RESEARCH ON GRADING METHOD OF PEPPER PLUG SEEDLINGS BASED ON MACHINE VISION
/
基于机器视觉的辣椒穴盘苗分级方法研究

Fengwei YUAN[1)], Shuaiyin CHEN[1)], Zhang XIAO[*1)], Erjie SUN[1)], Guoning MA[1)], Gengzhen REN[1)], Zhenlong LI[1)] Zhenhong ZOU[2)], Xiangjiang WANG[1)]

[1)] College of Mechanical Engineering, University of South China, Hengyang/China.

[2)] Hengyang Vegetable Seeds Co. Hengyang/ China.

Tel: +8615200596268; E-mail: 2013000885@usc.edu.cn

DOI: https://doi.org/10.35633/inmateh-76-105

Keywords: *Machine vision, Plug seedlings, Image segmentation, Grading methods*

ABSTRACT

This study proposes a grading method using combined target detection and segmentation models to enhance grading accuracy for pepper plug seedlings. We constructed separate datasets for target detection and image segmentation of pepper plug seedlings, then trained various detection models including EfficientDet, Faster R-CNN, SSD, and YOLO-series architectures. After comprehensive evaluation, YOLOv5 achieved optimal performance with 99.5% mAP, a compact 3.8 MB model size, and rapid 5.33ms inference time per image. To avoid the impact of the culture medium and adjacent pepper seedlings on the image, we developed an improved U-Net model incorporating an Efficient Channel Attention (ECA) mechanism, enhancing segmentation accuracy by 1.29% to 93.01% while reducing processing time by 69.7% (33.32 ms). Subsequent feature extraction analysed lateral area, height, stem thickness, hypocotyl length, and divergence degree from segmented images. Using these extracted features, grading models were trained employing support vector machines, k-nearest neighbours, and random forests. The random forest model achieved a grading accuracy of 99.33%, validating that this method meets the accuracy requirements for pepper seedling grading.

摘要

为了提高辣椒穴盘苗分级检测精度，本研究提出了一种基于目标检测和分割模型相结合的辣椒穴盘苗分级方法。首先，构建辣椒穴盘苗目标检测和图像分割两个数据集；其次，训练了 Efficientdet，Faster-Rcnn，SSD 以及 YOLO 系列目标检测模型，经过综合对比 YOLOv5 检测效果最好，平均精度均值（mean average precision，mAP）达到 99.5%，模型大小 3.8 MB，单张图像检测时间 5.53ms；然后，为了避免培养基质以及相邻辣椒苗对图像的影响，使用 U-Net 算法作为本文的图像处理算法的基础模型，并引入通道注意力机制 ECA（Efficient Channel Attention Network）进行改进，与 U-Net 相比改进后模型分割准确率提高 1.29%，达到 93.01%，检测时间缩短 69.7%，为 33.32ms；最后基于图像分割算法处理后的图像，提取了辣椒穴盘苗面积，高度，茎粗，下胚轴长度和发散程度等参数，并利用提取的特征参数完成了支持向量机，K 最邻近算法，和随机森林算法的分级模型训练，其中随机森林模型分级准确率为 99.33%，验证了该方法能够满足辣椒穴盘苗分级精度要求。

INTRODUCTION

During the cultivation process of plug seedlings, factors such as seed quality, temperature, sunlight, and soil conditions can cause large differences in seedling morphology, which in turn affect subsequent transplanting and sales. Currently, the grading of plug seedlings primarily relies on manual labor, which is characterized by inconsistent grading standards and relatively high costs. Therefore, achieving automated grading for plug seedlings is a crucial task.

Subo Tian et al., (2017), expressed the cumulative maximum of 0 grey value pixels within a certain range of rows as the diameter of the seedling to achieve grading of melon seedlings. *Li Mingyong et al., (2021),* extracted two characteristics of leaf area and greenness per unit leaf area for grading of pepper plug seedling. *Jiang Huanyu et al., (2009),* determined whether the transplanting criteria were met by extracting the leaf area and perimeter of plug seedlings, with an identification accuracy of 98%. *Tong Junhua et al.,* (*2018*), used the centroid method and watershed algorithm to extract leaf area and the number of seedling leaves in a string red tray image for quality evaluation. *Fu Wei et al.,* (*2022*), employed threshold segmentation, erosion, dilation, and area division to extract the leaf area of lettuce seedlings in terms of pixel values, which was subsequently used to assess seedling quality.

Although the aforementioned literature obtained different methods for grading plug seedlings by obtaining phenotypic parameters, the algorithms such as threshold segmentation used are not suitable for changing environmental conditions and have low robustness.

In recent years, deep learning algorithms continue to develop, and convolutional neural networks are beginning to be widely used in grading detection, recognition and segmentation of plug seedlings. *Zhang Xiuhua et al., (2022)*, used spatial pyramid pooling and introduced the SAM spatial attention mechanism to improve the YOLOv3-Tiny model, with an average accuracy of 98.22% for tomato seedling robustness detection. *Xu Shengyong et al., (2023),* collected height and leaf area data of watermelon seedlings in the tray for the first three days, and used the LSTM neural network to predict the height and leaf area on the sixth day for grading, achieving an accuracy rate of 84%. *Li Yatao et al., (2024),* utilized a multi-layer perceptron neural network for image segmentation and calculated the area of plug seedlings by connected component centroid coordinates for grading, achieving an accuracy rate of 96.9%. The aforementioned studies indicate that there are various methods for grading plug seedlings; however, most of these methods focus primarily on seedlings with fewer than five leaves, which do not meet the requirements for sale.

To address this issue, this paper proposes a phenotypic parameter acquisition method for pepper plug seedlings, which combines target detection and segmentation models. The target detection model processes the entire row of pepper seedlings into individual states, thereby reducing the complexity and computational load of subsequent image processing. The segmentation algorithm is then used to segment the images processed by the detection model, avoiding the influence of the background and adjacent pepper seedlings. Feature extraction methods are utilized to obtain five parameters of the pepper plug seedlings, including lateral area, height, stem thickness, hypocotyl length, and divergence degree. By comparing the experimental results of various grading models, the random forest model, which has the highest accuracy, is selected to achieve accurate grading of pepper plug seedlings.

MATERIALS AND METHODS

Data set construction

A size of 5×10 was selected for the plug seedlings in this study. The images were captured using an industrial camera MI-230U150C, with a resolution of 1920px by 1200px. To comprehensively acquire the phenotypic parameters of the pepper seedlings, the camera is positioned at an inclined angle above the seedlings to avoid interference from the back row of seedlings. The target detection dataset comprised a total of 200 images of pepper seedling trays, while the image segmentation dataset, derived from the individual seedling trays processed by the target detection algorithm, contained 1000 images. Both datasets were proportioned into training and validation sets with a 7:3 ratio. The processing flow is shown in Fig. 1.

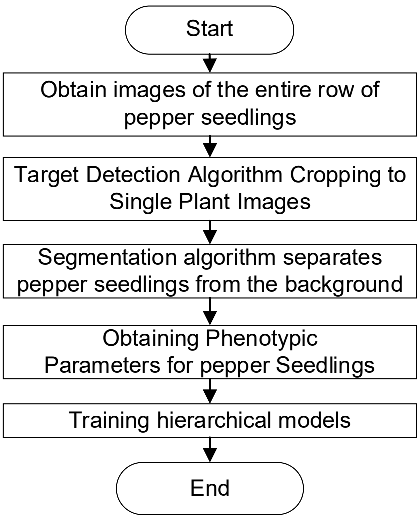

Fig. 1 - Flow chart of hierarchical processing

Algorithm for detection of pepper plug seedlings

In large scale nursery enterprises, most of them use plug seedlings, and the research object of this paper is to meet the transplanting conditions of pepper plug seedlings. The seedlings in the current growth cycle exhibit dense foliage, resulting in significant mutual shading between plants. Processing an entire row of seedlings into individual plants can effectively reduce the complexity of subsequent image processing.

After comprehensive comparison, this study adopts the YOLOv5 model for the detection of pepper plug seedlings. The backbone network of YOLOv5 employs a hopping fusion technique based on the Darknet architecture, which effectively enhances the information transfer efficiency of the network, allowing the model to learn rich feature information in a shorter period of time, thereby improving overall detection performance.

Segmentation Algorithm for Pepper plug Seedlings

The single target image is segmented based on the bounding box coordinates of the pepper seedling detected by YOLOv5. Since robust seedlings occupy a significant amount of space, the leaves of other seedlings may appear in the individual image of the pepper seedlings. Traditional segmentation algorithms that rely on features such as color, shape, and contour are incapable of distinguishing the leaves of other seedlings. In contrast, the improved U-Net segmentation algorithm, which is trained on convolutional neural networks, automatically learns the feature representation within the images, leading to superior segmentation outcomes. It effectively manages more intricate and delicate boundaries.

Network Architecture of Pepper Plug Seedling Segmentation Algorithm

The left side of the U-Net network structure comprises the contracting path, which doubles the feature channels and halves the image dimensions after each down sampling through convolutional layers and max pooling layers, primarily serving to extract shallow features of the target. The right side consists of the expansive path, which performs the inverse operation of down sampling through transposed convolution, meaning that the feature channels are halved while the image dimensions are doubled. The results of the transposed convolution are then concatenated with the corresponding parts of the feature map on the left side. This structure serves two main purposes: first, to extract the deep features of the image and perform feature fusion; second, to resize all extracted features back to the original image dimensions. Additionally, due to the inherent limitation of skip connections in the U-Net model for capturing long-range dependencies, an Efficient Channel Attention module (ECA)(*Wang et al.*, *2020*) is integrated between each pooling layer and subsequent convolutional layer on the contracting path (left side) of the U-Net architecture. This integration aims to strengthen inter-channel dependencies and enhance the model's capability to extract fine-grained surface features of seedlings. The enhanced U-Net network architecture is illustrated in Fig. 2.

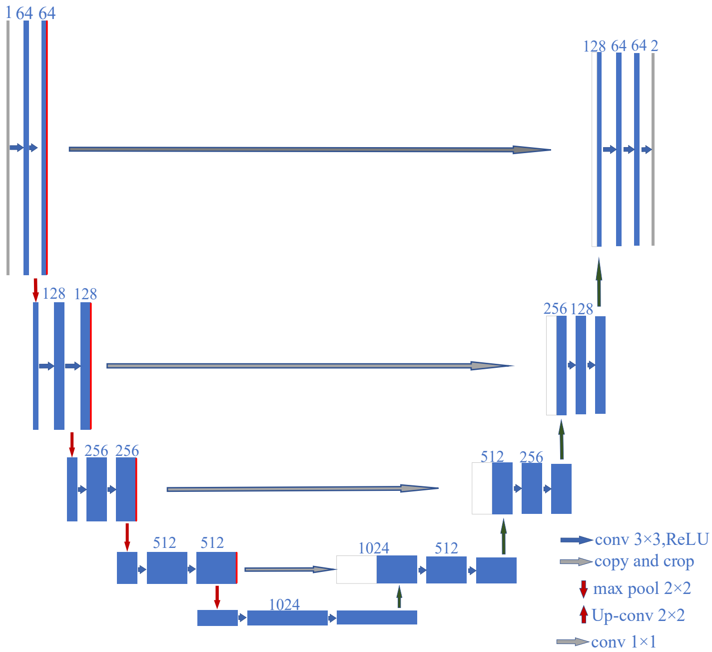

Fig. 2 - Enhanced U-Net network structure diagram

ECA Attention Mechanism

This study integrates ECA mechanism into the down-sampling part of the U-Net, effectively capturing channel interactions while avoiding dimensionality reduction. The inspiration for this attention mechanism is derived from the SE attention mechanism (*Hu et al.*, *2018*). The ECA, modified based on this foundation, does not increase the model's complexity and enhances the overall network performance. The ECA module initially computes the size of the one-dimensional convolutional kernel adaptively as indicated in Equation (1).

$$\left[k = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{r} \right|_{odd} \right] \tag{1}$$

In the equation, C denotes the channel count of the input features, while γ and b are the hyperparameters. The process of taking the absolute value and then rounding down to the nearest odd integer guarantees that the kernel size remains odd. The ECA module utilizes a one-dimensional convolution on the input features, as illustrated in Equation (2).

$$[out = Conv1D_k(in)] \tag{2}$$

The illustration of the attention mechanism is illustrated in Fig. 3.

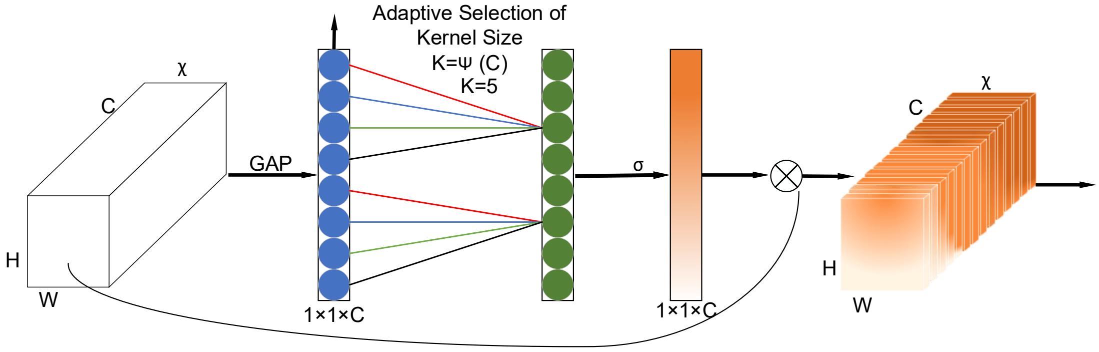

Fig. 3 - Schematic diagram of the ECA attention mechanism

Pepper plug Seedlings Feature Parameter Extraction

The characteristic parameters of plug seedlings serve as the basis for quality grading. Before transplanting the seedlings, it is essential to assess their quality based on these parameters to ensure that seedlings transplanted into the same tray are of uniform quality. In accordance with the agricultural industry standard of China, NY/T 3952-2021 (*Ministry of Agriculture and Rural Affairs of the People's Republic of China*, *2021*), this study selects five characteristic parameters for grading plug seedlings: lateral area, plant height, stem thickness, hypocotyl length, and divergence degree.

Parameter extraction of lateral area

This study employs the findContours function from the OpenCV vision library to detect the contours of pepper seedlings. This function retrieves all contours in the cv2.RETR_TREE mode, which prevents the inability to recognize internal contours due to overlapping leaves, thereby increasing the lateral area of the seedlings. The contourArea function is then used to count the pixel points and calculate the lateral area. The calculation formula is as follows:

$$A_i = contourArea(contours_i) \tag{3}$$

$$A = \sum_{i=1}^{n} A_i \tag{4}$$

Within the formula, Ai refers to the area of the ith contour, A signifies the aggregate area, and n indicates the total count of contours.

Extraction of height parameters

The method for extracting the height parameter typically involves locating the smallest enclosing rectangle around the plug seedlings and using the height of this rectangle as the plant height. Since the images of the plug seedling captured from a frontal view can have part of the stem obscured by the tray, this study tilted the camera at a certain angle during image acquisition, as shown in Fig. 4.

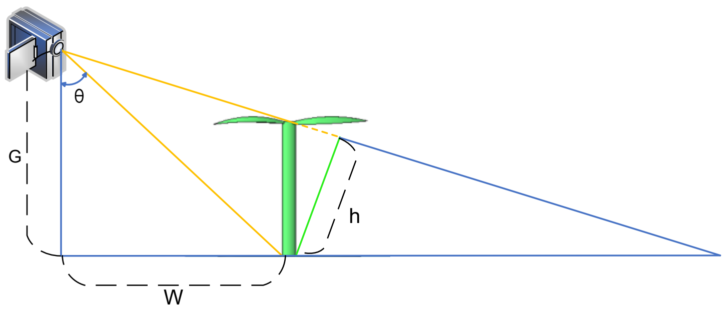

Fig. 4 - Schematic diagram of the plug seedling height error

(θ is the camera shooting angle; G is the camera height; W is the distance between the camera and the plug seedling; h is the plug seedling shooting height).

As can be discerned from Figure 4, the height of the plug seedlings depicted in the image, which is due to the angle of photography, represents the projection of their actual height. Prior to the training of the grading model, it is essential to perform a height correction on the tray seedlings to mitigate errors. The formula for height correction is presented as Equation (5).

$$H = \frac{h}{\cos(\sin^{-1}(\frac{\sqrt{G+W^2 - h^2}}{\sqrt{G+W^2}}) - \theta)} \tag{5}$$

The height parameter of the plug seedling is based on the contour calculation method during the extraction of the area parameter. Based on the known contour of the pepper seedling, the minimum rectangular bounding box of the seedling is calculated by using the boundingRect function in the OpenCV vision library, and the height of this rectangle is approximated as the height of the pepper seedling.

Extraction of width parameters

The stem thickness of plug seedlings can serve as an indicator of the "robustness" of the seedlings to a certain degree. In this study, the Canny edge detection algorithm is employed to initially identify the root location of the pepper seedlings. Following this, the contours within the area between the root and 20 pixels above it are detected. Then, using the root and the 20 pixels above as reference points, and the detected contours as criteria, the minimum bounding rectangle of the stem section is fitted to represent the stem thickness of the pepper seedlings, offering enhanced stability.

Extraction of Hypocotyl Parameters

Image scanning begins at the bottom of the seedling image, where the lowest pixel row generally has a value of 0 (black). At the root position, the pixel value changes from 0 to 1 (white) and then back to 0, which is recorded as two transitions (0 → 1 → 0). The row at which the first transition occurs is taken as the root position. Scanning continues upward, and the distance to the first true leaf is defined by the number of transitions encountered, normally equal to two. The line showing a different number of transitions is recorded as the first leaf position, and the vertical pixel distance between the root row and the leaf row represents the hypocotyl length. To avoid errors caused by occluded or deformed stems, a threshold is introduced: if the hypocotyl length is calculated to be less than 10 pixels, scanning proceeds further upward until a valid leaf position is identified. This procedure ensures accurate and reliable determination of hypocotyl length. The calculation method is shown in Equation 6. For each pixel point P_{ij}, in row i, it is checked whether it differs from the previous pixel point $P_{i(j-1)}$.

$$C_i = \sum_{j=1}^{W} P_{ij} \neq P_{i(j-1)} \tag{6}$$

During scanning, the current line number is recorded when C_i equals 2, after which the upward scan continues. If C_i is not equal to 2 and the distance between two consecutive lines exceeds 10 pixels, the vertical difference between these lines is considered the measured hypocotyl length.

Extraction of the dispersion degree parameters

n this paper, the image of a single seedling is obtained using a target detection algorithm and a segmentation algorithm, and the resulting image contains the overall characteristics of the seedling. In fact, the dispersion degree of the seedling leaves can also indicate the 'strength' of the seedling in a certain aspect. Considering that representing the dispersion degree of plug seedlings solely through the minimum bounding rectangle is susceptible to the accuracy of contour detection results, and individual plug seedling images are obtained via object detection algorithms, this paper proposes to utilize the total pixel area of the entire individual plug seedling image as the parameter to characterize the dispersion degree of single-pepper plug seedlings.

Grading of plug seedlings

Data standardisation

This study utilizes the measured lateral area, plant height, stem thickness, hypocotyl length, and dispersal degree of plug seedlings as grading criteria. The extracted pixel values for the lateral area of the seedlings typically range from 5 to 6 digits, while the pixel values for plant height generally consist of 3 digits, and stem thickness is represented by only 2 digits.

The substantial differences among these five parameters can easily lead to situations where larger parameters overshadow the characteristics of smaller ones during the training of the grading model. To mitigate this issue, it is necessary to standardize these five feature parameters.

In this paper, the StandardScaler method was used to normalise the data with the processing equation:

$$X^* = \frac{x-\mu}{\sigma} \tag{7}$$

In the equation, x denotes the original feature value, μ represents the mean of the feature, and σ signifies the standard deviation of the feature.

Classification methodology

As the research subjects of this paper are pepper seedlings that have already satisfied the requirements for transplanting, there are essentially no empty slots. Therefore, this paper divides the plug seedlings into two categories: Grade 1 and Grade 2. In machine learning, a variety of classification algorithms have shown good performance on binary classification issues. This paper discusses Support Vector Machines (SVM), Random Forest (RF) classification algorithms, and K-Nearest Neighbor (KNN) classification algorithms, with the aim of training a grading model that is most suitable for this project.

Scikit-learn (Sklearn) is a third-party machine learning library for Python that is both simple and efficient, including a range of machine learning algorithms such as Support Vector Machines, K-Nearest Neighbor classification algorithms, and Random Forest. This paper employs Sklearn directly for the training of the grading model.

RESULTS

Evaluation indicators

This research aims to identify a suitable target detection model by assessing model performance based on precision (P), recall (R), mean average precision (mAP), mean intersection over union (MIOU), model size in megabytes (MB), and detection time in milliseconds (ms). The metrics calculation formulas are presented below:

$$p = \frac{TP}{TP+FP} \times 100\% \tag{7}$$

$$R = \frac{TP}{TP+FN} \times 100\% \tag{8}$$

$$AP = \frac{1}{n}\int_0^1 P(R)dR \tag{9}$$

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

$$MIOU = \frac{1}{k+1}\sum_{i=0}^{k} \frac{TP}{TP+FP+FN} \tag{11}$$

In the equation, TP denotes the true positive samples, FP denotes the false positive samples, FN denotes the false negative samples, and N refers to the number of detection categories.

Comparison of detection algorithm results

With the continuous development of computer hardware, machine vision-based object detection methods have entered a prosperous period. The current mainstream object detection algorithms include EfficientDet (*Tan et al.*, *2020*), Faster R-CNN (*Ren et al.*, *2015*), SSD (*Liu et al.*, *2016*), and the YOLO (*Bochkovskiy et al.*, *2020; Ge et al.*, *2021; glenn-Jocherglenn-Jocher*, *2022; Redmon & Farhadi*, *2018; Ultralytics*, *2023; Wang et al.*, *2023*) series of algorithms. We conducted experimental comparisons of these popular object detection models, and the results are presented in Table 1.

Table 1

Performance of mainstream target detection models on pepper seedling dataset

Model	Precision	Recall	mAP	Parameter	Inference Time
	[%]	[%]	[%]	[MB]	[ms]
Efficientdet	77.54	84	82.29	15.8	14.83
Faster-Rcnn	96.86	99.35	99.92	113.4	39.24
SSD	92.68	98.06	98.81	95.1	5.88

Model	Precision	Recall	mAP	Parameter	Inference Time
	[%]	[%]	[%]	[MB]	[ms]
YOLOv3	94.86	89.35	95.76	246.5	9.77
YOLOv4	99.67	98.39	99.69	256.3	12.85
YOLOv5	99.10	99.10	99.50	3.8	5.33
YOLOv7	100	99.70	99.60	74.8	19.60
YOLOv8	99.03	99.03	99.33	44.8	6.11
YOLOX	99.68	99.03	99.03	36.1	6.52

Experimental results demonstrate that YOLOv5 exhibits outstanding comprehensive performance in pepper seedling detection: achieving a precision rate of 99.10% and a mean average precision (mAP) of 99.50% while maintaining high accuracy, with a model size of only 3.8 MB (representing a reduction exceeding 95% compared to YOLOv4/YOLOv7) and an inference speed of 5.33 ms per frame (equivalent to 187.6 FPS). Compared to SSD, it achieves a 6.42% precision improvement alongside a 25-fold reduction in model size (3.8 MB vs. 95.17 MB). When benchmarked against Faster-RCNN, YOLOv5 operates 7.36 times faster (5.33 ms vs. 39.24 ms) with merely a 0.42% decline in mAP. This model provides an efficient, resource-optimized, and lightweight detection solution for agricultural robotics applications requiring high robustness and low-power deployment.

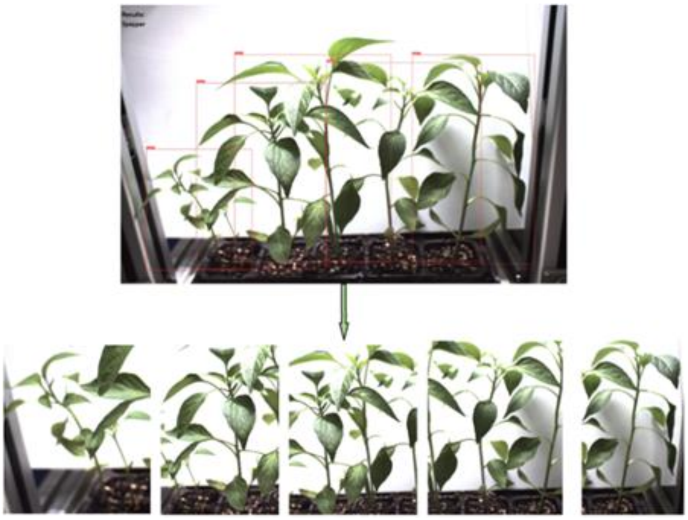

Fig. 5 - YOLOv5 test result chart

Comparison of different segmentation algorithms

To systematically evaluate the practical application performance of the enhanced U-Net model, this study adopts the Mean Intersection over Union (MIoU) and pixel-wise accuracy as core metrics for assessing segmentation precision, while additionally incorporating model size and single-frame detection time to characterize hardware deployment characteristics of segmentation performance. The improved U-Net was compared with other segmentation models, including U-Net(*Ronneberger et al.*, *2015*), U-Net++(*Zhou et al.*, *2018*), U-Net+++(*Huang et al.*, *2020*), SegNet (*Badrinarayanan et al.*, *2017*), DeepLabv3+(*Chen et al.*, *2018*), and AGNet (*Zhang et al.*, *2019*). The performance of these segmentation models on the seedling tray dataset is detailed in Table 2.

Table 2

Performance of different segmentation models on pepper seedling dataset

Model	MIoU	Precision	Parameter	Inference Time
	[%]	[%]	[MB]	[ms]
U-Net	86.30	91.72	53.65	110.01
UNet++	81.51	91.15	12.16	64.98
UNet+++	77.34	87.73	16.15	55.43
SegNet	85.28	92.95	117.9	26.36
DeepLabv3+	85.01	91.40	238	22.84
AGNet	83.60	91.73	37.34	45.41
improved U-Net	86.94 (↑0.6%)	93.01 (↑1.29%)	81.69	33.32

The experimental results presented in Table 2 indicate that the accuracy of the U-Net model with the ECA module improved by 1.9% compared to the original model. Most importantly, the detection time decreased substantially from 110 ms to 33.32 ms, thus better satisfying practical application requirements. The images processed by the segmentation algorithm are shown in Fig.6. It is clear that the improved U-Net model demonstrates superior control over the contour details of pepper seedling segmentation compared to other models. By enhancing its focus on surface feature information, the improved U-Net effectively captures the intricate details of the pepper seedling contours, resulting in better segmentation performance. This model can effectively distinguish adjacent plug seedlings and handles both the target seedling contours and interference features with notable efficacy, achieving segmentation results that closely approximate those of manual segmentation.

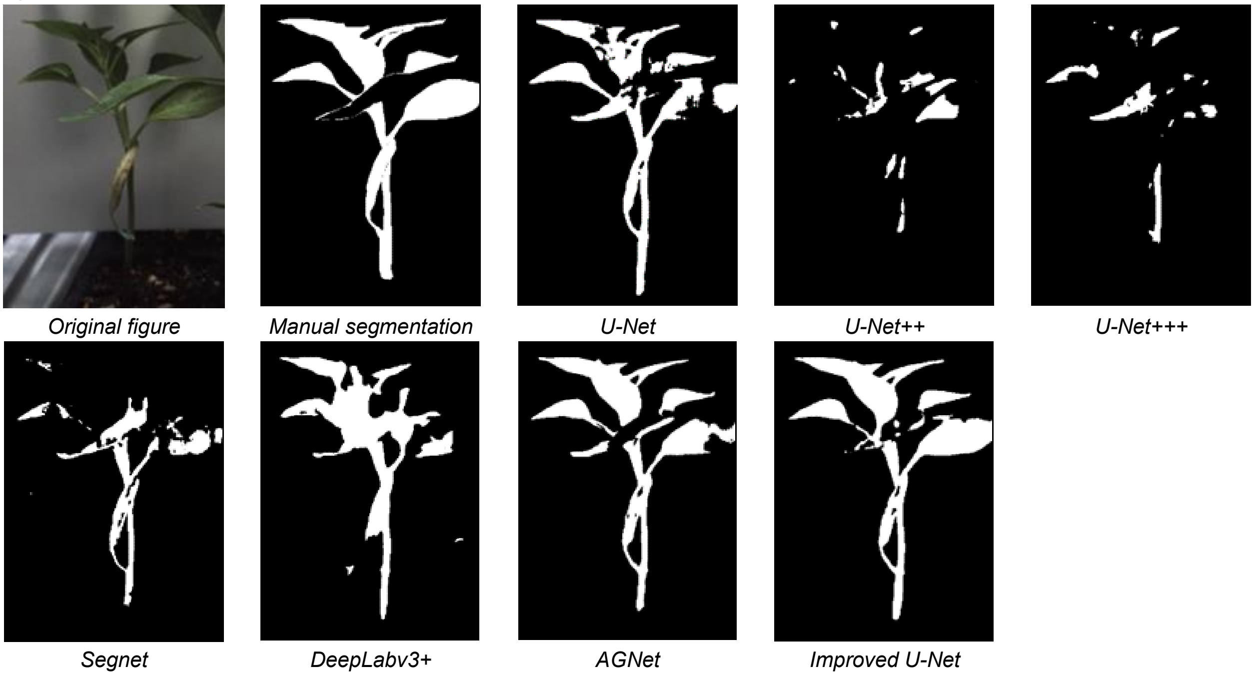

Fig. 6 - Comparison of segmentation results of different algorithms

Parameter extraction results

This study extracts feature parameters of plug seedlings from images that were pre-processed using target detection and segmentation algorithms. Parameters such as the lateral area, plant height, stem thickness, and hypocotyl length of the seedlings are extracted based on pixel points. Because the resolution of individual pepper seedling images segmented by the detection algorithm was inconsistent, all images used for feature extraction in this study were standardized to a uniform resolution of 640 × 640 pixels to avoid error accumulation. If the resolution of an image intended for feature extraction is below the target resolution, linear interpolation is used for upscaling. Conversely, if the resolution exceeds the target, area interpolation is applied for downscaling. All diagrams for parameter extraction are depicted in Fig.7.

Fig. 7 - Parameter extraction

Classification results

150 images of pepper plug seedlings were selected, which were firstly classified into first-grade and second grade seedlings according to the technical regulations of pepper greenhouse cultivation (*Hunan Provincial Administration for Market Regulation*, *2022*) as well as the opinions of relevant growers.

Parameter data such as height, stem thickness, lateral area, hypocotyl length and degree of dispersion of pepper seedlings were obtained using feature extraction methods, and they were divided into training set and validation set according to the ratio of 8:2. The accuracy and recall of Support Vector Machine, K Nearest Neighbour Classification Model and Random Forest Classification Model for grading the 150 pepper dataset are presented in Table 3.

Table 3

Performance of three classification models on the pepper dataset

	Precision	Recall
SVM	96.67%	90.42%
KNN	98%	93.74%
Random Forest	99.33%	99.62%

The results in Table 3 demonstrate that the Random Forest classification algorithm achieves higher accuracy and recall across all datasets, surpassing the performance of the other two classification algorithms. Moreover, the overall speed of the algorithm is efficient, making it suitable for the grading algorithm proposed in this paper.

CONCLUSIONS

The paper begins with the application of the YOLOv5 algorithm to detect and automatically crop entire rows of pepper seedlings into individual plant states. Subsequently, a channel attention mechanism is integrated into the down sampling section of the U-Net architecture to enhance the extraction of surface features of the seedlings, thereby facilitating their segmentation. Finally, feature extraction methods are employed to derive five parameters—side area, plant height, stem thickness, hypocotyl length, and dispersal degree—of the plug seedlings. These parameters are then used to train a Random Forest classification model for the grading of plug seedlings. The main conclusions are as follows:

1) The training results of the U-Net model enhanced with ECA attention on the pepper seedling dataset cropped by YOLOv5 indicate that the accuracy of the improved U-Net segmentation algorithm reached 93.01%, an increase of 1.29 percentage points. The image processing time was 33.32 ms, only 30% of that of the original model, demonstrating a significant speed advantage.

2) By employing feature extraction techniques to derive parameters such as area, height, stem diameter, hypocotyl length, and divergence degree of pepper plug seedlings, a grading model was trained using the random forest classification algorithm. This resulted in an impressive grading accuracy of 99.33%, facilitating the swift and precise classification of pepper plug seedlings in real-world agricultural applications.

ACKNOWLEDGEMENT

This paper was supported by the Hunan Provincial Natural Science Foundation(2023JJ50125) and Hengyang Science and Technology Innovation Programme Projects (202330066437).

REFERENCES

[1] Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence,* 39(12), 2481-2495.

[2] Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

[3] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European conference on computer vision (ECCV)*.

[4] Fu, W., Gao, J., Zhao, C., Jiang, K., Zheng, W., Tian, Y. (2022). Detection method and experimental research of leafy vegetable seedlings transplanting based on a machine vision. *Agronomy,* 12(11), 2899.

[5] Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J. (2021). Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv:2107.08430*.

[6] glenn-Jocherglenn-Jocher. (2022). YOLOv5 by Ultralytics (Version v7.0). *https://github.com/ultralytics/yolov5*

[7] Hu, J., Shen, L., Sun, G. (2018). Squeeze-and-excitation networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*.

[8] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., Wu, J. (2020). Unet 3+: A full-scale connected unet for medical image segmentation. *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*.

[9] Hunan Provincial Administration for Market Regulation. (2022). *Technical regulations for medium cultivation of chili pepper production (DB43/T 2396-2022)*. Hunan Provincial Administration for Market Regulation.

[10] Jiang, H., Shi, J., Ren, Y., Ying, Y. (2009). Application of machine vision on automatic seedling transplanting. *Transactions of the Chinese Society of Agricultural Engineering,* 25(5), 127-131.

[11] Li, M., Jin, X., Ji, J., Li, P., Du, X. (2021). Design and experiment of intelligent sorting and transplanting system for healthy vegetable seedlings. *International Journal of Agricultural and Biological Engineering,* 14(4), 208-216.

[12] Li, Y., Wei, H., Tong, J., Qiu, Z., Wu, C. (2024). Evaluation of health identification method for plug seedling transplantation robots in greenhouse environment. *Biosystems engineering,* 240, 33-45.

[13] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A. C. (2016). Ssd: Single shot multibox detector. *Computer Vision–ECCV 2016: 14th European Conference,* Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14.

[14] Ministry of Agriculture and Rural Affairs of the People's Republic of China. (2021). *General technical requirements for whole industrial chain management of pepper in solar greenhouses (NY/T 3952-2021)*. Ministry of Agriculture and Rural Affairs of the People's Republic of China. https://hbba.sacinfo.org.cn/attachment/onlineRead/b9c4af4f9f64a23f98509674b4f50b6648583ad9f97e06e8cbea99538098fcb8

[15] Redmon, J., Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

[16] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems,* 28.

[17] Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference,*. Munich, Germany, October 5-9, 2015, proceedings, part III 18.

[18] Tan, M., Pang, R., Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

[19] Tian, S., Wang, Z., Yang, J., Huang, Z., Wang, R., Wang, L., Dong, J. (2017). Development of an automatic visual grading system for grafting seedlings. *Advances in Mechanical Engineering,* 9(1), 1687814016686265.

[20] Tong, J., Shi, H., Wu, C., Jiang, H., Yang, T. (2018). Skewness correction and quality evaluation of plug seedling images based on Canny operator and Hough transform. *Computers and electronics in agriculture,* 155, 461-472.

[21] Ultralytics. (2023). YOLOv8 by Ultralytics. *https://github.com/ultralytics/ultralytics/tree/main/ultralytics/models/v8.*

[22] Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

[23] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q. (2020). ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

[24] Xu, S., Zhang, Y., Dong, W., Bie, Z., Peng, C., Huang, Y. (2023). Early identification and localization algorithm for weak seedlings based on phenotype detection and machine learning. *Agriculture,* 13(1), 212.

[25] Zhang, S., Fu, H., Yan, Y., Zhang, Y., Wu, Q., Yang, M., Tan, M., & Xu, Y. (2019). Attention guided network for retinal image segmentation. *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen*. China, October 13–17, 2019, Proceedings, Part I 22.

[26] Zhang, X., Jing, M., Yuan, Y., Yin, Y., Li, K., Wang, C. (2022). Tomato seedling classification detection using improved YOLOv3-Tiny. *Trans. Chin. Soc. Agric. Eng,* 38(1).

[27] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018*. Granada, Spain,September 20, 2018, Proceedings 4.