Seed Biology

ARTICLE

Open Access

https://doi.org/10.48130/seedbio-0026-0027

Seed Biology **2026**, 5: e030

An intelligent monitoring and warning system for seed germination in plant factories based on deep learning and computer vision

Shumao Wang[1,2], Luyang Kang[1*], Ping Zha[1], Jing Li[1], Hetong Zhang[1], Jigang Li[2] and Rongcheng Lin[1*]

[1] *Biotechnology Institute, Xianghu Laboratory, Hangzhou 311231, China*

[2] *State Key Laboratory of Plant Environmental Resilience, College of Biological Sciences, China Agricultural University, Beijing 100193, China*

* Correspondence: kangluyang@xhlab.ac.cn (Kang L); linrongcheng@xhlab.ac.cn (Lin R)

Abstract

Plant factories with artificial lighting (PFALs) are sophisticated engineering systems that integrate interdisciplinary approaches, including plant science, environmental control, and computer science. Previous studies have predominantly focused on monitoring plant phenotypes and the development of light, temperature, water, and fertilizer management technologies for PFALs, while insufficient attention has been paid to seed germination management. As a fundamental link in PFAL production, managing seed germination is of great importance, underscoring the imperative to develop an intelligent monitoring system. Using lettuce (*Lactuca sativa* L.) as the model plant, this study established an image dataset of 6,173 images covering seeds, germinating seeds (germination), and seedlings cultivated in PFALs. Subsequently, we designed depthwise separable wavelet transform convolutions and integrated them into *You Only Look Once, Version 11*(YOLOv11). After modification, the YOLOv11-s and YOLOv11-m models achieved mAP^{50-95} values of 93.29% and 95.26%, respectively, with both precision and recall exceeding 99.50%. In terms of different growth stages, the improved model achieved mAP^{50-95} values of 89.97%, 94.54%, and 95.37% for seeds, germinating seeds, and seedlings, respectively. Furthermore, an automatic warning mechanism was developed, which reached an average accuracy of 98.47%. Meanwhile, dedicated management software was developed to facilitate practical deployment. The system can be implemented in PFALs, enabling the adjustment and visualization of key monitoring parameters. This study followed a systematic workflow encompassing dataset construction, model establishment, application validation, and software development. It provides valuable references for the cultivation and breeding of other plants in vertical farm and PFALs.

Keywords: Plant factories; Seed germination; Deep learning; Computer vision; YOLOv11

Citation: Wang S, Kang L, Zha P, Li J, Zhang H, et al. 2026. An intelligent monitoring and warning system for seed germination in plant factories based on deep learning and computer vision. *Seed Biology* 5: e030 https://doi.org/10.48130/seedbio-0026-0027

Introduction

Plant factories with artificial lighting (PFALs) represent an advanced agricultural production scenario and are pivotal for the sustainable development of agriculture[1,2]. Plants cultivated in open fields are highly vulnerable to extreme weather events, including storms, heatwaves, and floods, which severely restrict plants' growth and yield[3]. PFALs enable precise regulation of light, temperature, water, gas, and fertilizer, allowing plants to grow in accordance with their inherent developmental rhythms, independent of the external climatic conditions[4–6]. Meanwhile, this cultivation mode reduces excessive application of chemical fertilizers, conserves water resources, and mitigates environmental pollution[7–9]. Owing to these advantages, PFALs are increasingly applied to secure fresh vegetable supplies in polar, marine, and alpine regions. Consequently, PFALs have garnered growing attention for their merits, including clean production, expanded arable land, high efficiency, and short growth cycles[6,10]. Against the backdrop of a growing global population and frequent extreme weather, PFALs offer a novel solution to address these pressing challenges.

Current research on PFALs primarily concentrates on several important domains. A major research direction is the optimization of light recipes tailored to diverse plant species, which mainly involves three key parameters: Light quality, light intensity, and photoperiod. These factors collectively regulate plants' photosynthesis and carbohydrate synthesis, thereby exerting profound effects on their growth and development. The objective of these studies is to identify the optimal light combination that maximizes plants' growth performance. A red–blue light mixture at a ratio of 1:0.8 under a 16/8 h (light/dark) photoperiod has been shown to significantly enhance photosynthetic efficiency and chlorophyll accumulation in tomato (*Solanum lycopersicum*) plants[11], while promoting their biomass accumulation and improving fruit quality, providing a theoretical basis for light regulation in tomato cultivation within PFALs. In addition, the design and regulation of nutrient solutions are critical for the stable operation of PFALs. pH, temperature, and solution concentration determine plants' growth quality, and nutrient requirements vary with plant species and growth stages. A 0.25× nutrient solution concentration combined with a light intensity of 350 $\mu mol \cdot m^{-2} \cdot s^{-1}$ has been identified as a more suitable condition for industrial lettuce (*Lactuca sativa*) production[12]. High-concentration polyphosphate nutrient solutions have also been demonstrated to boost lettuce's growth rates, including indices such as leaf area and root length[13]. Another vital research avenue is the application of PFALs as breeding accelerators. Compared with field conditions, PFALs can provide optimal growth environments for plants and can be further utilized in plant breeding. PFALs have been found to significantly accelerate the growth of rice (Oryza sativa), with *indica* rice reaching maturity and harvest in only 63 d[10], enabling five or six breeding generations per year. Airflow design is also a key factor influencing plant growth in PFALs. Poor air mobility can trigger plant diseases such as downy mildew, and CO_2 concentration represents another key factor affecting crop growth[14]. How to realize favorable airflow patterns while matching the light-emitting diode

© The Author(s)

(LED) lighting layout remains a practical problem to be solved for PFALs. Finally, artificial intelligence (AI) applications are becoming increasingly prevalent in PFALs. AI-optimized light and climate systems in PFALs have been shown to reduce energy consumption in both cool and warm climates[1]. However, despite the high demand for AI systems in PFALs, deep learning-based plant phenotyping studies in PFALs remain relatively scarce compared with those on field crops. Computer vision and deep learning techniques have been extensively applied to monitoring plant growth[15,16], but the monitoring of seed, germination, and seedling transition remains underexplored.

Previous studies on monitoring seed germination have been confined to ideal laboratory environments, with limited applicability to practical commercial production scenarios[17,18]. To date, research focusing on monitoring seed germination in PFALs remains scarce, making it highly significant to supplement this research area for improving intelligent PFAL systems. A previous study monitored the seed germination process in a small laboratory-scale growth chamber and proposed the SeedAI model[17]. This model adopts a two-stage deep learning framework: in the first stage, Mask R-CNN is used for instantaneous seed segmentation; in the second stage, the segmented regions are input into a customized convolutional neural network (CNN) model for classifying seed status. The model outperformed mainstream models such as ResNet50 and Inception in key metrics, including pixel accuracy (0.88), intersection over union (IoU, 0.77), and precision (0.89). However, it only supports single-category seed detection, and the monitored seeds are relatively large, simplifying the monitoring task. Additionally, it fails to achieve category-wide monitoring across the entire growth process from seed to germination to seedling. Furthermore, the study was conducted under laboratory conditions, and its performance in practical production applications requires further validation. The YOLO-seed germination detection (SGD) model was proposed based on the YOLOv7-l baseline[18], specifically addressing challenges in detecting cabbage seed germination, such as root entanglement and difficulty in small-object recognition. For cabbage seeds, the model achieved a mean average precision (mAP) of 98.0% and an F1 score of 93.3%. For other seeds, including tomato, eggplant, and pepper, mAP values all exceed 95%, with the tomato mAP reaching 97.8%. However, the model mainly focuses on fully germinated seeds and lacks the ability to monitor the intermediate germination process. Similarly, this study was conducted under controlled laboratory conditions, and its adaptability to complex production scenarios requires further verification.

Seed quality lays the foundation for consistent crop performance across plant production systems[19]. Reliable seed-germination monitoring, in turn, plays an indispensable role in both yield stability and plant breeding programs. In plant factories, conventional germination management predominantly relies on manual inspection and camera observations, which suffer from poor real-time responsiveness and are easily restricted by image resolution limitations and human subjective bias. For large-scale production and breeding inside PFALs, manual labor alone is incapable of supporting efficient, dynamic, and large-area monitoring of germination. AI has driven transformative changes in agriculture[20–22], creating new opportunities for intelligent monitoring of germination. Computer vision, a prominent branch of AI, plays a key role in monitoring plant phenotypes and plant breeding[23,24]. Therefore, this study leverages the distinct morphological states of seed germination and deep learning models to construct an AI-based real-time monitoring and warning system for germination status, offering a viable solution for cruise camera applications in plant production and breeding within large-scale PFALs. Using lettuce as the model plant, this study first constructed a dataset comprising images of seeds, germinated seeds, and seedlings. Next, YOLOv11 was adopted to detect the three classifications. Depthwise separable wavelet transform convolutions (DSWTCs) were designed and integrated into YOLOv11 to further improve detection performance. The model was validated in PFALs under different light formulas, and tracking of detected objects was implemented. An automatic warning system was developed, and dedicated management software was finally compiled. Users can visualize detailed detection information, including objects' location and confidence score. This work is expected to increase the intelligence and automation level of PFALs.

Materials and methods

Dataset establishment

PFALs used in this study

Seed germination experiments were conducted in a dedicated PFAL capable of fully independent regulation of the plants' growth climate and environment via a touchscreen control panel. Regulated parameters included temperature, humidity, light, and other key environmental factors. The light system consisted of white, red, blue, and near-infrared LEDs, with each light source adjustable and combinable to form composite spectra. Lettuce seeds were cultivated using hydroponics, and the nutrient solution was automatically diluted and circulated at a fixed ratio to maintain a stable nutrient supply. Ventilation units were vertically distributed inside the PFAL to promote air circulation and reduce the incidence of plant diseases.

Seed germination and image acquisition

Two lettuce cultivars, *L. sativa* 'Rex' and *L. sativa* 'Lalique', were sown in seedling trays filled with moist sponges. After sowing, the seedling trays were placed in a dark environment for the first 2 d to induce germination. Once the seeds entered the radicle protrusion stage (characterized by the emergence of a white tip), they were exposed to various LED light treatments and gradually developed into seedlings. Two camera types, Canon EOS 90D and Vivo Dual Cameras, were used to capture images throughout the seed germination process, with the detailed parameters listed in Supplementary Table S1. To enhance the model's robustness, seeds, germinated seeds (germination), and seedlings were separately placed under different dominant and mixed light conditions, including white light, blue light, red light, and composite light. The camera's height above the seedling trays was adjusted within the range of 15–30 cm (Supplementary Fig. S1).

Image annotation and dataset establishment

After removing duplicate images, 6,173 images were annotated with the LabelImg tool (tray-level). Three categories (seed, germination, and seedling) were labeled for object detection, as illustrated in Fig. 1a. The width and height distribution of all annotated bounding boxes is presented in Fig. 1b. Medium and large objects dominate the samples of the three classifications. The entire dataset was split into training, validation, and test datasets at a ratio of 8:1:1, as listed in Supplementary Table S2. Multiple data augmentation strategies were applied before model training, including color enhancement, geometric transformation, mosaic augmentation, and random erasing, which effectively improved the model's robustness.

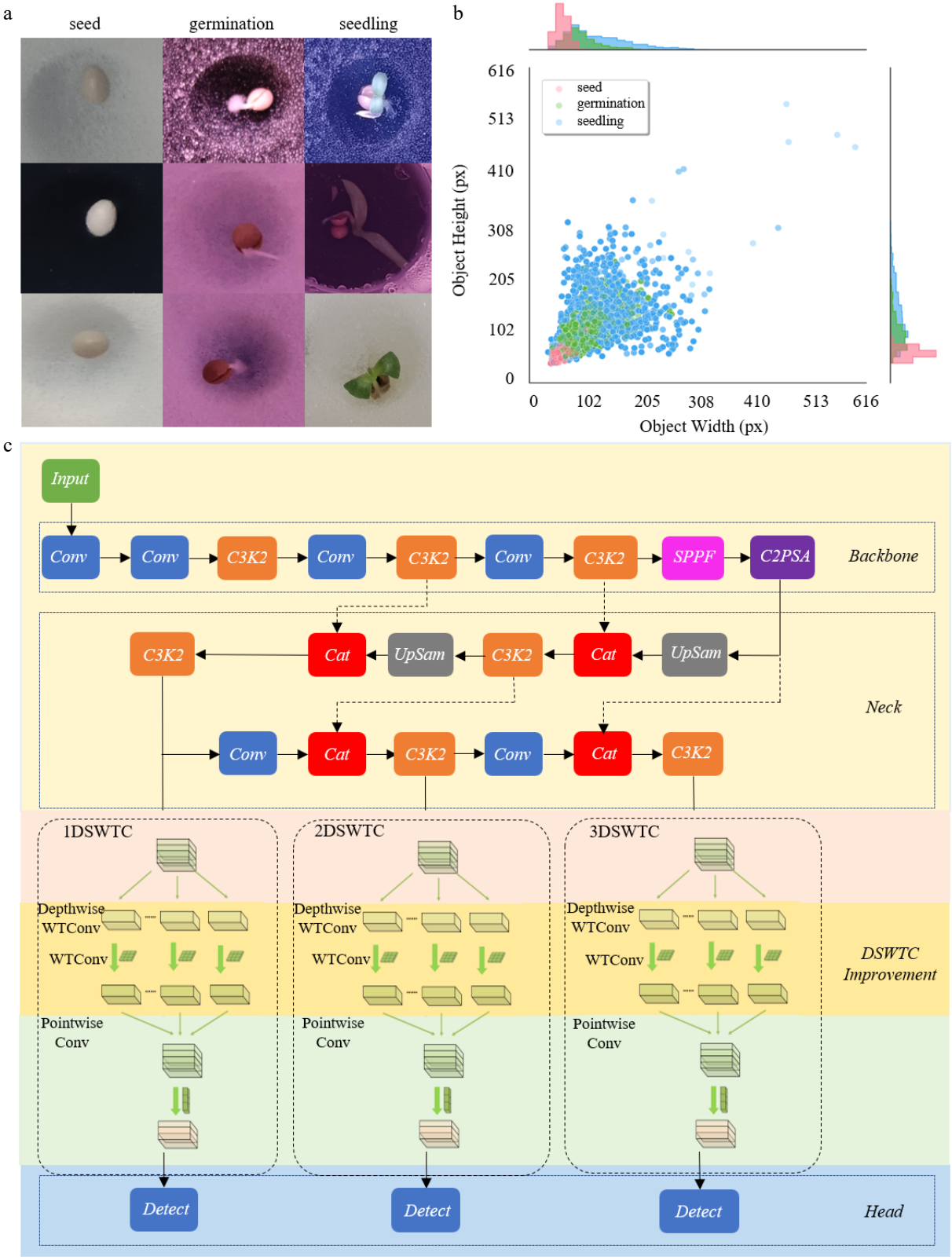


Fig. 1 Overview of the datasets and detection model. (a) Three growth stages: Seeds, germination and seedlings. (b) Size distribution of three types of annotation bounding boxes in the dataset. (c) The proposed YOLOv11-DSWTC object detection model in this study.

Model

Depthwise separable wavelet transform convolutions

Monitoring hydroponic lettuce from seed to two-cotyledon seedlings in plant factories faces multiple technical challenges. One challenge lies in the tiny size of lettuce seeds. There are subtle morphological differences between germinating and nongerminating seeds. Weak multiscale visual cues also exist throughout the germination and cotyledon emergence stages. Furthermore,

breeding applications require treatments with various light qualities, especially red and blue light. This poses considerable challenges to detection and monitoring. Conventional depthwise convolution (DWConv) suffers from a limited receptive field, insufficient low-frequency shape perception, and missing cross-channel information interaction, degrading the detection ability. In this study, wavelet transform convolutions (WTConvs) were integrated into DWConv to construct the DSWTC module[25,26]. The innovation of the proposed module lies in cascaded wavelet decomposition, which enlarges the model's effective receptive field and strengthens low-frequency feature perception and shape recognition capability for tiny germinating tissues. Benefiting from the inherent multiscale analysis property of WTConv, the module captures multiscale features more efficiently than conventional DWConv, and resolves DWConv's cross-channel fusion deficiency via multiscale collaborative feature processing to boost the monitoring of lettuce germination under fluctuating light.

Improvement of YOLOv11 with the DSWTC module

YOLOv11, released by Ultralytics in 2024[27], is a typical one-stage object detector evolved from previous YOLO frameworks that is capable of simultaneous object detection and localization. Beyond basic detection, it also supports semantic segmentation, image classification, pose estimation, and oriented object detection. The overall architecture of YOLOv11 consists of the input, backbone, neck, and head components (Fig. 1c). In the backbone network, the C3k2 module replaces the original C2f module for efficient extraction of critical image features. The SPPF module at the end of the backbone enhances small-object detection performance. A key upgrade of YOLOv11 is the integration of the C2PSA module for stronger multi-scale feature extraction. Built upon the classic CSP structure, C2PSA embeds the position-sensitive attention (PSA) mechanism. Each PSA block adopts multihead self-attention followed by a convolutional feed-forward network to model long-range spatial dependencies and perform adaptive weighting for spatial and channel features. In this study, DSWTC modules were embedded into YOLOv11 (Fig. 1c) to optimize the detection of seeds, germination, and seedlings.

Model evaluation

Precision, recall, mAP^{50-95}(mAP over IoU from 0.50 to 0.95 with a step of 0.05), F1 score, and overall accuracy (OA) were selected to evaluate the model's performance. The essential indices for evaluating the model are calculated as follows.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

where, *TP* indicates true positives and *FP* indicates false positives when the model detects seeds, germination, and seedlings.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

where, *FN* indicates the false negatives in the prediction.

$$mAP^{50-95} = \frac{1}{10}\int_{\theta=0.5}^{0.95} AP(\theta,\ \Delta\theta = 0.05) \tag{3}$$

where, *AP* is defined as follows:

$$AP(\theta) = \int_0^1 P(r)\,\mathrm{d}r \tag{4}$$

where, θ is threshold value of IoU, *r* is recall, and *P*(*r*) is the precision value for the recall *r*.

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

For better observation of the model's ability to apply in PFALs, *OA* with IoU and confidence value of 0.50 was introduced. The expression is as follows:

$$OA\ (\%) = \frac{N}{M} \times 100\% \tag{6}$$

where, *N* denotes the number of correctly warned objects and *M* denotes the total number of objects.

The warning mechanism

To enable effective monitoring of seed germination, an automatic warning mechanism was designed to alert managers to germination status, supporting intelligent management including environmental parameter optimization, improved space utilization, inactive seed elimination, and timely replanting. The simplified workflow of the mechanism is shown in Fig. 2. Commercially available conventional seeds normally achieve a germination rate of 80% or higher under optimal environments. For lettuce breeding scenarios, seed germination can be suppressed by diverse molecular regulatory effects and environmental stresses, which commonly lead to reduced actual germination performance. Thus, to improve the model's suitability for real-world application, a relatively conservative threshold (65%) was adopt rather than directly using the ideal condition standard. The threshold also could avoid recognition errors caused by individual differences in seed germination. When the program initiates, it automatically counts the total number of objects in each classification. During the germination process (Fig. 2), if only a single category is detected (seeds only, germination only, or seedlings only), the state is marked as 'Normal'. When both seeds and germination are detected simultaneously, the proportion of germination is calculated; if the proportion reaches or exceeds 65%, the seeds are marked 'Warning!'. When both germination and seedlings are detected together, the proportion of seedlings is computed; if the proportion reaches or exceeds 65%, the germinated seeds are labeled 'Warning!'. Furthermore, if seeds, germination, and seedlings are all detected concurrently, all seeds are directly marked 'DEAD!'. Meanwhile, the proportion of seedlings relative to the total number of germinated seeds and seedlings is calculated; if this proportion reaches or exceeds 65%, the germinated seeds are assigned a 'Warning!' label. If seeds and seedlings are both detected concurrently, all seeds are directly marked as 'DEAD!'. All detection results outside the abovementioned scenarios are uniformly marked 'Normal'. The system performs periodic monitoring every 24 h and updates the growth status in real time.

Experimental parameters

The experimental platform was equipped with an Intel® Core™ i7-14700KF central processing unit (CPU) (base frequency: 3.40 GHz), an NVIDIA GeForce RTX 4070 graphics processing unit (GPU) (12 GB dedicated video memory), and 32 GB of random access memory. The detailed training parameters are listed in Supplementary Table S3. Notably, no pretrained weights were used in all experiments to ensure a fair comparison of the model's performance. Notably, the random seed was set to 0 to guarantee experimental reproducibility.

Results

Performance of the DSWTC module in different YOLOv11 versions

Comparative experiments were conducted on YOLOv11 models with and without the DSWTC module. The trend of mAP^{50-95} with increasing training epochs is shown in Fig. 3a. As the model's scale increases from YOLOv11-n to YOLOv11-l, the model's convergence speed and mAP^{50-95} values rise gradually. Embedding the DSWTC

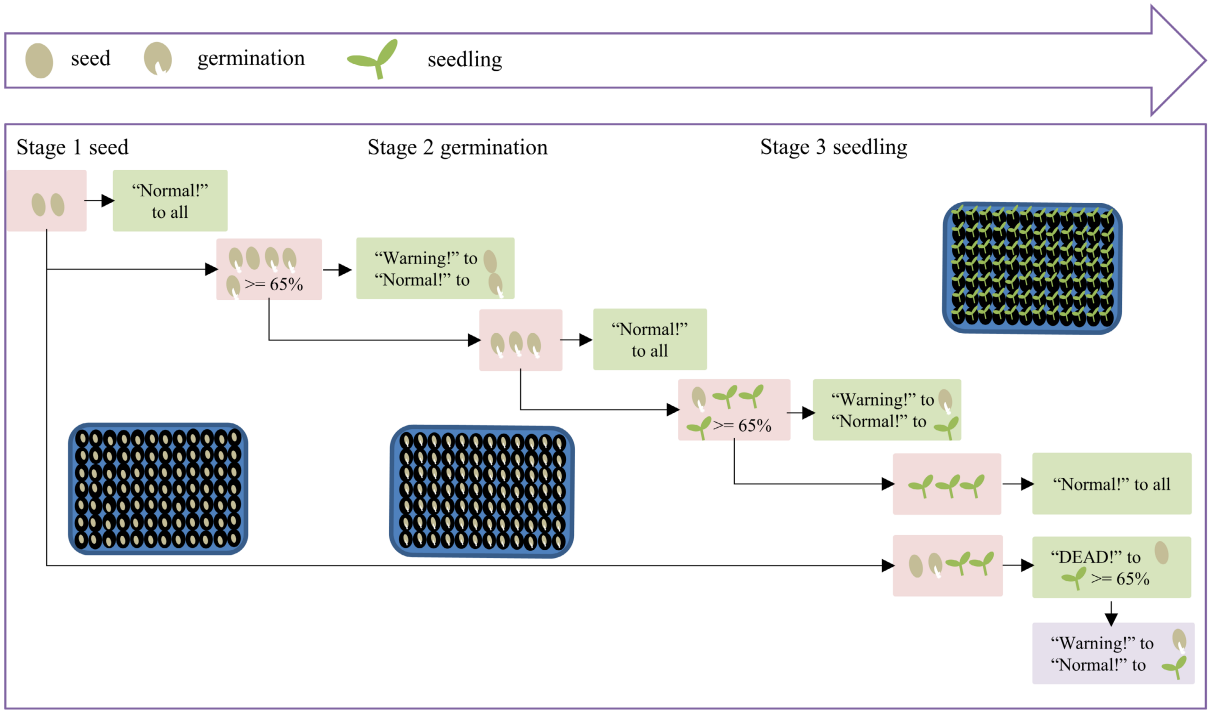


Fig. 2 Process of the warning mechanism.

module effectively boosts the overall performance of YOLOv11 models at all scales. As the model's size increases from YOLOv11-m to YOLOv11-l, the performance gain gradually narrows, indicating that the model's capacity approaches saturation. The specific performance gains in the first 50 epochs and the last 50 epochs are presented in Fig. 3b, c, respectively.

Comparative analyses of models with and without the DSWTC module were carried out on all datasets, generating comprehensive quantitative data. As listed in Table 1, on the test dataset, the original YOLOv11-n, YOLOv11-s, YOLOv11-m, and YOLOv11-l models yield mAP^{50-95} values of 89.73%, 93.01%, 94.79%, and 95.59%, respectively. After integrating the DSWTC module, the corresponding mAP^{50-95} values increase to 90.07%, 93.29%, 95.26%, and 95.69%. On the training dataset (Supplementary Table S4), the initial mAP^{50-95} values of the four models are 90.87%, 94.32%, 96.25%, and 96.84%, which rise to 91.13%, 94.60%, 96.62%, and 96.90% after DSWTC's integration. On the validation dataset (Supplementary Table S5), the original mAP^{50-95} values are 90.24%, 93.19%, 94.92%, and 95.73%. After adding the DSWTC module, the mAP^{50-95} of YOLOv11-n remains unchanged at 90.24%, although YOLOv11-l decreases slightly to 95.67%; the other two models achieve mAP^{50-95} values of 93.49% and 95.30%.

The detection performance for classifying seeds, germination, and seedlings is summarized in Fig. 3d–g. In terms of mAP^{50-95}, seedlings exhibit the highest detection accuracy, followed by germination and seeds. Embedding the DSWTC module into YOLOv11-n improves the detection performance of seeds and seedlings. Integration into YOLOv11-s enhances the detection accuracy for seeds and germination. When embedded in YOLOv11-m, the DSWTC module boosts detection performance for all three classifications. For YOLOv11-l, only the detection accuracy of seeds is improved, but that of the other two classifications declines. The DSWTC module delivers prominent performance gains for the classification of seeds and retains its effectiveness even in large-parameter models such as YOLOv11-l.

Radar charts of the minimum loss values on the training and validation datasets before and after embedding the DSWTC module are shown in Supplementary Fig. S2. During model training, the box loss of the improved models is generally lower than that of the original YOLOv11, with the same trend observed for classification loss. YOLOv11-n and YOLOv11-m achieve significant loss reduction on both the training and validation datasets. YOLOv11-n and YOLOv11-m exhibit clear optimization in box loss, classification loss, and distribution focal loss. YOLOv11-s mainly improves in terms of box loss and classification loss, with a slight decrease in distribution focal loss. In contrast, YOLOv11-l shows no notable loss reduction on either dataset and only minor optimization in training box loss, consistent with its limited mAP^{50-95} improvement after the integration of DSWTC. The most significant optimization after embedding the DSWTC module is observed in box loss, verifying that the DSWTC module accelerates model convergence and enhances detection performance. However, the DSWTC module shows limited improvement in the convergence of distribution focal loss.

Overall, the DSWTC module effectively enhances the overall detection capability of YOLOv11, according to the mAP^{50-95} results. Considering the practical application requirements of PFALs, YOLOv11-s combined with the DSWTC module was selected for real-world deployment. Further ablation experiments are still required before formal deployment to fully validate the effectiveness and robustness of the proposed improvement.

Ablation experiments of the DSWTC module

In light of the results described above, this section presents ablation experiments of the DSWTC module based on YOLOv11-s. The 1DSWTC, 2DSWTC, and 3DSWTC modules were embedded into the detection heads, corresponding, respectively, to large, medium, and small feature maps for detecting seeds, germination, and seedlings (Fig. 1c). The model achieves optimal performance when the DSWTC module is embedded in all three detection heads, with a mAP^{50-95} of 93.29%, and precision, recall, and F1 scores of 99.86%, 99.77%, and

https://doi.org/10.48130/seedbio-0026-0027

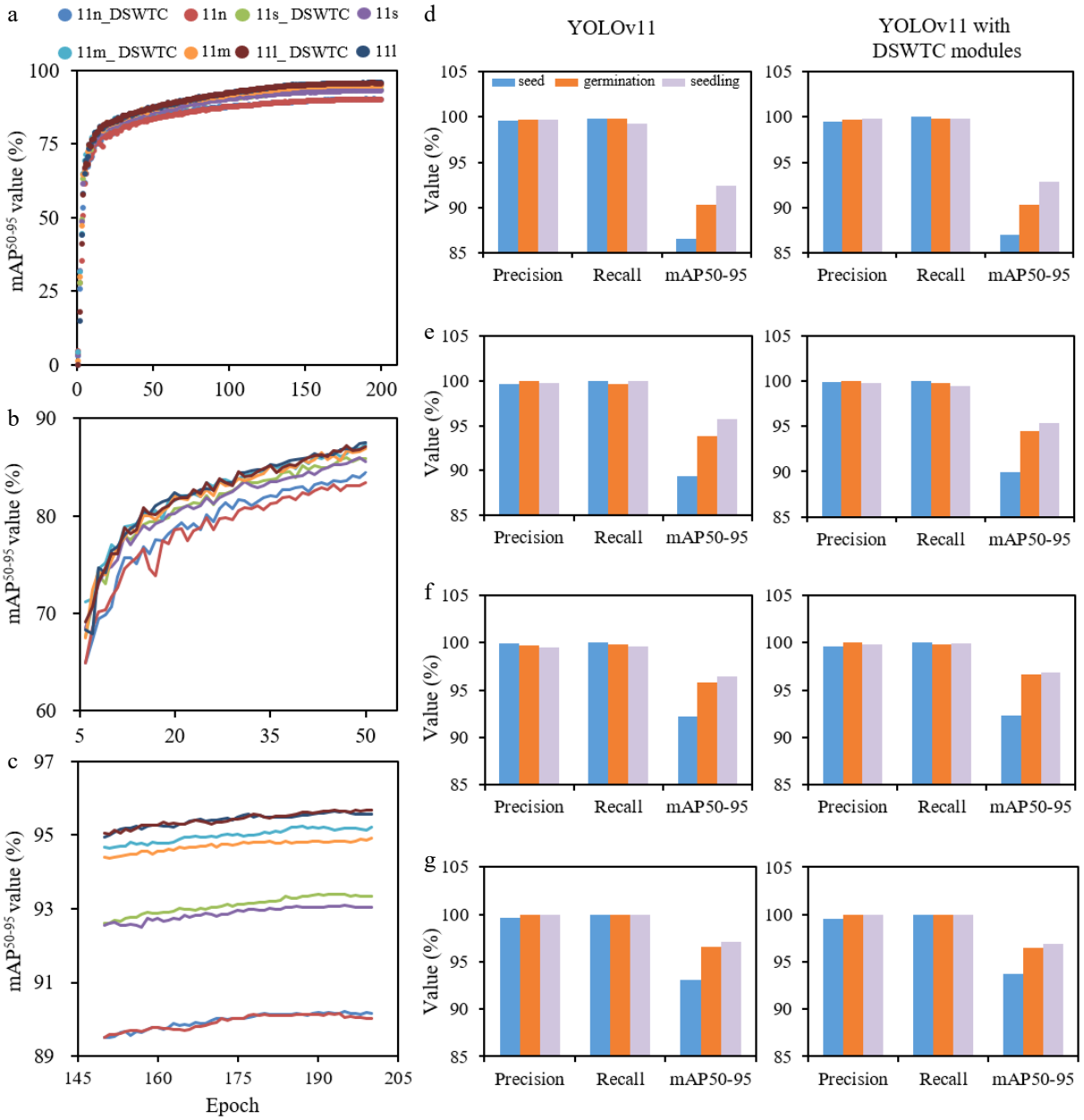


Fig. 3 mAP^{50-95} changes of different YOLOv11 versions with and without the DSWTC module, and detection performance for seeds, germination, and seedlings. The overall mAP^{50-95} values in (a) 0–200 epochs, (b) 5–50 epochs, and (c) 150–200 epochs, respectively. (d), (e), (f), and (g) showed the seed, germination, and seedling detection performance in different sizes before and after embedding the DSWTC module on test dataset.

Table 1. Overall evaluation parameters of different-sized models on the test dataset before and after embedding the DSWTC module.

Model	Precision (%)	Recall (%)	F1 score (%)	mAP^{50-95} (%)
YOLOv11-n	99.70	99.60	99.65	89.73
+DSWTC	99.65	99.86	99.75	90.07
YOLOv11-s	99.85	99.84	99.85	93.01
+DSWTC	99.86	99.77	99.81	93.29
YOLOv11-m	99.73	99.80	99.76	94.79
+DSWTC	99.79	99.88	99.83	95.26
YOLOv11-l	99.86	99.99	99.92	95.59
+DSWTC	99.83	99.96	99.90	95.69

99.81%, respectively (Supplementary Table S6). In contrast, the worst performance occurs when the DSWTC module is embedded only in the medium and large object detection heads, yielding a mAP^{50-95} of 92.70%, with precision, recall, and F1 scores of 99.75%, 99.88%, and 99.81%. Ablation results were also evaluated on the training and validation datasets. For the training dataset (Supplementary Table S7), the model with the DSWTC module embedded in all three detection heads achieves the highest mAP^{50-95} of 94.60%, with precision, recall, and F1 scores of 99.97%, 99.98%, and 99.98%, ranking first among all DSWTC embedding schemes. Embedding the DSWTC module in the small and large object detection heads produces a mAP^{50-95} of 94.58%, slightly lower than the full embedding scheme. The 23DSWTC scheme (the DSWTC module inserted into both the medium and large object detection heads) produces the lowest mAP^{50-95} of 94.28%.

In the validation dataset (Supplementary Table S8), the model with the DSWTC module embedded in all detection heads still performs best, with a mAP^{50-95} of 93.49% and corresponding

precision, recall, and F1 scores of 99.93%, 99.91%, and 99.92%. The lowest mAP^{50-95} among all schemes is 93.04%. Overall, the scheme with the DSWTC module embedded in all three detection heads maintains the best comprehensive performance across the training, validation, and test datasets. Moreover, embedding the DSWTC module into the small-object detection head consistently improves detection accuracy on all three datasets. The variation curves of mAP^{50-95} across training epochs are presented in Supplementary Fig. S3. In the early training stage, all models show little difference in their mAP^{50-95} values, with no sustained numerical superiority, and their performance fluctuations remain relatively consistent (Supplementary Fig. S3a). From Epoch 150 to 200, the model equipped with the DSWTC module gradually outperforms other schemes (Supplementary Fig. S3b).

The confusion matrix illustrates the distribution of model predictions across the seed, germination, seedling, and background categories (Supplementary Fig. S4a). The original model without the DSWTC module produced five false detections. After embedding the DSWTC module into all three detection heads, the number of false detections decreased to four. Different DSWTC module embedding positions resulted in 5–8 false detections (Supplementary Fig. S4b). The confusion matrix also reveals obvious background mispredictions: The full DSWTC embedding scheme has only two background false predictions, whereas the 1DSWTC scheme has five, and other schemes produce three or four mispredictions. Notably, the original model already had a low error rate, and the improved model further reduced the total number of prediction errors, verifying the effectiveness of DSWTC module optimization.

Application of the improved model in PFALs

Detection performance under different LED light formulas

In this subsection, the performance of the improved model under different LED light formulas is explored further (Fig. 4). A 7 × 7 box was used to intuitively visualize the detection results. Two typical LED light environments were set in this study (Supplementary Fig. S5), one with a reddish spectrum and the other with a bluish spectrum. The corresponding confusion matrices and quantitative detection results are displayed in Supplementary Fig. S6.

Supplementary Fig. S5a shows the results of seed detection under the two light conditions. All seeds are accurately identified in both reddish and bluish light environments. The scene in Supplementary Fig. S5b contains three seeds and multiple germinated samples. Under reddish light, the model correctly detects two seeds and

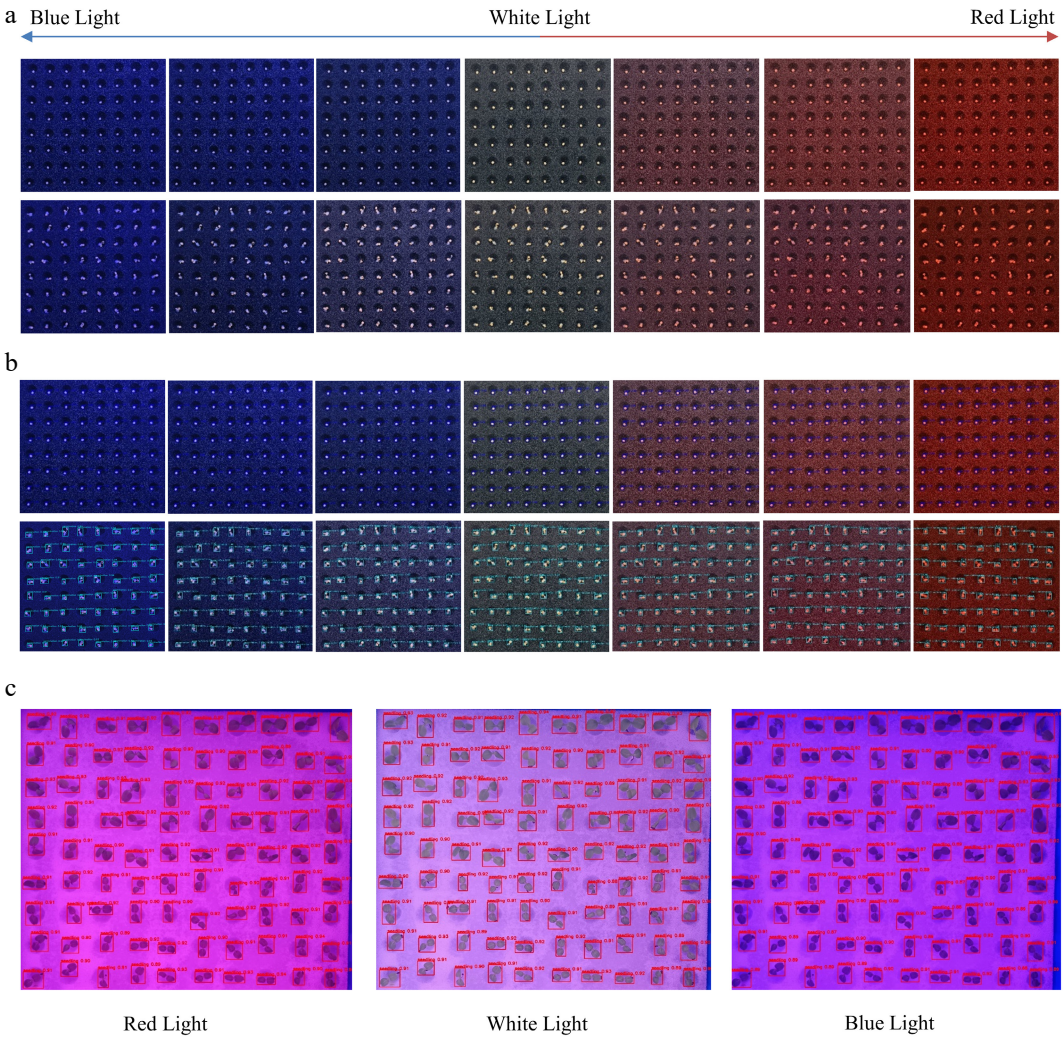


Fig. 4 Seed, germination, and seedling detection under different LED formulas in the PFALs. (a) Seed and germination under different LED lights. (b) Detection performance under the corresponding LED lights. (c) For better observation, the results of seedling detection performance are presented separately.

misclassifies one seed as germination. Under bluish light, the model additionally misidentifies one germinated sample as a seed. The scenario in Supplementary Fig. S5c includes one seed, one germination, and several seedlings. The model successfully detects all seedlings in both light environments. However, one seed is misclassified as a germinated sample under both reddish and bluish light. This study further constructs a more complex mixed scene to validate the model's ability to generalize (Supplementary Fig. S5d). The model achieves satisfactory overall detection performance, accurately recognizing all seven seeds and nine germinated samples in this scene.

These results demonstrate that the proposed model exhibits high robustness for practical application in PFALs. Furthermore, multiple heatmaps were adopted for visualizing features (Supplementary Fig. S7) to reveal the response distribution of predictive features. This approach enables the observation of features' response intensity at different levels and an analysis of features' contributions to object localization and classification confidence. GradCAM, GradCAMPlus-Plus, and XGradCAM were adopted to visualize the feature prediction ability for seeds, germination, and seedlings, presenting features' responses at the bounding box and classification levels, respectively.

Application of the warning mechanism

This subsection presents typical warning cases to illustrate the practical application of detection results for managing seed germination in PFALs. A 7 × 7 box was used for clearer observation and presenting the results. The embedded program automatically counts the number of seeds, germinated examples, and seedlings within the monitoring area.

When only a single classification is detected, the warning mechanism marks the status as 'Normal'. As shown in Fig. 5a-I, after sowing, only seeds are identified, and the system assigns 'Normal' labels. In the germination stage (Fig. 5a-II), both seeds and germination are detected. When the proportion of germinated plants reaches or exceeds 65%, the remaining seeds are marked 'Warning!' to alert managers to slowly germinating seeds. Similarly, when germinating plants develop into seedlings (Fig. 5a-III), the system detects both germination and seedlings. If the seedling proportion reaches or exceeds 65%, the corresponding germinated samples are labeled 'Warning!'. Figure 5a-IV shows an extreme scenario where seeds, germination, and seedlings coexist. Once seeds are detected in this mixed state, the warning mechanism directly marks these seeds as 'DEAD!'. After excluding dead seeds from the statistics, the system calculates the proportion of seedlings among the remaining germination and seedlings; if this proportion reaches or exceeds 65%, the germinated samples are labeled 'Warning!'.

For a clearer visual comparison, Fig. 5a-V–VIII are provided as enlarged subgraphs corresponding to Fig. 5a-I–IV, intuitively demonstrating the models' detection performance. As summarized in Table 2, all OA values exceed 95.00%. Specifically, OA values of 100.00% were achieved in Fig. 5a-I and IV; 95.92% was reached in Fig. 5a-II; and 97.96% were obtained in Fig. 5a-III. The difference in performance between Fig. 5a-II and III are mainly reflected in correctly warned objects counts (47 and 48, respectively).

Object tracking and software development

To accommodate supervision of daily seed germination in large-scale PFALs with mobile cameras, DeepSort was introduced for object tracking (Supplementary Fig. S8). The model first detects seedlings and locates each object at the center of its predicted bounding box. As the mobile camera moves, the object tracking function activates, and colored trajectory lines visualize the movement path of each detected object. Meanwhile, the complete motion trajectory of each object is continuously recorded.

In addition, dedicated monitoring software was developed, consisting of a login page and a management page (Supplementary Fig. S9). On the management page, multiple alternative models are provided in the 'Model Selection' module. The IoU and confidence thresholds can be adjusted via sliders. The software supports both image and video processing according to the users' demands. Detection and warning results are displayed in a data table, which records the classification, coordinate position, confidence score, and abnormal status of each object.

Discussion

Comparison with other YOLO series models

The development of modern object detection began with AlexNet, which laid the foundation for deep learning applications in computer vision[28]. The R-CNN series established the classic two-stage detection paradigm by integrating deep CNNs with region proposal mechanisms, and the subsequent Fast R-CNN and Faster R-CNN algorithms effectively optimized detection efficiency and accuracy[29–31]. In contrast, the YOLO algorithm pioneered a unified regression-based single-stage detection paradigm, removing the separate region generation branch, enabling end-to-end real-time detection with balanced speed and precision[32]. In recent years, object detection has evolved in two major directions: anchor-free methods such as FCOS and CenterNet that improve generalization for targets with irregular shapes and variable sizes by abandoning predefined anchors, and transformer-based detectors represented by DETR[33–34]. Among these mainstream frameworks, the YOLO series stands out for its lightweight architecture, fast inference speed, and excellent deployment compatibility, making it particularly suitable for real-time monitoring of tiny lettuce seeds in PFAL environments. Therefore, this study has adopted YOLOv11-s as the baseline model for detecting seed germination.

Comparisons with YOLOv12-s[35] and YOLOv13-s[36] on the seed germination dataset were conducted, with the quantitative results presented in Table 3 and Fig. 5b. The YOLOv11-DSWTC achieves an F1-score of 99.81% and a $mAP^{50–95}$ of 93.29%, which outperforms YOLOv12 (92.41%) and is slightly inferior to YOLOv13 (93.47%). It delivers the fastest inference speed at 166.16 FPS, much higher than YOLOv12 (156.25 FPS) and YOLOv13 (125.58 FPS). The $mAP^{50–95}$ values for germination and seedlings are both above 90%, whereas that for seeds reaches 89.97%, close to 90%. The remaining metrics also exhibit favorable performance for all categories (Supplementary Table S9). Notably, Ultralytics officially discourages the adoption of YOLOv13 in practical applications, reporting training instability, increased architectural complexity, elevated inference latency, and nonreproducible benchmark results relative to YOLOv11, which partly motivated the selection of YOLOv11 as the backbone for improvement in this study. Nevertheless, YOLOv11-DSWTC has its limitations. It has the largest floating point operations (FLOPs, 22.00 G) and model file size (9.83 MB) among the three competitors, leading to relatively higher computational and storage consumption. How to reduce the computational cost without degrading detection accuracy is an important direction for future improvements. The effectiveness of the DSWTC module is derived from wavelet transform convolution, which captures both high-frequency texture details and low-frequency contour features corresponding to

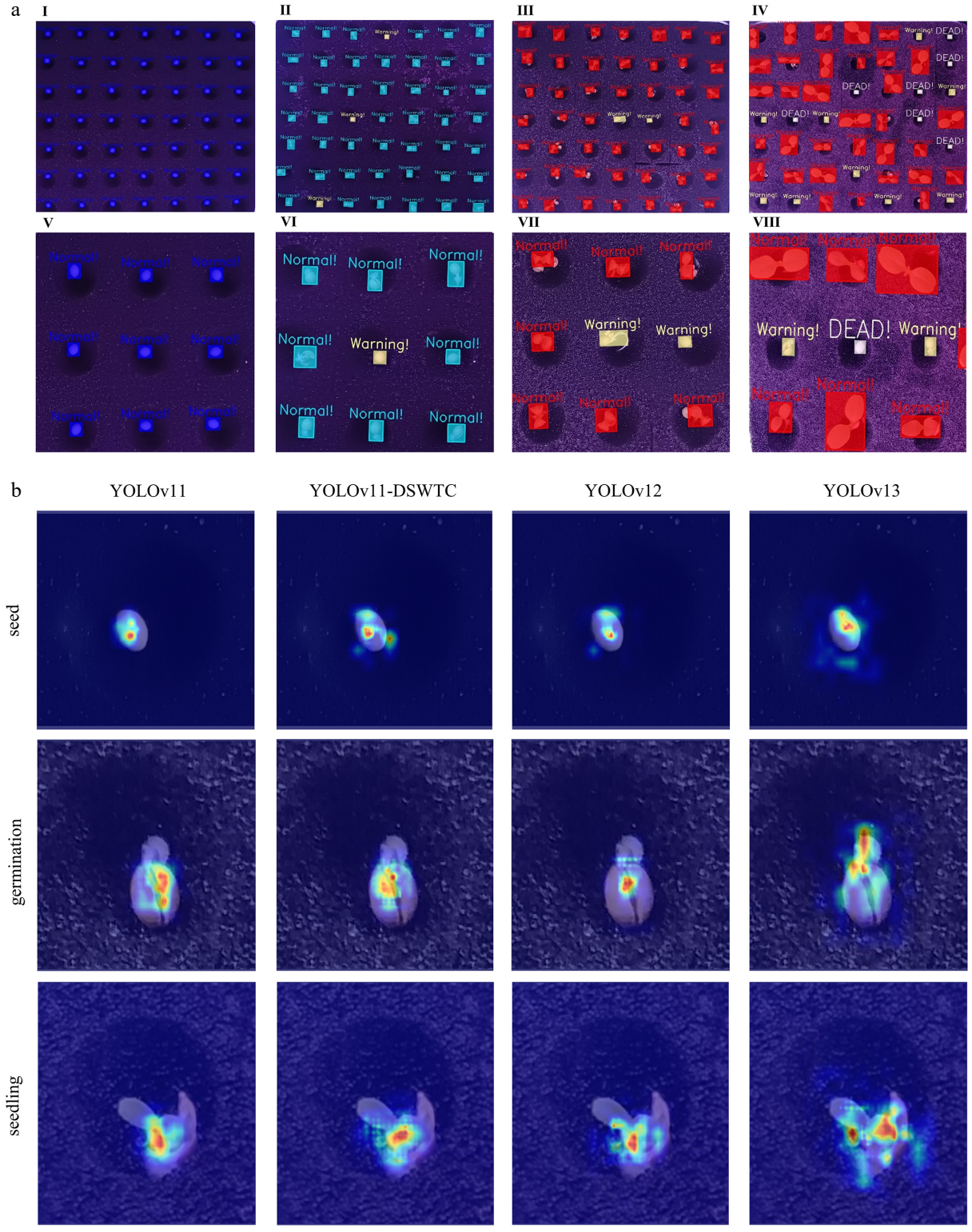


Fig. 5 Application of the warning mechanism in different scenes and comparison of performance with other advanced models. (a) I–IV are applications of the algorithm in different scenes; V–VIII are partly enlarged figures of I–IV, respectively. (b) Heatmap-based comparison of small versions among different YOLO models.

different germination stages[26]. Meanwhile, the embedded depth-wise structure enables efficient multiscale feature extraction, further improving the model's feature learning efficiency and adaptability to complex lighting environments inside PFALs. Despite the improved model limitations in parameter size and computational cost, the model exhibits little performance discrepancy in real-world deployment. It maintains strong adaptability to the complex and dynamic lighting environments of large-scale PFALs.

Table 2. Overall accuracy of seed germination warnings in Fig. 5a-I–IV.

Image	M	N	OA (%)
I	49	49	100.00
II	49	47	95.92
III	49	48	97.96
IV	49	49	100.00
		Average	98.47

Table 3. Comparison experiment with other small versions of YOLO.

Model	P (%)	R (%)	F1 (%)	$mAP^{50–95}$ (%)	Inference speed (FPS)	FLOPs (G)	Size (MB)
YOLOv11-DSWTC	99.86	99.77	99.81	93.29	166.16	22.00	9.83
YOLOv12	99.79	99.76	99.77	92.41	156.25	21.20	9.12
YOLOv13	99.92	99.76	99.84	93.47	125.58	20.70	9.00

Note: P, R, and F1 are precision, recall, and F1 score, respectively. FLOPs denotes floating-point operations.

Practical innovation and application value for monitoring germination in PFALs

Compared with open-field cultivation, seeds in conventional farmland are usually covered by soil after sowing, which hinders nondestructive observation of the germination process and limits real-time perception of early growth status. Most relevant evaluations can only be performed after seedlings' emergence, resulting in insufficient research progress on monitoring seed germination. In sharp contrast, PFALs, particularly hydroponic systems, provide an occlusion-free growing environment that supports nondestructive, full-cycle, and continuous monitoring of seed germination. In such systems, seeds germinate on moist nursery substrates without soil coverage, enabling dynamic recording of the complete developmental process, including seed imbibition, radicle protrusion, radicle elongation, and cotyledon expansion, and thus supporting refined germination management. Nevertheless, research targeting the monitoring of seed germination within PFALs remains relatively limited, highlighting the necessity and significance of further enriching this research field to advance the intelligent operation of plant factory systems.

Although previous studies have achieved reliable detection performance under laboratory constraints[17–18], most lack deployment-related exploration and verification of robustness in real production contexts. Unlike those laboratory-oriented investigations, the present study targets the practical demands of large-scale seedling cultivation in commercial PFALs. Beyond conventional identification of seeds' germination states, the established system integrates a production-oriented warning function that was rarely included in previous studies. Moreover, a dedicated object tracking module is incorporated to adapt to the imaging characteristics of cruising cameras, which is more consistent with the actual spatial layout and automatic monitoring mode of PFAL production lines. Overall, this work constructs a complete technical workflow covering dataset establishment, model optimization, and on-site deployment, effectively bridging the gap between algorithm-oriented laboratory research and practical industrial application in the field of monitoring seed germination in PFALs.

Limitations and future work

This study still has several limitations despite its achievements in monitoring seed germination in PFALs. First, the current algorithm only generates discrete categorical warning labels and fails to quantitatively describe dynamic physiological changes of the seed during germination. Such discrete evaluations cannot reflect subtle variations in seed vigor, thus providing limited support for precise environmental regulation. Second, the integration between the monitoring platform and the PFAL environmental control system is inadequate. At this stage, the model only has a warning function, without forming an automatic control strategy linked to temperature, humidity, and other key parameters, meaning that environmental regulation still depends on manual operation.

Future work will focus on further improving the practicality and intelligence of the monitoring system. Although this multiobject detection approach yields improved performance under the stringent $mAP^{50–95}$ metric, these gains come at the expense of increased computational overhead. Future work will focus on targeted optimizations to mitigate this issue. In addition, subsequent optimization will abandon simple discrete labels and construct a continuous quantitative evaluation system based on physiological indicators of germination such as radicle emergence and hypocotyl growth. Combining multidimensional feature information extracted by the improved model will enable refined quantitative perception of seeds' status. Furthermore, an integrated framework of monitoring and automatic regulation will be established to realize real-time adaptive adjustment of environmental parameters according to dynamic germination changes, thereby achieving full-process intelligent management for seedlings cultivated in commercial PFALs.

Conclusions

As a viable pathway to mitigate global food insecurity and facilitate crop breeding research, plant factories still suffer from a lack of intelligent monitoring systems for seed germination. This work upgrades YOLOv11 to boost performance in detecting seed germination. It also builds a early warning algorithm grading seeds' health states and integrates DeepSort for continuous tracking of seedling targets. A dedicated monitoring software package is finally developed to visualize detailed seed growth data for practical plant factory deployment.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Lin R, Jigang Li; data collection: Wang S, Zha P, Jing Li; analysis and interpretation of results: Wang S, Kang L, Zhang H; draft manuscript preparation: Wang S, Lin R, Kang L. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by the Leading Innovation and Entrepreneurship Team Project of Hangzhou City (TD2024008), National Natural Science Foundation of China (32402657), Key R&D Program of Hangzhou City (2025SZD1A17), and Zhejiang Province Postdoctoral Research Selective Funding Project (ZJ2025137). We thank the handling editor and the reviewers for their professional suggestions in manuscript improvement.

Conflict of interest

The authors declare no conflict of interest.

Supplementary information accompanies this paper online at: https://doi.org/10.48130/seedbio-0026-0027.

Dates

Received 30 June 2026; Revised 6 September 2026; Accepted 6 October 2026; Published online 19 November 2026

References

[1] Decardi-Nelson B, You F. 2024. Artificial intelligence can regulate light and climate systems to reduce energy use in plant factories and support sustainable food production. *Nature Food* 5(10):869–881
[2] Martin M. 2024. AI-driven optimization in plant factories. *Nature Food* 5(10):805–806
[3] Du S, Xiong W. 2024. Weather extremes shock maize production: current approaches and future research directions in Africa. *Plants* 13(12):1585
[4] Ahamed MS, Sultan M, Monfet D, Rahman MS, Zhang Y, et al. 2023. A critical review on efficient thermal environment controls in indoor vertical farming. *Journal of Cleaner Production* 425:138923
[5] Chen WH, Decardi-Nelson B, Kubota C, You F. 2025. AI applications in the environmental control of controlled environment agriculture in the digital age. *Modern Agriculture* 3(2):e70027
[6] Zhang L, Yang X, Li T, Gan R, Wang Z, et al. 2022. Plant factory technology lights up urban horticulture in the post-coronavirus world. *Horticulture Research* 9:uhac018
[7] Asseng S, Guarin JR, Raman M, Monje O, Kiss G, et al. 2020. Wheat yield potential in controlled-environment vertical farms. *Proceedings of the National Academy of Sciences of the United States of America* 117(32):19131–19135
[8] Pinstrup-Andersen P. 2018. Is it time to take vertical indoor farming seriously? *Global Food Security* 17:233–235
[9] Guo M, Jin Z, Ma L, Ou S. 2026. Application of plant factory with artificial lighting in horticultural production: current progress and future trends. *Horticultural Plant Journal* 12(4):747–763
[10] Liu Y, Li ZG, Cheng H, Yang X, Li MY, et al. 2025. Plant factory speed breeding significantly shortens rice generation time and enhances metabolic diversity. *Engineering* 50:259–269
[11] Zhang Y, Zhu K, Wang X, Yan J, Zhu H, et al. 2025. Manipulation of artificial light environment improves plant biomass and fruit nutritional quality in tomato. *Journal of Advanced Research* 75:79–93
[12] Song J, Huang H, Hao Y, Song S, Zhang Y, et al. 2020. Nutritional quality, mineral and antioxidant content in lettuce affected by interaction of light intensity and nutrient solution concentration. *Scientific Reports* 10(1):2796
[13] Lyu KS, Hwang HS, Jeong HW, Hwang SJ. 2024. Application of a polyphosphate fertilizer treatment enhances the growth rate of lettuce seedlings in a commercial plant factory. *Horticultural Science and Technology* 42(4):389–398
[14] Kang L, Zhang Y, Kacira M, van Hooff T. 2024. CFD simulation of air distributions in a small multi-layer vertical farm: impact of computational and physical parameters. *Biosystems Engineering* 243:148–174
[15] Ge X, Wu S, Wen W, Shen F, Xiao P, et al. 2025. LettuceP3D: a tool for analysing 3D phenotypes of individual lettuce plants. *Biosystems Engineering* 251:73–88
[16] Wang Z, Li T, Du R, Yang N, Ping J. 2025. A high-efficiency lettuce quality detection system based on FPGA. *Computers and Electronics in Agriculture* 231:109978
[17] Reddy DR, Cheruku R, Kodali P. 2023. SeedAI: a novel seed germination predictionsystem using dual stage deep learning framework. *Environmental Research Communications* 5(12):125014
[18] Yang T, Peng B, You L, Zhang J, Zhang D, et al. 2025. YOLO-SGD: precision-oriented intelligent detection of seed germination completion. *Agronomy* 15(9):2146
[19] Xiong M, Zhang H, Huang Y, Li Y. 2024. G protein regulation in rice seed biology. *Seed Biology* 3:e010
[20] Wang SM, Ma JH, Zhao ZM, Yang HZY, Xuan YM, et al. 2023. Pixel-class prediction for nitrogen content of tea plants based on unmanned aerial vehicle images using machine learning and deep learning. *Expert Systems with Applications* 227:120351
[21] Agarwal A, de Jesus Colwell F, Dinnis R, Correa Galvis VA, Hill TR, et al. 2025. Infrared thermography in plant factories: solving spatiotemporal variations via machine learning. *Modern Agriculture* 3:e70012
[22] Wang S, Kang L, Zhu Q, Zha P, Li J, et al. 2026. Chuiyuan: a large language model-driven AI agent for plant factories-applications and perspectives. *Modern Agriculture* 4:e70044
[23] Jiang Y, Li C. 2020. Convolutional neural networks for image-based high-throughput plant phenotyping: a review. *Plant Phenomics* 2020:4152816
[24] Wang SM, Yu CP, Ma JH, Ouyang JX, Zhao ZM, et al. 2024. Tea yield estimation using UAV images and deep learning. *Industrial Crops and Products* 212:118358
[25] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, et al. 2017. MobileNets: efficient convolutional neural networks for mobile vision applications. *arXiv* Prprint:1704.04861
[26] Finder SE, Amoyal R, Treister E, Freifeld O. 2024. Wavelet convolutions for large receptive fields. In *Computer Vision – ECCV 2024*, eds. Leonardis A, Ricci E, Roth S, Russakovsky O, Sattler T, et al. Cham: Springer. pp. 363–380 doi: 10.1007/978-3-031-72949-2_21
[27] Ultralytics. YOLOv11. https://github.com/ultralytics/ultralytics [EB/OL]. 2025
[28] Krizhevsky A, Sutskever I, Hinton GE. 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM* 60(6):84–90
[29] Girshick R, Donahue J, Darrell T, Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition, 23–28 June, 2014, Columbus, OH*. USA: IEEE. pp. 580–587 doi: 10.1109/CVPR.2014.8
[30] Girshick R. 2015. Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV), 7–13 December 2015, Santiago, Chile*. USA: IEEE. pp. 1440–1448 doi: 10.1109/iccv.2015.169
[31] Ren S, He K, Girshick R, Sun J. 2015. Faster R-CNN: towards real-time object detection with region proposal networks. *arXiv* Preprint:1506.01497
[32] Redmon J, Divvala S, Girshick R, Farhadi A. 2016. You only look once: unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27–30 June, 2016, Las Vegas, NV, USA*. USA: IEEE. pp. 779–788 doi: 10.1109/CVPR.2016.91
[33] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, et al. 2017. Attention is all you need. *arXiv* Preprint:1706.03762
[34] Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, et al. 2020. End-to-end object detection with transformers. In *Computer Vision – ECCV 2020*, eds. Vedaldi A, Bischof H, Brox T, Frahm JM. Cham: Springer. pp. 213–229 doi: 10.1007/978-3-030-58452-8_13
[35] Tian Y, Ye Q, Doermann D. 2025. YOLOv12: attention-centric real-time object detectors. *arXiv* Preprint:2502.12524
[36] Lei M, Li S, Wu Y, Hu H, Zhou Y, et al. 2025. YOLOv13: real-time object detection with hypergraph-enhanced adaptive visual perception. *arXiv* Preprint:2506.17733

Copyright: © 2026 by the author(s). Published by Maximum Academic Press on behalf of Hainan Yazhou Bay Seed Laboratory. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit https://creativecommons.org/licenses/by/4.0/.