# RESEARCH

# **Open Access**



Xuefei Chen<sup>1</sup>, Shouxin Sun<sup>1</sup>, Chao Chen<sup>2</sup>, Xinlong Song<sup>1</sup>, Qiulan Wu<sup>1\*</sup> and Feng Zhang<sup>1\*</sup>

# Abstract

A small object Lentinus Edodes logs contamination detection method (SRW-YOLO) based on improved YOLOv7 in edge-cloud computing environment was proposed to address the problem of the difficulty in the detection of small object contaminated areas of Lentinula Edodes logs. First, the SPD (space-to-depth)-Conv was used to reconstruct the MP module to enhance the learning of effective features of Lentinula Edodes logs images and prevent the loss of small object contamination information, and improve the detection reliability of resource-limited edge devices. Meanwhile, RepVGG was introduced into the ELAN structure to improve the efficiency and accuracy of inference on the contaminated regions of Lentinula Edodes logs through structural reparameterization. This enables models to run more efficiently in mobile edge computing environments while reducing the burden on cloud computing servers. Finally, the boundary regression loss function was replaced with the WIoU (Wise-IoU) loss function, which focuses more on ordinary-quality anchor boxes and makes the model output results more accurate. In this study, the measures of Precision, Recall, and mAP@0.5 reached 97.63%, 96.43%, and 98.62%, respectively, which are 4.62%, 3.63%, and 2.31% higher compared to those for YOLOv7. Meanwhile, the SRW-YOLO model detects better compared with the current advanced one-stage object detection model, providing an efficient, accurate and practical small object detection solution in mobile edge computing environments and cloud computing scenarios.

**Keywords** Lentinula Edodes logs, Contamination detection, Mobile edge computing, Cloud computing, Small object detection

# Introduction

Lentinula Edodes logs are critical carriers for Lentinula Edodes cultivation and are frequently contaminated by sundry bacteria during the cultivation process [1-3], causing substantial economic losses to enterprises. Currently, the contamination status of Lentinula Edodes logs

\*Correspondence:

Qiulan Wu

zxylsg@sdau.edu.cn

Feng Zhang

zhangfeng@sdau.edu.cn

<sup>1</sup> School of Information Science & Engineering, Shandong Agricultural University, Tai'an 271018, China

<sup>2</sup> Shandong Century Intelligent Agriculture Technology Co., Ltd,

Ji'ning 273500, China

still relies on manual inspection. Manual inspection is not only high in labor cost and low in efficiency, but it also requires high professional quality of inspectors and usually can only detect contamination of Lentinula Edodes logs when it is more obvious. Timely, accurate detection of initial contamination in Lentinula Edodes logs is crucial for preventing further spread and generation of contamination, ensuring quality and yield improvements.

In recent years, with the development of deep learning theory and edge cloud computing, detection algorithms based on deep learning have been widely used due to their good generalization ability and cross-scenario capabilities [4–7], and related needs for mobile edge computing and cloud computing have gradually emerged [8–13]. Many academic institutions and



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

industry researchers have invested in the field of edge cloud computing, and these studies have promoted the development of edge cloud computing and computer vision technology [14-17]. In this context, the use of deep learning technology to process crop disease images has gradually become a research hotspot [18–20]. Zu et al. [21] used a deep learning method to identify contamination of Lentinula Edodes logs for the first time and proposed an improved ResNeXt-50 (32×4d) model for Lentinula Edodes log contamination identification. The method improves the model by fine-tuning the six fully connected layers in the ResNeXt-50 $(32 \times 4d)$  model to improve the accuracy of Lentinula Edodes logs contamination recognition, thereby breaking the situation of relying on manual detection with low efficiency and easy selection errors. However, this method has complex network structure and low detection efficiency, so it is not suitable for deployment in mobile devices or embedded devices, and cloud computing may provide it with more computing resources. To this end, Zu et al. [22] proposed Ghost-YoLoV4 for Lentinula Edodes logs contamination identification, which used a lightweight network, GhostNet, instead of a back-bone feature extraction network. This lightweight approach is well-suited for mobile edge computing, allowing real-time contamination identification on edge devices and alleviating the burden on cloud computing servers. Although scholars have achieved certain results in Lentinula Edodes logs contamination detection using deep learning techniques, the detection effect of existing studies is unsatisfactory in the early stage of contamination due to small contamination areas, requiring small object detection. The difficulty of small object detection has been an important problem faced by object detection algorithms, and many scholars have conducted indepth studies for this purpose [23-27]. However, no relevant literature on deep learning for small object Lentinula Edodes logs contamination detection has been found in previous studies. This study combined cloud computing and edge computing to process shiitake mushroom stick data, and designed an edge cloud computing framework for image enhancement and real-time detection on edge devices, as shown in Fig. 1. Edge devices with good network establish wireless network or data transmission connections with cloud servers. The cloud server receives and processes edge devices requests and performs corresponding algorithm calculations. When receiving multiple requests from multiple mobile edge devices, the cloud server performs parallel computing in the cloud. After the calculation is completed, real-time responses are provided to the edge devices. The overall data training and testing framework maintains the accuracy of edge computing resources in small object Lentinula Edodes logs contamination detection.

The YOLOv7 model proposed by Wang et al. [28] has faster speed and higher accuracy on the COCO dataset, making it ideal for mobile edge computing and resource-constrained environments. This performance improvement is of great significance for applications deployed in mobile edge computing systems and



Fig. 1 The edge cloud computing framework for image enhancement and real-time detection on edge devices

providing services in cloud computing environments. In this study, we improved the YOLOv7 algorithm and proposed a model (SRW-YOLO) applicable to small object Lentinula Edodes logs contamination detection. First, SPD-Conv was introduced in the MP module of the network to highlight small object contamination object features, which helps to perform object detection more effectively on resource-constrained edge devices. Then, RepVGG was used to re-parameterize the ELAN structure in the backbone network, reducing the pressure on mobile edge inference computing and cloud server resources, and further improving the detection of small object Lentinula Edodes logs contamination. Finally, the object location was regressed using the WIoU loss function paying more attention to ordinary-quality anchor boxes to improve overall detection performance of bacteriophage contamination condition.

## **Related work**

#### **Cloud computing**

As a paradigm of distributed computing, cloud computing decomposes large-scale data into sub-modules through a network center and then distributes it to a system composed of multiple servers for processing and analysis. The calculation results are finally fed back to the central node [29, 30]. Cloud computing technology combines the characteristics of distributed computing, parallel computing and grid computing to build massive computing clusters and storage clusters to provide users with scalable computing resources and storage space at low cost. Currently, many companies have enterprise-level cloud computing platforms, such as Amazon Cloud Computing, Alibaba Cloud Computing, Baidu Cloud Computing, etc. Compared with traditional application platforms, cloud computing platforms have the advantages of powerful computing power, unlimited storage capacity, and convenient and fast virtual services. However, for individuals and small companies, renting a cloud computing server involves additional costs. Therefore, in order to reduce cloud computing costs, a variety of new lightweight networks have been proposed for target detection. Common strategies include avoiding full connections in the network, reducing the number of channels and convolution kernel size, and optimizing down-sampling, weight pruning, weight discretization, model representation, and encoding [31, 32]. However, there is currently a lack of a small object Lentinula Edodes logs contamination detection model suitable for mobile edge computing and cloud computing environments.

#### **Edge computing**

The architecture design of edge computing originated from cloudlet [33] proposed by Carnegie Mellon University in 2009. In 2016, the Wayne State University team [34] formally defined edge computing and conducted in-depth research on its application scenarios. Subsequently, artificial intelligence solutions based on edge computing became a research hotspot. Hu et al. [35] proposed a method to build a face detection video surveillance system based on mobile edge computing (MEC). This method uses different detection algorithms at the edge and the cloud, and decides whether it needs to be sent to the cloud based on the confidence of edge detection. Jia [36] discussed the application prospects of edge computing models based on distributed data collection and processing in intelligent video detection. Wang et al. [37] proposed an online monitoring system architecture for transmission lines based on the ubiquitous Internet of Things by studying image recognition and mobile edge computing technology based on deep learning. However, it is difficult for existing edge detection models in agriculture to maintain a balance between accuracy and real-time performance.

#### **Crop disease detection**

In the field of agricultural production, ignoring the early signs of plant disease may lead to losses in food crops, which could eventually destroy the world's economy. Anh et al. [38] introduced a multi-leaf classification model based on a benchmark dataset, utilizing a pre-trained MobileNet CNN model. Their approach demonstrated efficiency in classification, achieving a reliable accuracy of 96.58%. In another study [39], a multi-label CNN was proposed for the classification of various plant diseases, employing transfer learning approaches such as DenseNet, Inception, Xception, ResNet, VGG, and MobileNet. The authors claimed the novelty of their research as the first to classify 28 classes of plant diseases using a multi-label CNN. The Ensemble Classifier was employed for plant disease classification in [40], evaluated with two datasets—PlantVillage and Taiwan Tomato Leaves. Pradeep et al. [41] presented the EfficientNet model, a convolutional neural network designed for multi-label and multi-class classification. The inclusion of a secret layer network in the CNN positively impacted the identification of plant diseases. However, the model exhibited underperformance when validated with benchmark datasets. In [42], an effective, loss-fused, resilient convolutional neural network (CNN) was proposed using the benchmark dataset PlantVillage, achieving a notable classification accuracy of 98.93%. Despite enhancing classification

accuracy, the model faced challenges in real-time image classification under varying environmental conditions.

## **Materials and methods**

#### Data acquisition

The dataset used in this study was sourced from the meticulously built database of the Smart Village Laboratory at Shandong Agricultural University, which was specifically developed for this research. The data was collected from a factory culture shed located in Shandong Province, China. A unique aspect of the data collection process was the installation of LED strip lights at regular intervals within the Lentinula Edodes log culture shed. At the same time, using cloud computing technology, Lentinula Edodes logs cultivators can monitor the lighting conditions in the Lentinula Edodes logs cultivation shed in real time, and remotely control the brightness and position of the LED strip lights to ensure a normal Lentinula Edodes logs cultivation environment to the greatest extent. The acquisition equipment used in the study was composed of two devices: A Canon EOS 600D camera and an IQOO8 cell phone. The image resolution captured by these devices ranged from 1900 to 4000 pixels in width and from 2200 to 4000 pixels in height. Based on the collected images of Lentinula Edodes logs, the logs were categorized into three distinct groups: Normal Lentinula Edodes logs, Aspergillus flavus-contaminated Lentinula Edodes logs, and Trichoderma viride-contaminated Lentinula Edodes logs, as distinctly delineated in Fig. 2. The dataset also included images of Lentinula Edodes logs that were contaminated by small objects. A comprehensive total of 3156 images were amassed, which comprised of 1734 images of normal Lentinula Edodes logs, 700 images of Aspergillus flavus-contaminated Lentinula Edodes logs, and 722 images of Trichoderma viride-contaminated Lentinula Edodes logs.

## Data pre-processing

In the realm of deep learning, it's essential for models to undergo rigorous training using copious amounts of data, a practice that is paramount to avert the issue of overfitting. The adequacy and comprehensiveness of the dataset employed assume a pivotal role in the endeavor to bolster the accuracy of the model being proposed. In a quest to widen the sample size, this study embraced the technique of data augmentation. The strategy of data enhancement incorporated a plethora of morphological operations such as rotation of angle, adjustment of saturation, alteration of exposure, flipping of images either up or down, and the application of random cropping, as distinctly delineated in Fig. 3. Through the implementation of these methodologies, an enlarged pool of samples could be generated, thereby enhancing the generalization



Fig. 2 Example images of Lentinula Edodes logs. **a**, **b** normal, **c**, **d** Aspergillus flavus-contaminated, and **e**, **f** Trichoderma viride-contaminated

capability and sturdiness of the model under consideration. The expanded set of data samples comprised 2988 images capturing normal Lentinula Edodes logs, 1912 images showcasing Aspergillus flavus-tainted Lentinula Edodes logs, and a further 1512 images depicting Trichoderma viride-contaminated Lentinula Edodes logs. The grand total of these images reached 6412, each of which was archived in the jpg format.

Concurrently, labeling was deployed as an image annotation tool. The types of annotations were meticulously categorized into three distinct groups: Normal, Aspergillus flavus, and Trichoderma viride. The label files, in turn, were preserved in the yolo format. To conclude the dataset's preparation, it was systematically partitioned into three distinct subsets: a training set, a validation set, and a test set. The proportions of these sets were calculated at a ratio of 8:1:1. Specifically, the training set encompassed a total of 5130 images, while the validation and test sets each contained 641 images.

# SRW-YOLO model construction

In this study, an SRW-YOLO network model suitable for mobile edge computing and cloud computing environments was designed for the problem of small object Lentinula Edodes log contamination detection, as shown in Fig. 4. Firstly, the MP module was improved



Fig. 3 Renderings of data enhancements



Fig. 4 The network structure of SRW-YOLO

by SPD-Conv to enhance the learning of small object features in the Lentinula Edodes log images and avoid the loss of fine-grained information. Secondly, RepVGG was introduced into the ELAN structure, and the structure was reparameterized to decouple the multi-branch structure and inference ordinary structure during model training, which further improved the efficiency and accuracy of inference for small object contaminated regions. Finally, the original boundary regression loss function was replaced with the WIoU loss function, which weakens the influence of high-quality anchor boxes and lowquality sample features and focuses on ordinary-quality anchor boxes, making the model output results more accurate. During the Lentinula Edodes logs cultivation phase, the mobile device collects images and transmits them to the cloud data processing center to generate the final mushroom stick detection image.

#### MP module based on SPD-Conv

YOLOv7 uses an MP structure to downsample the input. Downsampling is usually implemented using convolutional layers, pooling layers, or convolution with a step size greater than 1 to gradually reduce the spatial size of the input tensor and, thus, increase the perceptual field of the network. However, in the process of downsampling, it was easy to cause the resolution of Lentinula Edodes log images to decrease too fast, which would lead to a loss of information about the location and size of Lentinula Edodes log contamination, thus reducing the accuracy of detection. Therefore, to solve this problem, the MP module was improved by introducing SPD-Conv [43]. SPD-Conv consists of a space-to-depth (SPD) layer and a non-stride convolutional layer. The SPD layer slices an intermediate feature map  $X(S * S * C_1)$  into

a series of sub-maps  $f_{(x,y)}$  by downsampling the feature maps inside the convolutional neural network and the entire network.

$$f_{0,0} = X[0:S:scale, 0:S:scale], f_{1,0} = X[1:S:scale, 0:S:scale], ..., f_{scale,0} = X[scale - 1:S;scale, 0:S;scale]$$
(1)

$$f_{0,1} = X[0:S:scale, 1:S:scale], f 1, 1, \dots, f_{scale-1,1} = X[scale - 1:S:scale, 1:S:scale];$$
(2)

$$f_{0,scale-1} = X[0:S:scale,scale-1:S:scale], f_{1,sclae-1}, \dots,$$
(3)  
$$f_{clae-1,clae-1} = X[scale-1:S:scale,scale-1:S:scale]$$

Given any (original) feature map X,  $f_{x,y}$  which consists of the feature map X(i,j) is composed of the region where i + x and j + y are divisible by the scale.

:

Thus, each subsample is mapped down by a scale factor X. Finally, the sub-feature maps are stitched along the channel dimension to obtain a feature map X. Adding a non-stride convolution after the SPD feature transformation preserves all the discriminative feature information as much as possible, and the SPD-Conv structure is shown in Fig. 5

A total of five MP modules were constructed in the original model for the backbone network and the feature fusion network. Since there is a convolution of step 2 in the second branch of the MP module, this study used SPD-Conv to replace the convolution of step 2 in the MP of the feature fusion network, as shown in Fig. 6. Considering the large input image pixels, the number of parameters, and the computational efficiency of the model, all convolutions with step size 2 in the network were not replaced in this study.





## RepVGG-based efficient aggregation network module

The efficient aggregation network module proposed in the original model is mainly divided into an ELAN [44] structure and an E-ELAN structure. The ELAN uses a special jump connection structure to control the longest gradient path, and the deeper network can learn and converge efficiently. The E-ELAN is an expansion, channel rearrangement, and transition layer architecture without destroying the original gradient path of the ELAN or changing the merging bases to enhance the learning ability of the network. However, the efficient aggregation network module may assign some important information to different groups and affect model performance. In addition, this module uses fewer convolutional layers, which can be challenging when dealing with the task of detecting small object contaminated areas of Lentinula Edodes logs. Therefore, in this study, the efficient aggregation network module was improved using RepVGG [45]. RepVGG decouples the training multi-branch topology and inference singleway structure using structural reparameterization, as shown in Fig. 7. The structural reparameterization is mainly divided into two steps: the first step is mainly to fuse Conv2d and BN (Batch Normalization) as well as



Fig. 7 Sketch of RepVGG architecture

to convert the branches with only BN into one Conv2d, and the second step fuses the  $3 \times 3$  convolutional layers on each branch into one convolutional layer; this structure can increase the nonlinearities of the model while reducing the computation during inference. At the same time, the reparameterization reduces the computation and memory usage, which helps to handle small object contamination detection tasks. The specific improvement in this study is to introduce the RepVGG module in all ELAN structures in the backbone network.

#### Boundary regression loss function

In an object detection task, the bounding box regression loss function is critical to the performance of the model used. The role of the bounding box regression loss function is to measure the difference between the model-predicted bounding box and the true bounding box, which affects the detection effectiveness of the model. Low-quality samples, such as small object contamination, exist in the dataset of Lentinula Edodes logs, and the geometric factors, such as distance and aspect ratio, taken into account by the traditional bounding box loss function will aggravate the penalty of low-quality examples, which may reduce the generalization performance of the model. Therefore, in this study, WIoUv3 [46] was used as the boundary regression loss function for the model. WIoUv3 proposes outliers instead of IoU to evaluate the quality of anchor boxes and provide a sensible gradient gain allocation strategy. This strategy reduces the competitiveness of high-quality anchor boxes while minimizing harmful gradients generated by low-quality examples, which contributes to the speed of model convergence and the accuracy of inference, thus improving the overall performance of model detection. This is achieved by assigning outlier  $\beta$  an appropriate gradient gain depending on its size, with smaller or larger outliers  $\beta$  being assigned smaller gradient gains that are more focused on ordinary-quality anchor boxes, with outlier  $\beta$  being defined as follows:

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \tag{4}$$

where  $L_{IoU}^*$  is the monotonic focus factor and  $\overline{L_{IoU}}$  is the sliding average of the momentum of m.

Distance attention was also constructed based on the distance metric, and a WIoUv1 with two layers of attention mechanisms was constructed as follows:

$$L_{W \cdot IoUv1} = R_{WIoU}L_{IoU}$$
$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{\left(W_g^2 + H_g^2\right)^*}\right)$$
(5)

where  $L_{IolI}$  is the degree of overlap between the prediction box and the real box; (x, y) is the center coordinate of the predicted box;  $(x_{gt}, y_{gt})$  is the center coordinate of the real box; and  $W_g$  and  $H_g$  are the length and width of the real box and the predicted box, respectively.

At this point, applying the outlier degree to  $L_{WIoUv1}$  obtains  $L_{WIoUv3}$ :

$$L_{W \cdot IoUv3} = r L_{W \cdot IoUv1}, r = \frac{\beta}{\delta \alpha^{\beta - \delta}}$$
(6)

where  $L_{WIoUv1}$  is the attention-based boundary loss, and  $\delta$  with  $\alpha$  is the hyperparameter.

When the outlier degree of the anchor box satisfies  $\beta = C$  (*C* is a constant value), the anchor box will obtain the highest gradient gain. Since  $\overline{L_{Ioll}}$  is dynamic, the quality classification criteria of the anchor boxes are also dynamic, which allows WIoUv3 to construct a gradient gain allocation strategy that best fits the current situation at each moment.

# Model training and evaluation *Model training*

In this study, SRW-YOLO used the default hyperparameters of YOLOv7. The learning rate was set to 0.01, SGD was selected for hyperparameter optimization, and the learning rate momentum was set to 0.937. Meanwhile, a pre-trained model was used for training assistance, which could help the model achieve better initial performance. The configuration of the experimental environment in this study is shown in Table 1.

#### Model evaluation

To verify the performance of Lentinula Edodes log contamination detection, Precision, Recall, mAP, and FPS were used for evaluation in this study. The calculation equations are as follows.

$$Precision = \frac{TP}{TP + FP}$$
(7)

$$Recall = \frac{TP}{TP + FN}$$
(8)

$$AP = \int_{0}^{1} P_{(r)} dr \tag{9}$$

$$mAP = \frac{\sum_{i=1}^{C} AP_i}{C} \tag{10}$$

where TP indicates that the object is a certain type of Lentinula Edodes logs and the network model detection also indicates a certain type of Lentinula Edodes logs. FPindicates that the object is not a type of Lentinula Edodes logs, but the network model detects a type of Lentinula Edodes logs. FN indicates that the object is a certain type of Lentinula Edodes logs, but the network model detection indicates it is not a certain type of Lentinula Edodes logs. AP is the area enclosed by Precision and Recall on the curve. mAP is the average of all categorized AP values; when IoU is set to 0.5, it is mAP@0.5, and

Table 1 Experimental environment configuration

Experimental Environment	Details		
Programming language	Python 3.8.10		
Operating system	Linux		
Deep learning framework	Pytorch 1.11.0		
CUDA Version	11.3		
GPU	RTX 3090		
CPU	Intel(R) Xeon(R) Platinum 8358P @ 2.60 GHz		

mAP@0.5:0.9 means that the IoU threshold is between 0.5 and 0.9.

### **Results and analysis**

#### Model visualization analysis

After the training of the model, the feature extraction results of the first convolutional layer, the backbone module, and the last convolutional layer were visualized and analyzed in this study using class activation mapping (CAM) [47]; the information of interest to the network model can be seen from the visualized feature map. This study randomly selects an image of small object Lentinula Edodes logs contamination from the training set to visualize its characteristics. The red box area is the area contaminated by the Lentinula Edodes log. The visual analysis results are shown in Fig. 8. The figure shows the feature visualization images of the three improvement strategies of SPD-Conv, RepVGG, and WIoUv3 regression loss function and the three stages of the SRW-YOLO comprehensive improvement model. The three stages are the first convolutional layer, the feature extraction backbone layer and the last convolutional layer. The darker the red part, the more the model pays attention to this part of the image; this is followed by the yellow part. The bluer the heat map is, the more the model considers this part as redundant information.

As can be seen from the first layer convolutional feature map, the three improvement strategies mainly focus on the low-level features of the Lentinula Edodes logs, such as edges and textures. The feature map of the feature extraction backbone convolu-tional layer shows more advanced feature attention, and the focus is more localized. SRW-YOLO accurately locates small object contaminated areas, and the three im-provement strategies all focus on the contaminated areas of the bacterial sticks rela-tively accurately. However, the three improvement strategies all focus on more back-ground redundant information to varying degrees. It can be seen from the feature map of the last convolutional layer that the features extracted by different im-provement strategies are more abstract and refined, revealing how the model focuses on discriminative features in the final stage. The above improvement strategies ulti-mately focused on two contaminated areas. However, SPD-Conv paid too much attention to the two areas and considered more redundant pixels; Rep-Conv and WIoU3 also paid too much attention to the right areas. The feature extraction ability of the contaminated area below is weak; while SRW-YOLO focuses on key pixel areas and is more accurate.

YOLOv7 +SPD-Conv YOLOv7+RepVGG Original image YOLOv7+WIoUv3 SRW-YOLO Stage0\_Conv\_feature Backbone Conv feature Last Conv feature

Fig. 8 Visualization of the feature map

It can be observed from the feature maps from the backbone module to the last layer that the algorithm model proposed in this study plays a good role in reinforcing the feature maps, suppressing unnecessary features, and enabling better extraction of small object contamination feature information from the images.

#### Analysis of experimental results

To verify the positive impact of the improvement strategy proposed in this study on the network, ablation experiments were conducted on the Lentinula Edodes log da-taset in this paper. Five sets of experiments were conducted, and different improve-ment modules were added for comparison with YOLOv7, with Precision, Recall, mAP@0.5, and FPS being used as the measures. The results of the ablation experiments are shown in Table 2.

As can be seen from the above table, Experiment 1 provides the detection results of the original YOLOv7 network. In Experiment 2, Precision, Recall, and mAP@0.5 improve by 2.33%, 1.93% and 1.97%, respectively. This indicates that during the model downsampling process, SPD-Conv effectively alleviates the impact of the rapid decrease in resolution of the mushroom stick image, thereby strengthening the learning of effective feature representation of the Lentinula Edodes logs image and avoiding the loss of fine-grained information, and helping improve the accuracy of mobile edge device detection. In Experiment 3, Precision, Recall and mAP@0.5 improve by 1.77%, 0.51% and 1.56%, respectively. This indicates that after using structural re-parameterization to improve the efficient aggregation network module in the model, RepVGG can reduce the computational load and memory usage of model inference while improving the efficiency and accuracy of inference on Lentinula Edodes logs contaminated areas, and reducing the pressure of mobile edge computing inference. and the burden on cloud servers. In Experiment 4, YOLOv7 improves Precision by 1.56%, Recall by 1.05% and mAP@0.5 by 1.58% over the YOLOv7 algorithm after using WIoUv3 as the boundary regression loss function of the network. This indicates that when YOLOv7 adopts WIoUv3, through a wise gradient gain allocation strategy, the model is more focused on ordinary-quality Lentinula Edodes logs detection anchor boxes, making the model output results more accurate. In Experiment 5, the Precision improves by 3.15% and mAP@0.5 improves by 2.64% over the YOLOv7 algorithm. This shows that when the SPD-Conv module and RepVGG module are introduced into the original network, the network inference efficiency is improved while avoiding the loss of location and size information of bacteriophage contamination, which in turn improves the accuracy of detection. Experiment 6 integrated the above improved methods, and it can be clearly seen that the detection effect is the best. Precision reaches 97.63%, which is 4.62% better than YOLOv7; Recall reaches 96.43%, 3.63% higher than YOLOv7; and mAP@0.5 reaches 98.62%, which is 2.31% better than YOLOv7. At the same time, it also maintains good real-time detection, which can meet the requirements of small object Lentinula Edodes logs contamination detection in mobile edge computing and cloud computing environments.

The ablation experiments can only verify the effectiveness of the improved strategy in this study relative to the original algorithm, but whether it can reach the leading level in different models still needs further proof. Therefore, under the same experimental conditions, a series of comparative experiments were conducted in this study to compare the performance of the improved method with the current mainstream one-stage object detection method using the Lentinula Edodes log dataset.

A comparison of the training results of different models is shown in Fig. 9. From the figure, it can be seen that the value of mAP@0.5 of the improved algorithm in this study is significantly higher than the other three models.

Figure 10 presents a comparison of the regression loss curves for different models with training time. After 40 iterations, the loss curves of different models gradually and steadily converge. It can be seen that YOLOv6m has poor loss convergence in this dataset and YOLOv51 has an overfitting problem after 100 training iterations. YOLOv51 and YOLOv6m are much less effective than YOLOv7 in terms of regression loss. The model proposed in this study shows a better drop rate and convergence ability than YOLOv7, thus proving that the improvement

Number	Model	Precision	Recall	mAP@0.5	FPS
1	YOLOv7	93.01%	92.80%	96.31%	56
2	YOLOv7 + SPD-Conv	95.34%	94.73%	96.88%	52
3	YOLOv7 + RepVGG	94.78%	93.31%	96.47%	53
4	YOLOv7+WIoUv3	94.57%	93.85%	96.49%	64
5	YOLOv7 + SPD-Conv + RepVGG	96.16%	96.05%	97.55%	51
6	YOLOv7 + SPD-Conv + RepVGG + WIoUv3	97.63%	96.43%	98.62%	62

Table 2 Results of ablation experiments



Fig. 9 Comparison of training box\_loss curves of different models



Fig. 10 Comparison of training box\_loss curves of different models

of the boundary regression loss function improves the convergence ability of the network.

Table 3 lists the comparison results of the evaluation metrics of different models. ResNeXt-50 ( $32 \times 4d$ ), MobilenetV3-YOLOv4 and Ghost-YOLOv4 are Zu's research methods. Compared with the mainstream YOLO series algorithms, the performance of these methods in small object Lentinula Edodes logs contamination detection needs to be improved. Compared to other models, although the detection speed of the SRW-YOLO

Model	mAP@0.5	Recall	mAP@0.5:0.9	FPS
ResNeXt-50 (32×4d)	93.32%	91.12%	87.13%	17
MobilenetV3-YOLOv4	92.16%	89.11%	84.77%	30
Ghost-YOLOv4	93.04%	91.15%	86.56%	36
YOLOv5I	94.79%	92.68%	87.51%	67
YOLOv6m	95.82%	93.29%	91.32%	48
YOLOv7	96.31%	92.80%	91.66%	56
SRW-YOLO	98.62%	96.43%	92.39%	62

class	model					
	YOLOv7			SRW-YOLO		
	Precision	Recall	mAP@0.5	Precision	Recall	mAP@0.5
Normal	95.62%	94.44%	99.53%	97.45%	96.81%	99.71%
Aspergillus flavus	87.71%	93.32%	95.37%	95.83%	98.65%	97.73%
Trichoderma viride	95.75%	86.16%	93.22%	99.73%	91.24%	97.31%

Table 4 Comparison of evaluation indicators of different classes

model proposed in this study is not the highest, it is much better than other models in the evaluation metrics of mAP@0.5, Recall, and mAP@0.5:0.9, This allows the model to maintain a good balance between detection accuracy and real-time performance.

At the same time, in order to further demonstrate the superiority of the SRW-YOLO model improvement strategy, Table 4 lists the comparison results of the evaluation indicators of YOLOv7 and SRW-YOLO in three classes of Lentinula Edodes logs contamination detection. Compared with YOLOv7, SRW-YOLO has improved to varying

degrees in the evaluation indicators of Precision, Recall and mAP@0.5. Among them, the original model has the worst effect in detecting Aspergillus flavus contaminated Lentinula Edodes logs, but the SRW-YOLO model improves Precision, Recall and mAP@0.5 by 8.12%, 5.33% and 2.36% respectively compared with YOLOv7. This shows that the SRW-YOLO model proposed in this article has more advantages in actual detection and can accurately detect different classes of Lentinula Edodes logs.

For a more intuitive understanding of the performance of the models, Fig. 11. Shows the detection results of



Original picture



YOLOv51



Fig. 11 Detection results of different models



YOLOv6m



SRW-YOLO

the four models for a randomly selected image in the test set, with the red box in the figure showing the area contaminated by Trichoderma viride. Although all four models are able to detect the type of Lentinula Edodes logs, YOLOv5l, YOLOv6m, and YOLOv7 have lower confidence in the detection of the object and poorer detection results. In contrast, SRW-YOLO has obvious superiority with 95% object confidence and accurately detects small object contaminated areas.

In summary, the Lentinula Edodes log contamination detection model proposed in this study has strong generalization ability and robustness. During the Lentinula Edodes logs cultivation stage, this model can better locate areas with small contamination objects in Lentinula Edodes logs and accurately detect the type of Lentinula Edodes log contamination.

# Conclusion

In this study a model for small object Lentinula Edodes logs contamination detection (SRW-YOLO) suitable for mobile edge computing and cloud computing environments was constructed based on YOLOv7. SPD-Conv was introduced in the MP module of the feature fusion network to improve the learning ability of the model for small object contamination location and semantic information of Lentinula Edodes logs, which helps to enhance the accuracy of mobile device detection with limited resources; the ELAN structure in the backbone network was reparameterized and the RepVGG architecture was used to realize the decoupling of training and inference to efficiently and accurately detect the types of Lentinula Edodes logs and reduce mobile edge computing inference pressure and cloud server burden; the WIoU loss function was set as the boundary regression loss of the network function to reduce the competitiveness of highquality anchor boxes while minimizing harmful gradients generated by low-quality samples to improve the overall performance of Lentinula Edodes logs contamination condition detection. Compared to the current mainstream one-stage object detection model, the experimental results showed that the detection of small object Lentinula Edodes log contamination by SRW-YOLO is significantly better. In summary, SRW-YOLO provides an efficient, accurate and practical small object contamination detection method that can be deployed to Android mobile devices or embedded devices. In addition, companies or individuals using the network proposed in this study can reduce the performance of cloud computing servers and reduce the cost of renting cloud computing servers.

However, there are still some areas for improvement. The current Lentinula Edodes logs dataset has a relatively simple background, and the model may not perform well when the background is more complex or the data collection environment is dimmer. Therefore, in subsequent work, the dataset will be further improved and the proposed Lentinula Edodes logs contamination detection method will be optimized.

#### Abbreviations

YOLO	You only look once
FPS	Frames per second
IOU	Intersection over union
AP	Average precision
mAP	Mean average precision
ELAN	Efficient layer aggregation networks
E-ELAN	Extended efficient layer aggregation networks

#### Acknowledgements

We would like to thank Shandong Century Intelligent Agricultural Technology Co., Ltd. for providing the experimental site and data for this study.

#### Authors' contributions

Conceptualization, X.C. and Q.W.; methodology, X.C.; software, C.C.; validation, F.Z., Q.W., S.S. and X.S.; formal analysis, C.C.; investigation, X.C.; resources, Q.W.; data curation, F.Z., S.S. and X.S.; writing—original draft preparation, X.C.; writing—review and editing, F.Z., S.S., and X.S.; visualization, X.C. and C.C.; supervision, Q.W.; project administration, Q.W.; funding acquisition, Q.W. All authors have reviewed and agreed to the published version of the manuscript.

#### Funding

This research was funded by the Major Scientific and Technological Innovation Project of Shandong Province (Grant No. 2022CXGC010609), and the Talent Project of Zibo City.

### Declarations

#### **Competing interests**

The authors declare no competing interests.

#### Received: 14 October 2023 Accepted: 21 December 2023 Published online: 10 January 2024

#### References

- Cao Z, Wang S, Zheng S et al (2022) Identification of Paecilomyces variotii and its interaction with Lentinula Edodes mycelium. North Horticulture 509(14):116–125
- 2. Wang Y, Liu Z, Feng Y et al (2021) Study on the infection process of Trichoderma in the production of Lentinus Edodes. Seed 40(6):131–141
- Yao Q, Gong Z, Si H et al (2020) Study on the formulation of culture substrate of lentinus Edodes with resistance to hybrid bacteria. Chinese J Edible Fungi 39(10):56–58
- Kim JH, Kim BG, Roy PP et al (2019) Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE access 7:41273–41285
- Jeppesen JH, Jacobsen RH, Inceoglu F et al (2019) A cloud detection algorithm for satellite imagery based on deep learning. Remote Sens Environ 229:247–259
- Li Z, Xu X, Hang T, et al (2022) A knowledge-driven anomaly detection framework for social production system. IEEE Transactions on Computational Social Systems
- Ling Z, Yu K, Zhang Y et al (2022) Causal learner: A toolbox for causal structure and markov blanket learning. Pattern Recogn Lett 163:92–95
- Yang Y, Ding S, Liu Y et al (2022) Fast wireless sensor for anomaly detection based on data stream in an edge-computing-enabled smart greenhouse. Digital Commun Netw 8(4):498–507

- 9. Xu X, Tang S, Qi L et al (2023) CNN Partitioning and Offloading for Vehicular Edge Networks in Web3. IEEE Commun Mag 61(8):36–42
- 10. Xu X, Li H, Li Z et al (2022) Safe: Synergic data filtering for federated learning in cloud-edge computing. IEEE Trans Industr Inf 19(2):1655–1665
- Paranjothi A, Atiquzzaman M (2022) A statistical approach for enhancing security in VANETs with efficient rogue node detection using fog computing. Digit Commun Netw 8(5):814–824
- Gong W, Zhang W, Bilal M et al (2021) Efficient web APIs recommendation with privacy-preservation for mobile app development in industry 4.0. IEEE Trans Industr Inform 18(9):6379–6387
- Wenwen G, Chengming Z, Qing C, et al. A Trust Model for Secure and Reliable Cloud Service Systems[C]//2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS). IEEE, 2018; 95–99
- 14. Yan K, Zhou X (2022) Chiller faults detection and diagnosis with sensor network and adaptive 1D CNN. Digit Commun Netw 8(4):531–539
- Kumar A, Abhishek K, Ghalib MR et al (2022) Intrusion detection and prevention system for an IoT environment. Digit Commun Netw 8(4):540–551
- Xu X, Gu J, Yan H et al (2022) Reputation-aware supplier assessment for blockchain-enabled supply chain in industry 4.0. IEEE Trans Industr Inform 19(4):5485–5494
- 17. Weinger B, Kim J, Sim A et al (2022) Enhancing IoT anomaly detection performance for federated learning. Digit Commun Netw 8(3):314–323
- Liu J, Wang X (2021) Plant diseases and pests detection based on deep learning: a review. Plant Methods 17:1–18
- Meng-min Si, Ming-hui D, Ye H (2019) Using deep learning for soybean pest and disease classification in farmland. J Northeast Agric Univ 26(01):64–72
- Elhassouny A, Smarandache F. Smart mobile application to recognize tomato leaf diseases using convolutional neural networks. Collected Papers, 2019: 431
- 21. Zu D, Zhang F, Wu Q et al (2022) Disease identification of Lentinus Edodes sticks based on deep learning model. Complexity 2022:1–9
- Zu D, Zhang F, Wu Q et al (2022) Sundry bacteria contamination identification of Lentinula Edodes logs based on deep learning model. Agronomy 12(9):2121
- Han J, Ding J, Xue N, et al. (2021) Redet: A rotation-equivariant detector for aerial object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 2786–2795
- Zand M, Etemad A, Greenspan M (2021) Oriented bounding boxes for small and freely rotated objects. IEEE Trans Geosci Remote Sens 60:1–15
- Yu D, Xu Q, Guo H et al (2022) Anchor-free arbitrary-oriented object detector using box boundary-aware vectors. IEEE J Sel Top Appl Earth Obs Remote Sens 15:2535–2545
- Zhu X, Lyu S, Wang X, et al. (2021) TPH-YOLOV5: Improved YOLOV5 based on transformer prediction head for object detection on drone-captured scenarios. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 2778–2788
- Benjumea A, Teeti I, Cuzzolin F, et al. (2021) YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles. arXiv preprint arXiv: 2112.11798
- Wang C, Bochkovskiy A, Liao H. (2022) YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696
- 29. Xue C, Lin C, Hu J (2019) Scalability analysis of request scheduling in cloud computing. Tsinghua Sci Technol 24(3):249–261
- Shen D, Luo J, Dong F, Zhang J (2019) Virtco: joint coflow scheduling and virtual machine placement in cloud data centers. Tsinghua Sci Technol 24(5):630–644
- Han S, Mao H, Dally W J. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv preprint arXiv:1510.00149, 2015
- Han S, Pool J, Tran J, et al. Learning both weights and connections for efficient neural network. Advances in neural information processing systems, 2015, 28
- 33. Satyanarayanan M, Bahl P, Caceres R et al (2009) The case for vm-based cloudlets in mobile computing. IEEE Pervasive Comput 8(4):14–23

- Shi W, Cao J, Zhang Q et al (2016) Edge computing: Vision and challenges. IEEE Internet Things J 3(5):637–646
- Hu H, Shan H, Wang C et al (2020) Video surveillance on mobile edge networks—a reinforcement-learning-based approach. IEEE Internet Things J 7(6):4746–4760
- 36. Xiaoqian J, Gang C, Baibing L (2020) Application of edge computing in video surveillance. Comput Eng Appl 56(17):86–92
- Wang Y, Liu H, Li L (2019) Application of image recognition technology based on edge intelligence analysis in transmission line online monitoring. Int Conf Electr Eng Inform Commun 17(7):35–40
- Anh P T, Duc H T M. A benchmark of deep learning models for multileaf diseases for edge devices[C]//2021 International Conference on Advanced Technologies for Communications (ATC). IEEE, 2021: 318–323
- Kabir M M, Ohi A Q, Mridha M F. A multi-plant disease diagnosis method using convolutional neural network. Com Vision Machine Learning Agri, 2021: 99–111
- Astani M, Hasheminejad M, Vaghefi M (2022) A diverse ensemble classifier for tomato disease recognition. Comput Electron Agric 198:107054
- Prodeep A R, Hoque A S M M, Kabir M M, et al. Plant disease identification from leaf images using deep CNN'S efficientnet[C]//2022 International Conference on Decision Aid Sciences and Applications (DASA). IEEE, 2022: 523–527
- 42. Gokulnath BV (2021) Identifying and classifying plant disease using resilient LF-CNN. Eco Inform 63:101283
- Sunkara R, Luo T (2023) No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects. Lecture Notes in Computer 13715:443–459
- 44. Wang C, Liao H, Yeh I. (2022) Designing network design strategies through gradient path analysis. arXiv preprint arXiv: 2211.04800
- Ding X, Zhang X, Ma N, et al. (2021) Repvgg: Making vgg-style convnets great again. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 13733–13742
- Tong Z, Chen Y, Xu Z, et al. (2023) Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. arXiv preprint arXiv: 2301.10051
- Selvaraju R, Cogswell M, Das A, et al. (2017) Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision, pp 618–626

# **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- ► Rigorous peer review
- Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com