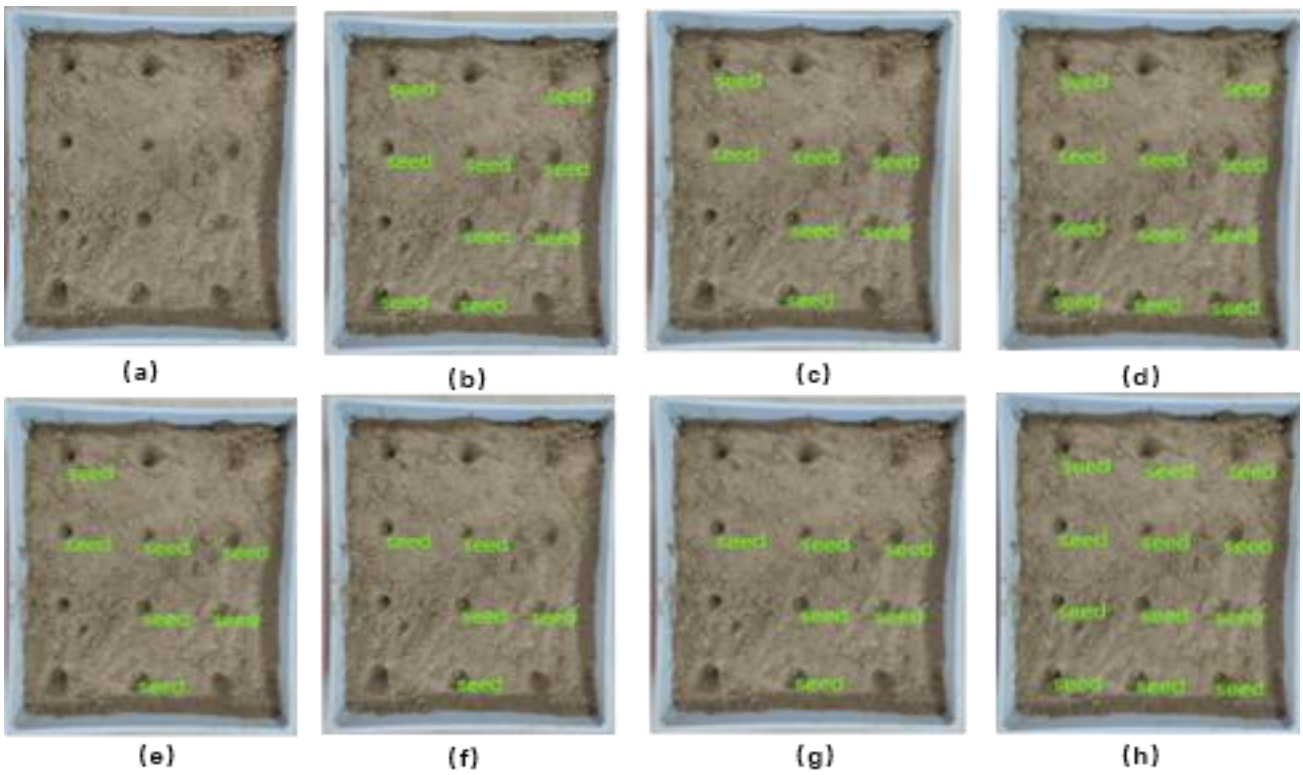

Fig. 17. Comparison of detection results.

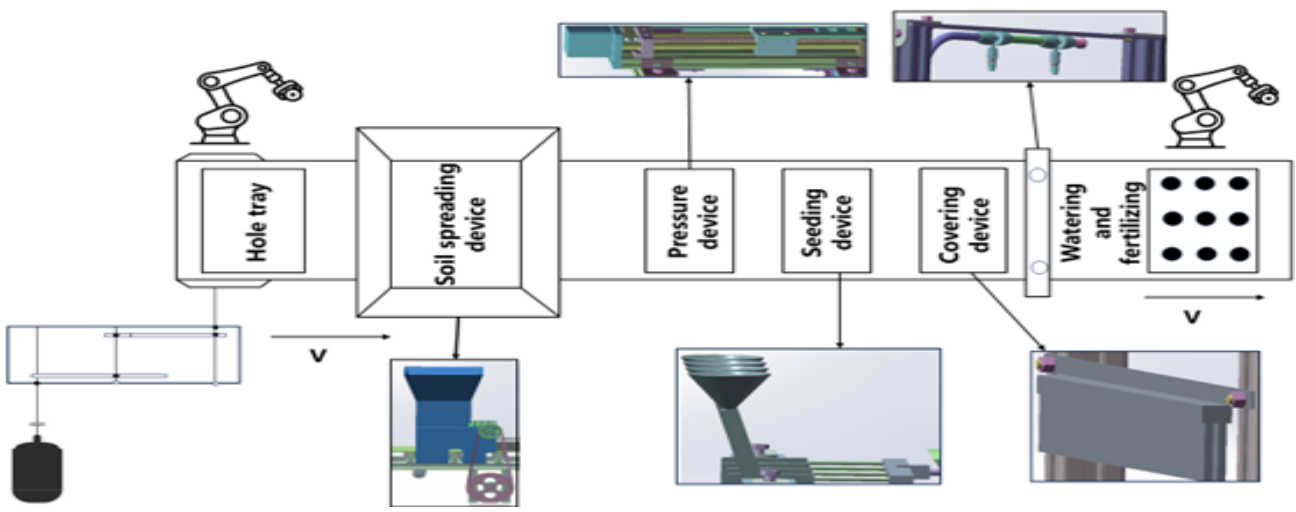

Fig. 18. Layout of complete production line for the plug tray machine.

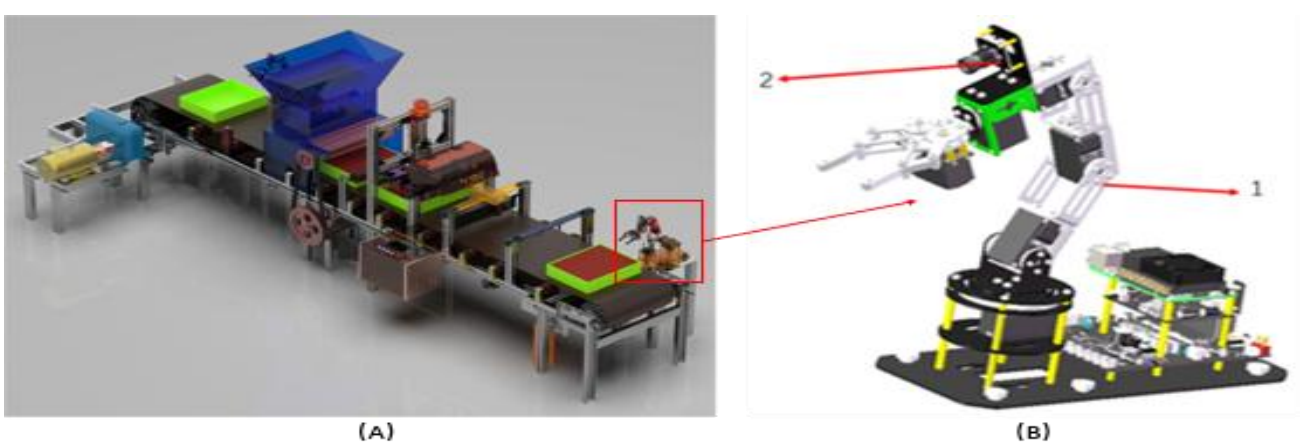

Fig. 19. Mechanical structure of the PTSC production line.

Tray seedling production line

System overall architecture framework: The integrated machine primarily consists of the reducer device, front-end feeding device, object recognition device, soil laying device, conveying device, hole pressing device, seeding device, soil covering device, and fertilizing and watering device, as shown in Figure 18. During operation, the robotic arm first grips the plug tray onto the conveyor belt using the vision algorithm. The motor output is high-speed and low-torque, which is converted into low-speed and high-torque by the two-stage reducer, enabling stable operation of the plug tray on the conveyor belt. The plug tray is recognized and automatically fed by the front-end robotic arm and the camera mounted on it. As it is transported by the conveyor belt and passes through the soil laying device, rapid soil laying is performed by the soil laying device. When passing through the hole pressing device, the plug tray briefly pauses, and the hole pressing device performs the hole pressing operation. After cell pressing is completed, the conveyor belt continues to transport the plug tray to the seeding position for triple seeding. The Soil covering operation is then performed by the brush. When the plug tray passes through the watering device, watering and Fertilization are carried out. Finally, the robotic arm places the plug tray back onto the rack.

Mechanical structure design: The Power System is the core component ensuring stable operation of the entire machine and enabling automation and high efficiency in Vegetable Seeding. This system consists of an Electric Motor, a Coupling, and a Two-stage Reducer, which converts the high-speed, low-torque output of the Electric Motor into low-speed, high-torque output suitable for the machine's operation. This provides sufficient traction for Plug tray conveying, ensuring smooth and efficient execution of the full process from loading, soil spreading, hole pressing, seeding to soil covering, watering, and fertilizing.

The entire machine is applicable to Planting Base, Seedling Enterprise, and Modern Agricultural Park, significantly improving Seedling Cultivation efficiency and reducing labor costs through Automated Production Line operation. Key components include: the Conveyor Belt Device for stable Plug tray conveying; the Soil Spreading Device utilizing a Belt Drive for uniform soil distribution; the Hole Pressing Device employing a Ball Screw and Hole Pressing Mold to form standard holes; the Soil Covering Device using a Brush to complete soil covering; the Watering and Fertilizing Device adopting an Atomizing Nozzle for precise water and fertilizer application; and the Automatic Loading Device integrating a Robotic Arm and a Vision Recognition System to automatically grasp and place Plug tray. The overall system features a rational structure and user-friendly operation, providing reliable support for the standardization, large-scale implementation, and automation of PTSC. The mechanical structure is shown in Figure 19.

The full production line of PTSC Machine is primarily used in multiple fields such as Planting Base, Seedling Cultivation Enterprise, and Modern Agricultural Park. Its Automated Production Line production mode enables efficient and large-scale cultivation of various vegetables and Flower Seedlings, meeting diversified demands while reducing labor costs and improving seedling efficiency. It provides strong support for the development of Agricultural Standardization, New Variety Breeding, research on Cultivation Technology, and enhancing the overall competitiveness of agricultural parks. The main Technical Parameters of the full production line Plug Tray Vegetable Seeding and Seedling Raising Integrated Machine are shown in Table 5.

Table 5. PTSC production line parameters.

Projects	Technical parameters
Overall dimension	3089mm×1110mm×1430mm
Motor rated power	1.5kW
Working width	380mm
Number of seeding rows	3
Number of seeding columns	4
Seeding depth	30mm
Row spacing	77mm
Degrees of freedom	5(6)
Body weight (kg)	0.6
Working radius (mm)	330.33
Working radius (mm)	250.45
Joint 1 – Range of motion (°)	-180 to +180
Joint 2 – Range of motion (°)	-90 to +90
Joint 3 – Range of motion (°)	-90 to +90
Joint 4 – Range of motion (°)	-90 to +90
Joint 5 – Range of motion (°)	-180 to +180
Joint 1 – Maximum speed (°/s)	60
Joint 2 –Maximum speed (°/s)	60
Joint 3 –Maximum speed (°/s)	60
Joint 4 –Maximum speed (°/s)	60
Joint 5 – Maximum speed (°/s)	60

Electrical control design: The PTSC production line is shown in Figure 20. The overall electrical system of the PTSC production line comprises an Advantech industrial control computer, a Siemens (1214C/DC/DC/DC) PLC, Siemens PLC expansion modules, two Siemens frequency converters, a control cabinet, and Hikvision cameras.

The Host Computer is developed using TIA Portal Software on the Advantech Industrial Control Computer, establishing Communication between the Host Computer and the Siemens PLC. Digital and analog signals are input from the Host Computer to the Siemens PLC, enabling the Siemens PLC to execute the corresponding procedures, thereby controlling multiple processes such as Soil Spreading and Hole Pressing. The Temperature Sensor and Humidity Sensor are acquired via the Analog Expansion Module of the Siemens PLC, displayed on the Host Computer interface after processing by the Siemens PLC, and the data is recorded for storage. The Host Computer also establishes Communication with the Hikvision Camera, enabling Vision Recognition for Tray Seeding. Simultaneously, Communication is established between the Host Computer and two Siemens Frequency Inverter units to control the inverters, thereby adjusting the Rotational Speed of the Conveyor Belt and Soil Spreading. Additionally, the Host Computer communicates with the Stepper Driver to control the operation of the Stepper Motor for completing Hole Pressing and seeding. The system layout is shown in Figure 21.

Figure 21 illustrates the Electrical Control Layout of a Six-tray Seedling Production Line, with the core component being the Electrical Control Box. The box contains a Frequency Inverter (for motor control), Contactor, Fuse, 220V source, Main Circuit Breaker, Relay, switches, and a switch for communication purposes. The system is equipped with multiple Function Switches and Temperature Sensors.

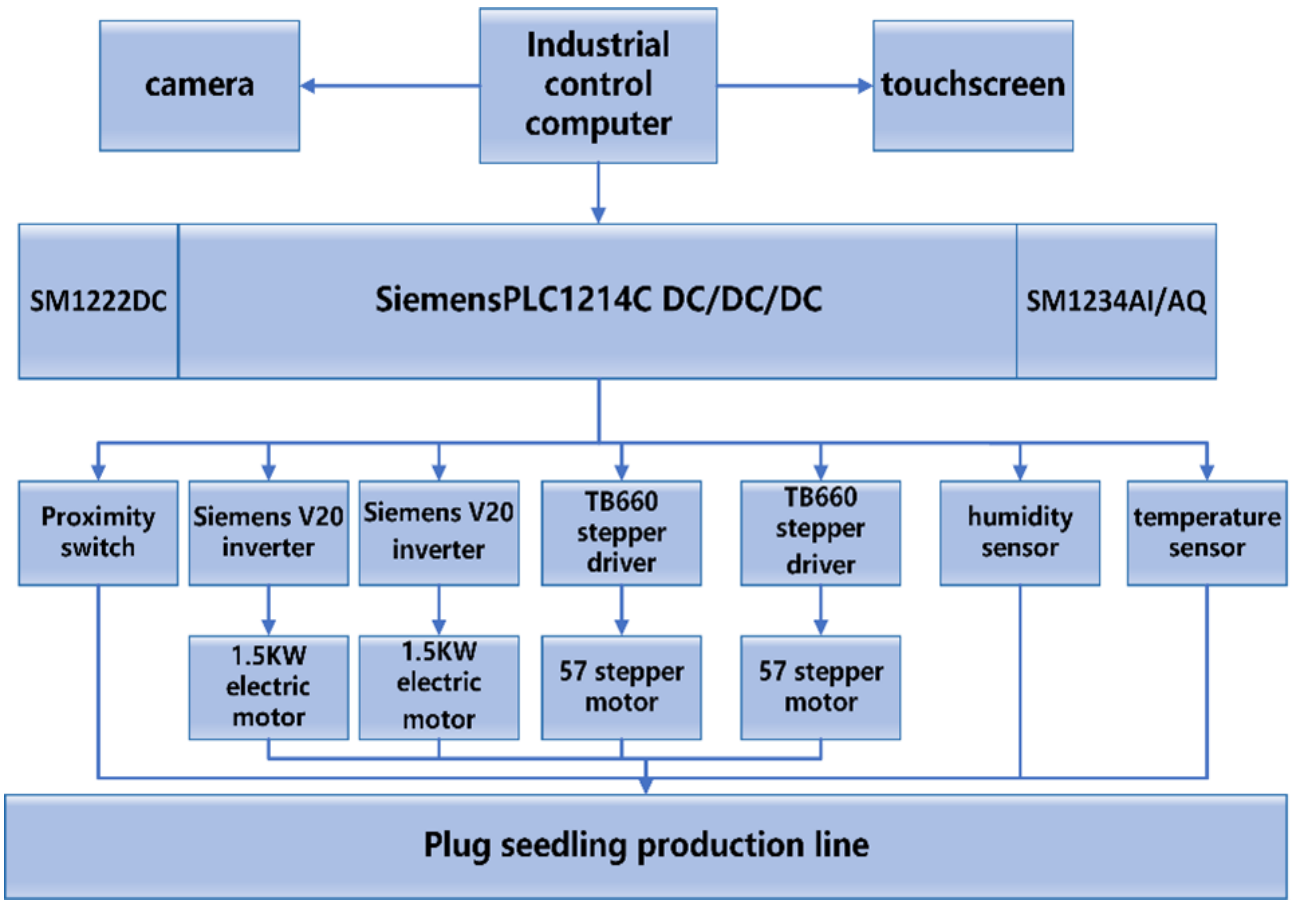

Fig. 20. Overall electrical architecture diagram of the production line.

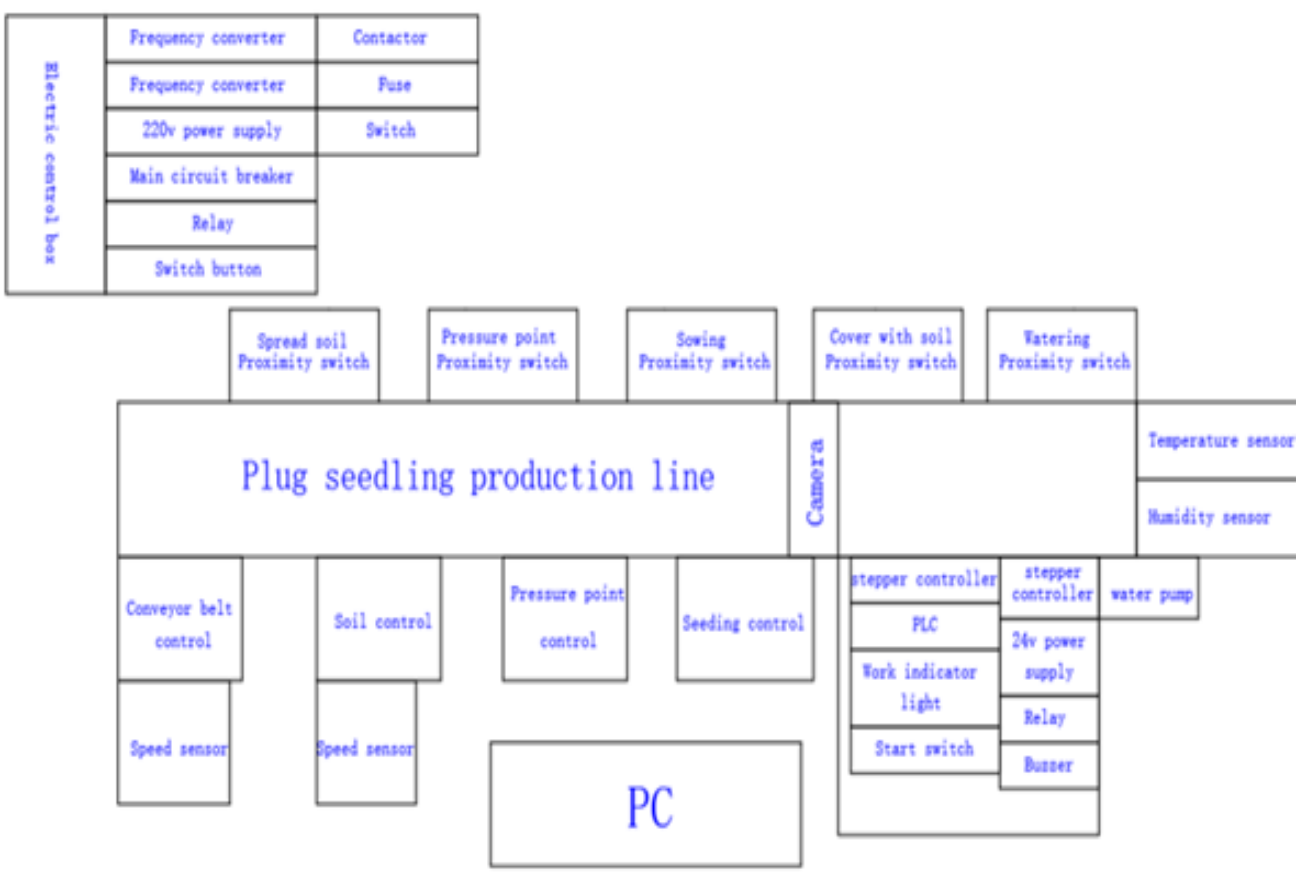

Fig. 21. Overall layout of the production line.

The production line operations are executed by Three-phase Asynchronous Motors (driving the Conveyor Belt and Soil Insertion, and Stepper Motors (controlling pressure and seeding. These are coordinated by the Stepper Controller within the Integrated PLC, which also includes a Work Indicator Light, Start Switch, 24V source, Relay, and Buzzer.

The Camera is used to detect the Precision of Seeding. In cases of missing Seeding or imprecise Seeding, a signal is sent to the Host Computer, triggering an alarm and stopping the production line for manual re-Seeding. External connections include the Water Pump and the Industrial Control Computer (monitored by the Host Computer). Together, they constitute a complete Seedling Automation Control System.

Human-computer interaction interface design

Overall structure of the human-machine interface: The human-machine interface of the plug seedling production line is developed using WinCC in TIA Portal. The test equipment consists of an Advantech industrial control computer and a control cabinet. In the design, commands are issued by the host computer software, and signals are transmitted to the Siemens PLC module via Communication. The Siemens PLC processor outputs control signals to actuate the corresponding components.

The control software adopts a multi-layer architecture design, which is mainly responsible for human-machine interface interaction, data acquisition, and execution of control instruction. It integrates the product, industrial control computer, operation management system, and Siemens PLC into a unified system. Based on the analysis of process requirements and current test conditions, the operator adjusts the operating equipment of the test bench, thereby modifying and changing the state of the test equipment. Meanwhile, the computer data processing module organizes and standardizes the acquired test process data to generate the test data file. The operating system intended for use in this computer measurement and control system is Windows 11 Professional developed by Microsoft Corporation, USA. The overall structure of the application software for this computer measurement and control system is shown in Figure 22.

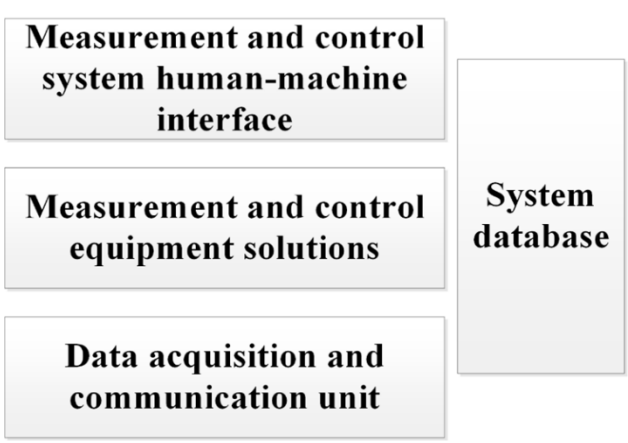

Fig. 22. Overall framework of the human-machine interface.

Design of human-computer interface

1st Layer: User interactive operation interface layer.

Adopts the standard Windows GUI flavor to provide users with a unified and efficient operation interface. It can show the Status Parameter Values of the Test Bench and Device Under Test online, enabling operators to understand the operational status of products and devices on the fly. It features comprehensive operations such as Test Data Acquisition, Data Processing, Data Storage, and Stored Data Playback, with a fully Chinese display interface.

2nd Layer: Test computing terminal

Responsible for the customization of the entire Test and Control Process, as well as System Management, Data Processing, Test Process Control, and Network Communication functions. It mainly completes functions such as Sampling Channel Configuration Management, Calibration and Verification, Signal Conditioning Adjustment, Engineering Unit Conversion, Data Transmission, Data Storage, Format Conversion, and Network Transmission. This tier is the core part of the Software Platform.

Tier 3: Data acquisition and communication unit

These Software Module and Component Unit are built upon the Windows System Kernel and Data Acquisition Equipment, forming the acquisition of various system AO, DI, and Measurement Parameters, as well as the DO Control Output and Communication Link.

Software functional architecture: The Control and Monitoring Software provides a fully Chinese Human-Machine Interface System with the following basic features.

(1) All operations during the test process are completed through Human-Machine Interaction and Menu Selection.
(2) The system is equipped with automatic or manual Storage Function for Test Process Parameters, and includes a Playback Function.
(3) The system supports Digital Display, Analog Knob, Analog Signal Light, and Analog Button for Test Process Parameters.
(4) Users with appropriate rights can access the Field Channel Calibration and Data Acquisition Channel to switch the human-machine interface for convenient tweaking.
(5) The system features powerful and comprehensive Data Management and processing capabilities to meet various data wrangling needs for testing.
(6) The software design adopts object-oriented Modular Technology programming, making it easy to expand and adapt. Based on the above basic characteristics, this Control and Monitoring Software consists of five parts: Calibration and Verification, Self-Testing, Data Processing, Parameter Setting, and System Management.

The functional settings of each part are as follows:

(1) Calibration and Verification: conducts the calibration and verification of the Test Equipment and products.
(2) Self-Testing: conducts self-testing of each Transmission Channel in the device.
(3) Data Processing: enables reading of data in the product, performs applicable parsing, and supports playback.
(4) Parameter Setting: allows tweaking of all variadic parameters in the system.
(5) System Management: responsible for managing all user sign-in info, Access Permission, etc.

Introduction to interface functions: After the system is unparked, users are required to perform Account Password Authentication via the Login Interface. Upon successful login, users can access the Function Selection Main Interface to enter core modules such as Parameter Monitoring, Status Monitoring, and Manual Control. The Parameter Monitoring interface supports the settings and Real-time Monitoring of key parameters such as Inverter Frequency and Seeding Depth, providing Color Alert and hints for abnormal values. The Status Monitoring interface visually showcases the Production Line Operating Status and Equipment Position through Indicator Light, and allows handoff to the Manual Control interface for Process Testing. The system is equipped with a Security Authentication Mechanism and Operation Log Recording functions to ensure operational safety and Traceability, as shown in Fig. 23.

The Visual Feedback Interface shown in Figure 23 is a key component enabling the coordination between Visual Detection and Tray Seeding. The system is preset with three Complete Seeding Rate Threshold values (60%, 90%, and 100%) and automatically determines equipment status and corresponding actions based on the actual number of seeds detected:

(1) When the number of detected seeds is ≤ 60% of the total number of seeds, the Equipment Fault Indicator Light is activated, indicating a seeding omission caused by equipment failure. Shut down for Maintenance immediately;
(2) When the number of detected seeds is between 60% and 90%, the Missing Seeding Indicator Light is activated, indicating a seeding omission due to a Process Issue. Perform Manual Intervention and initiate Re-detection;
(3) When the number of detected seeds reaches 100%, the Seeding Completion Indicator Light is activated, indicating successful completion of the seeding process. The design effectively distinguishes between Equipment Fault and Process Deviation}]]], ensuring the system can Real-time Response to different Seeding Status, thereby improving Reliability and Operational Efficiency.

Prototype test: Figure 24 illustrates the main components of the mechanical testing section, divided into five parts: experimental prototype, hole pressing test, seeding test, vision hardware, and vision software. Among them, vision hardware includes camera, lens, and light source; the vision software section involves the secondary development of Hikvision VM (machine vision platform). The overall structure reflects the integration of vision technology in mechanical testing for functional verification and system development.

Experimental prototype design: The experimental prototype was designed to implement the main functions of the Complete production line plug tray vegetable seeding machine, while ensuring stable equipment operation, ease of use, and convenient maintenance. Due to cost constraints, the current prototype uses a lead screw instead of a conveyor belt, and is primarily used for experimental seeding of a single Tray Seeding unit. carrot seeds with pelleting treatment were used in the experiment.

Acupoint pressure function test: The experimental prototype is driven by a 57-stepper motor and utilizes a ball screw to control the hole pressing die, enabling precise adjustment of the hole pressing depth. The hole pressing die consists of three rows and four columns, pressing 12 holes simultaneously. soil softness and humidity affect hole formation, so the hole pressing depth must be adjusted accordingly. Experimental results indicate that the prototype performs well in terms of depth adjustment, uniformity, and stability. The hole pressing depth is accurate and reliable, meeting the carrot seeds pre-burial depth requirements. When seeding other plants, adjust the hole pressing depth, soil covering thickness, and water and fertilizer amount based on the seed growth characteristics.

Seeding accuracy test: Pelleting treatment carrot seeds were purchased from the market, and a large number of vegetable seeds of various specifications were selected from multiple packages. Measurements showed that most of these seeds had a diameter of approximately 4.5 mm, with the majority being elliptical in shape. Multiple sowing tests were conducted under identical conditions.

By statistically analyzing the number of seeds per cell and the Missing Seeding rate, the Seeding Accuracy was evaluated. The test results indicated that the integrated machine demonstrated high Seeding Accuracy, with 95% of the cell achieving one seed per cell. Occasionally, multiple seeds would fall into the same cell or no seed would be sown, but the Missing Seeding rate remained below 1%, meeting the design requirements.

Visual hardware test: HikvisionMV-EM200M is a high-performance 20-megapixel Industrial Area Scan Camera. It is equipped with a high-performance Global Shutter CMOS Sensor, capable of delivering high-resolution image output at 5472×3648, ensuring rich detail and high clarity in captured images. The camera supports a maximum Frame Rate of 22.6 fps, making it well-suited for high-Precision Vision Inspection in static or low-to-medium speed motion scenarios. It features excellent Low Light Sensitivity and supports multiple functions such as Auto Exposure and Auto Gain, facilitating easy integration into various complex Industrial Vision Systems. It is widely used in applications requiring high image quality, such as Positioning, Measurement, Recognition, and Defect Detection.MVL-KF1628M-12MP is a high-quality Fixed Focal Length Lens specifically designed for high-resolution industrial camera applications. It features a focal length of 16 mm, an Image Circle Size of 17.6 mm, and a Field of View of D (17.52 mm): 54.8°, H (14.14 mm): 44.9°, V (10.35 mm): 33.9°. The lens is available in multiple commonly used focal lengths including 16 mm, 20 mm, 25 mm, and 28 mm, allowing users to select flexibly based on working distance and field of view requirements.

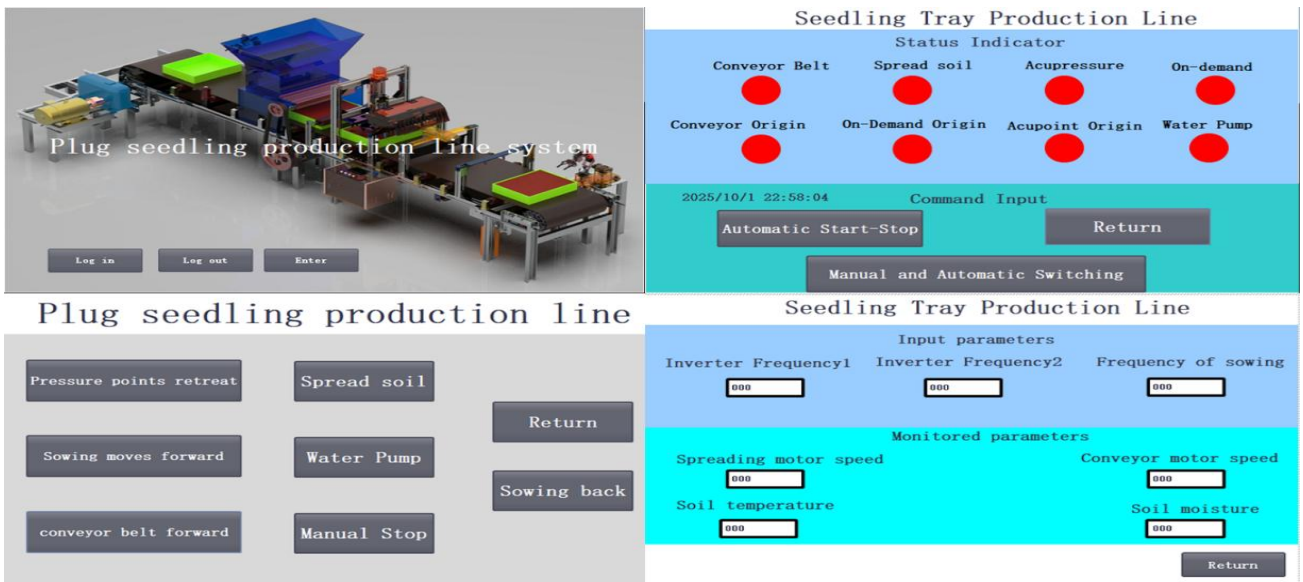

Fig. 23. Interface display.

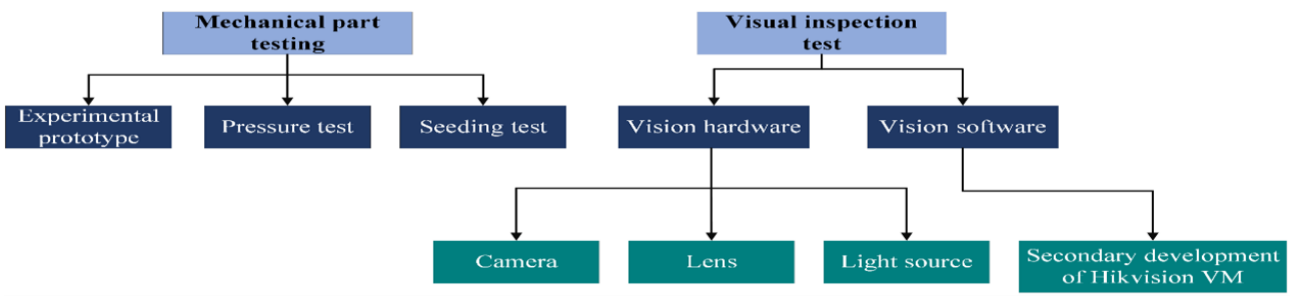

Fig. 24. Main test contents.

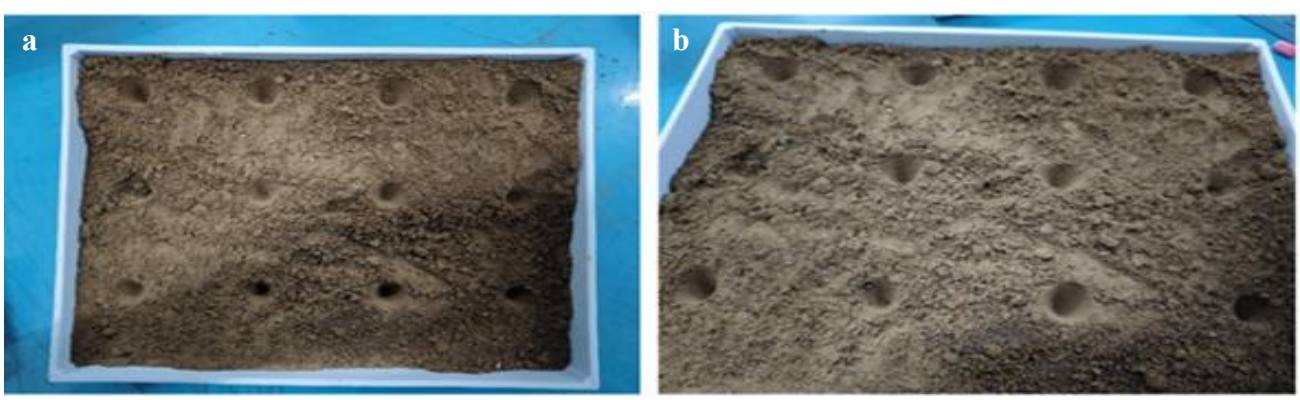

Fig. 25. Acupoint pressure test.

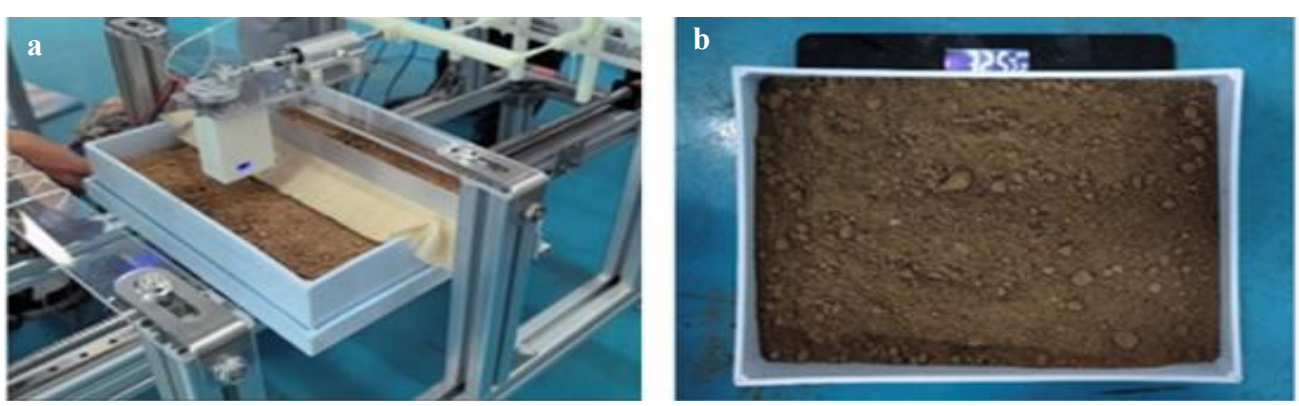

Fig. 26. Completion of seedling and covering soil.

Table 6. Sowing experiment structure.

Number of experiments	Number of acupoints	Theoretical seeding quantity	Actual number of seeds sown	Missed broadcast rate
1	3	36	35	0.027
2	5	60	58	0.03
3	7	84	80	0.047
4	2	24	24	0
5	9	108	106	0.01

Table 7. Seeder operating parameters.

	Soiling	Acupressure	Sowing			Covering soil	Fertilizing
			First sowing	Second sowing	Third sowing		
Distance (mm)	450	800	77.35	77.35	77.35	206	416
Transportation time (s)	2.25	4	0.38675	0.38675	0.38675	1.03	2.08
Working Hours (s)	1.23	8	2	2	2	1.23	1.23

It is equipped with a large Aperture of F2.8, ensuring bright and low-noise image capture even under low-light conditions. The Aperture can be manually adjusted to optimize the Depth of Field effect. The lens adopts a standard C-Mount (C-Mount), offering broad compatibility, making it an ideal choice for high-quality Machine Vision System applications in Detection, Measurement, and Recognition.MV-LRSS-H-80-W is a high-Brightness, compact white Ring Light Source. The "80" in its name indicates an outer diameter of 80 mm, while "H" typically denotes a "high" version, meaning the light-emitting body has a certain height (or thickness), enabling a more concentrated beam angle and enhanced penetration, which is particularly suitable for highlighting surface textures and contours of objects. This light source adopts a high-density LED Array to ensure uniform illumination with adjustable Brightness, and is professionally designed to effectively reduce Diffuse Reflection, making it well-suited for addressing illumination challenges on Specular Reflection surfaces. It is typically Coaxial Mounting with a lens, and is widely used in PCB Inspection, Precision Component Positioning, Character Recognition, and other machine vision inspection scenarios requiring shadow elimination and contour enhancement.

Software test: Hikvision Vision Master 4.2.0 is a piece of Machine Vision Algorithm Platform software that modularizes Vision Detection Tasks and provides a variety of Algorithm Tools, including Positioning, Measurement, Recognition, and detection. The platform offers a Graphical Development Environment based on Flowchart Mode, enabling users to build, debug, and deploy Vision Applications by dragging and connecting modules, which helps reduce the need for code writing. VM 4.2.0 has been updated in terms of Algorithm Performance, Compatibility with Hikvision Hardware Devices, User Interface, and support for Secondary Development, and provides API Interface support for invocation in languages such as C# and C++.

This project conducts Secondary Development to deploy the trained YOLOV11n-SEED Model on the VM4.2 Platform, and achieves Vision Detection Tasks through high-level integration with the Hikvision Camera System. Figure 30 shows the software interface. Figure 30(A) displays the Programming Interface and the Camera Display Interface of the software itself.

Visual feedback test: The precise implementation of Visual Detection in the plug seedling production line relies on the feedback function of the detection signal. A, when the detected number of Seeds is 12, the number of Seeds displayed on the Host Computer is also 12, which is consistent with the detection result. The Seeding Completion Indicator Light is illuminated, meeting the design requirements. when the detection result is 9, the number of Seeds displayed on the Host Computer is 9, which is consistent with the result of object detection. The Missing Seeding Indicator Light is illuminated, and the machine stops moving for Manual Intervention, meeting the design requirements.

Overall process testing: As summarized in Table 6, the integrated system was evaluated across varying operational scales (24 to 108 seeds per trial), demonstrating high precision with an average missed broadcast rate of approximately 2.28% and achieving a perfect success rate in specific configurations. The minor discrepancies observed between theoretical and actual seeding quantities are primarily attributed to inherent seed morphological variations and mechanical vibrations during conveyance, rather than deficiencies in the detection algorithm. These empirical results provide robust evidence of the system's operational stability and its readiness for large-scale vegetable seedling production.

The conveyor belt speed is set to 0.2 m/s. The seed tray dimensions are 245 mm × 310 mm × 65 mm, with a sowing spacing of 79 mm × 77.35 mm. A robotic arm places the seed tray onto the conveyor belt, which then transports it to the pressing position. The transport distance to the pressing station is 1250 mm, requiring 6.25 s. Upon arrival, the conveyor stops for an 8 s pressing operation. During pressing, the robotic arm simultaneously loads the next tray. After pressing, the conveyor resumes operation, moving the tray to the sowing position. The system performs three sequential sowing cycles, each taking approximately 2 s, with short intervals between cycles. After sowing, the tray is conveyed to the unloading position. The process is cyclic: while one tray is being unloaded, the next reaches the pressing station, enabling continuous operation. To avoid interference between sowing and soil spreading, the soil spreading device is positioned 450 mm from the loading point. The total processing time for the first tray is 26.5 s, while subsequent trays require only 3.84 s due to system overlap. Each tray covers an area of 0.07595 m², resulting in approximately 9.37 h to complete one acre. Key process timings are summarized in Table 7.

Conclusion

Project summary: This study successfully developed an integrated Plug Tray Seedling Cultivation (PTSC) machine coupled with an advanced YOLOv11-SEED vision detection system, specifically designed for large-scale vegetable and flower production. By integrating material feeding, soil spreading, precision seeding, and automated watering into a single production line, the system effectively addresses the long-standing bottlenecks of high labor intensity and unstable seeding quality in traditional agricultural practices.

Key technical advancements and findings include: System Integration: Through repeated experimental verification of mechanical processes, the structural stability and feasibility of the entire production line, from tray loading to final seedling emergence, were ensured.

Algorithmic performance: The improved YOLOv11-SEED model, deployed on the Hikvision VM platform, achieved high-efficiency recognition of pelleted carrot seeds. Experimental results demonstrated an average seeding accuracy of 97.72%, enabling real-time detection of missing or multiple seeding anomalies to ensure a high seedling emergence rate.

Economic viability: With a total hardware investment of only 4,000 RMB (comprising 1,000 RMB for the vision system and 3,000 RMB for mechanical components), the system offers an exceptionally low financial barrier for farmers and factories.

In conclusion, by leveraging open-source AI frameworks and cost-effective industrial hardware, this project provides a sustainable, high-precision, and affordable solution for modern agriculture. The integration of "hardware-software synergy" not only reduces recurring labor costs but also offers a rapid return on investment, significantly contributing to the intelligentization and digitalization of the seedling industry in China.

Project outlook and division of work: The research process represents not only a refinement of technical solutions but also a valuable pathway for intellectual growth and practical experience. Due to time and cost constraints, only part of the system functionality has been implemented at this stage. In future work, the project will be further improved, with particular emphasis on the optimization and full implementation of the material loading and unloading system. In addition, UG-based Model-Based Definition (MBD) will be employed to develop a digital twin model, providing robust support for integrated system control and process optimization.

Furthermore, future research will focus on integrating digital twin technology with the production line by abstracting the system into a mathematical model. This will enable the development of an adaptive production system capable of dynamic adjustment under varying operating conditions. By incorporating optimization algorithms, the system is expected to improve overall operational efficiency while ensuring equipment reliability and extending service life.

In terms of application expansion, the system will also be extended to accommodate different types of seeds, particularly non-pelleted seeds. This will require enhancements in seed handling mechanisms, precision control, and system robustness to address variations in seed properties, thereby improving the generalization and practical applicability of the proposed system.

Author's Contribution: Q.Z.: Responsible for electromechanical and experimental parts; Z.L.: Responsible for the visual framework; H.C.: Polish and check the article.

Declaration of Conflicting Interests: The author(s) declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Data Sharing Agreement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Chen, Y. 2023. Development of protected vegetable crops (Part II) Overview of China's Vegetable Mechanisation Development, 2022–2023. *Chin. Veg.*, 5: 5-10.

Chen, Z., G. Xie, X. Deng, J. Peng and H. Qiu. 2024b. DA-YOLOv7: a deep learning-driven high-performance underwater sonar image target recognition model. *J. Mar. Sci. Eng.*, 12: 1606.

Chen, Z., Z. He and Z.M. Lu. 2024a. DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Trans. Image Process.*, 33: 1002-1015.

Ding, X., Y. Zhang, Y. Ge, S. Zhao, L. Song, X. Yue and Y. Shan. 2024. Unireplknet: A universal perception large-kernel convnet for audio video point cloud time-series and image recognition. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5513-5524.

Han, J. and C. Moraga. 1995. The influence of the sigmoid function parameters on the speed of backpropagation learning. International Workshop on Artificial Neural Networks, 195-201.

Hao, X., L. Liu, R. Yang, L. Yin, L. Zhang and X. Li. 2023. A review of data augmentation methods of remote sensing image target recognition. *Remote Sens.*, 15: 827.

Jocher, G., A. Chaurasia and J. Qiu. 2023. Ultralytics YOLO (version 8.0.0). *Comput. Softw.*, 2023-1.

Kang, Z., F. Ma, C. Chen and J. Sun. 2024. YOSMR: A Ship Detection Method for Marine Radar Based on Customized Lightweight Convolutional Networks. *J. Mar. Sci. Eng.*, 12: 1316.

Li, Q. and H. Shi. 2024. Yolo-ge: An attention fusion enhanced underwater object detection algorithm. *J. Mar. Sci. Eng.*, 12: 1885.

Li, Q., W. Ma, H. Li, X. Zhang, R. Zhang and W. Zhou. 2024. Cotton-YOLO: Improved YOLOV7 for rapid detection of foreign fibers in seed cotton. *Comput. Electron. Agric.*, 219: 108752.

Luo, W., Y. Li, R. Urtasun and R. Zemel. 2016. Understanding the effective receptive field in deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.*, 29.

Ma, N., Y. Su, L. Yang, Z. Li and H. Yan. 2024. Wheat seed detection and counting method based on improved YOLOv8 model. *Sensors*, 24: 1654.

Niu, S., X. Xu, A. Liang, Y. Yun, L. Li, F. Hao, J. Bai and D. Ma. 2024. Research on a lightweight method for maize seed quality detection based on improved YOLOv8. *IEEE Access*, 12: 32927-32937.

Wang, P., S. Yang, G. Chen, W. Wang, Z. Huang and Y. Jiang. 2024. A ship's maritime critical target identification method based on lightweight and triple attention mechanisms. *J. Mar. Sci. Eng.*, 12: 1839.

Xin, Z., Y. Cui, X. Yang, L. Kong and Q. Lin. 2022. Current situation of global vegetable industry and research progress of vegetable breeding development path in China. *Mol. Plant Breed.*, 20: 3122-3132.

Xu, J., Z. Li, B. Du, M. Zhang and J. Liu. 2020. Reluplex made more practical: Leaky ReLU. 2020 IEEE Symposium on Computers and Communications, 1-7.

Yang, A. and B. Li. 2022. Machine vision-based crop sowing accuracy detection method. *J. Smart Agric.*, 2: 1-3.

Zhang, F., W. Cao, J. Gao, S. Liu, C. Li, K. Song and H. Wang. 2024a. Underwater object detection algorithm based on an improved YOLOv8. *J. Mar. Sci. Eng.*, 12: 1991.

Zhang, G., Y. Liu, L. Han, B. Li, H. Meng and S. Wang. 2024b. Research on the technology system and quality evaluation of vegetable plug seedling cultivation. *J. Anhui Agric. Sci.*, 1-6.

Zhou, H., H. Xia, C. Fan, T. Lan, Y. Liu, Y. Yang, Y. Shen and W. Yu. 2024. Intelligent detection method for surface defects of particleboard based on super-resolution reconstruction. *Forests*, 15: 2196.

Zou, X., D. Liu, F. Qing, X. Li, X. Xu and M. Gong. 2024. Factory seeding and nursery intelligent seeding performance detection technology application status and trend. *J. Intell. Agric. Mech.*, 1-12.