Computers and Electronics in Agriculture 190 (2021) 106443

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Segmentation of abnormal leaves of hydroponic lettuce based on DeepLabV3+ for robotic sorting

Zhenchao Wu [a], Ruizhe Yang [a], Fangfang Gao [a], Wenqi Wang [a], Longsheng Fu [a,b,c,*], Rui Li [a]

[a] College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
[b] Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China
[c] Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China

ARTICLE INFO

Keywords:
Uniform weights
Median frequency weights
ResNet
Semantic segmentation
Data imbalance

ABSTRACT

Hydroponic lettuce has been widely cultivated in plant factory and desiring for mechanical harvesting and packing. Sorting of hydroponic lettuce must be carried out before packing. Information perception and image processing of hydroponic lettuce is a crucial technology to develop a robotic sorting system. In this study, DeepLabV3+ models of deep learning technologies were employed with four backbones of ResNet-50, ResNet-101, Xception-65, and Xception-71 to design a vision system of segmenting abnormal leaves (yellow, withered, and decay leaves) of hydroponic lettuce. Two weights assignation methods, i.e., median frequency weights (MFW) and uniform weights (UW), were incorporated into DeepLabV3+ and compared for performance. Results showed that models trained by UW were better than that of MFW assignation method. ResNet-101 had the best segmentation performance in UW assignation method with pixel accuracy of 99.24% and mIoU of 0.8326. In terms of speed, ResNet-50 had the fast segmentation speeds with 154.0 ms per image. This study provided object detection methodology for automatic sorting device of hydroponic lettuce.

1. Introduction

Plant factory is an advanced stage of agriculture, which produces more products in less space. Hydroponics is a hot topic in the plant factory to feed the increasing population. As the cost of hydroponic equipment and electricity drops (Cui et al., 2021), hydroponics would potentially be worth within small cities and with common crops. However, due to current high energy requirements and costs of hydroponics (Wang et al., 2021), hydroponics is just worth within big cities and with high value crops. Hydroponic lettuce, a vegetable with a short growth cycle and high yield, has been widely cultivated in plant factory (Eshkabilov et al., 2021). They are harvested and packed using a harvesting – packing device, while not sorted before packing (Ma et al., 2019a, 2019b). There are abnormal leaves, i.e., yellow, withered, and decay leaves in some hydroponic lettuces. All these abnormal leaves may shorten shelf life and reduce commodity value of hydroponic lettuce (Firouz et al., 2021), and lettuce with that cannot be packaged directly. Hydroponic lettuces with yellow and withered leaves can be packed after removing these leaves. However, those with decay leaves need to be discarded since bacteria on it will breed and nitrites increase rapidly, which may harmful to human health (Herrmann et al., 2015). Therefore, the whole hydroponic lettuce containing decay leaves should be removed before packing. At present, sorting of hydroponic lettuce containing abnormal leaves is carried out manually with high labor cost and low efficiency. Therefore, there is a strong desire to introduce automatic sorting device of hydroponic lettuce.

Characteristic differences between hydroponic lettuce with different forms provided possibilities for automatic sorting device. There were some differences in appearance between healthy and abnormal hydroponic lettuce leaves (Hao et al., 2020). Healthy hydroponic lettuce leaves were green and smooth, as shown in Fig. 1a. Yellow color, withered, and decay leaves turned yellow, brown, and black, respectively, as shown in Fig. 1b. Compared with healthy leaves, texture of yellow leaves did not change much, while that of withered and decay leaves turned shrink and irregular, respectively. And the color of the leaves will be different when the degree of withered or decay of the leaves is changing. In addition, most of the abnormal leaves of hydroponic lettuce appear at the bottom of the hydroponic lettuce, which may cause light differences at the bottom of the lettuce for image capture. Therefore, the difference in color and texture between abnormal and healthy leaves made it possible to sort hydroponic lettuce containing abnormal leaves using machine vision and image processing.

* Corresponding author.
E-mail address: fulsh@nwafu.edu.cn (L. Fu).

https://doi.org/10.1016/j.compag.2021.106443
Received 30 June 2021; Received in revised form 29 August 2021; Accepted 3 September 2021
Available online 13 September 2021
0168-1699/© 2021 Elsevier B.V. All rights reserved.

Accurate segmentation of yellow, withered, and decay leaves in hydroponic lettuce was necessary for an automatic sorting device. Deep learning as a powerful technique in artificial intelligence field was becoming a prevalent way in agricultural field, although the deep learning may require more computational power. Compared with traditional machine vision, the deep learning could be more robust and learn more features of hydroponic lettuce abnormal leaves, such as color and texture of abnormal leaves. Cavallo et al. (2018) identified fresh-cut lettuce leaves with minimum color distortions based on Convolutional Neural Networks (CNNs) and found the performance loss was negligible (83% instead of 86%). Zhang et al. (2019) adopted Fully Convolutional Network (FCN) and realized segmentation of wheat ear with pixel accuracy (PA) of 92.50% in field. Stewart et al. (2019) utilized Mask R-CNN (Region-based Convolutional Neural Network) to segment northern leaf blight disease lesions in unmanned aerial vehicle images and achieved a mean intersection over union (mIoU) of 0.73. Wang et al. (2019) used Mask R-CNN with ResNet-101 (Residual Network with 101-layer) to detect and segment infected areas of tomato in field and reached a mean average precision (mAP) of 0.9964. Lin et al. (2019) proposed a semantic segmentation model based on convolutional neural networks (CNN) to segment the powdery mildew on cucumber leaf images at pixel level, which achieved an average pixel accuracy of 96.08% and intersection over union (IoU) of 72.11% on twenty test samples. Tassis et al. (2021) combined UNet and PSPNet networks to segment diseases of coffee leaves from in-field images and obtained a mIoU of 94.25%. These studies showed that semantic segmentation has been an effective method for abnormal leaves segmentation. There was no report on segmentation for abnormal leaves of hydroponic lettuce as we know. Therefore, it is promising to use a semantic segmentation network to segment yellow, withered, and decay leaves of hydroponic lettuce.

DeepLabV3+ proposed as a new semantic segmentation network has been proven to achieve good results in many studies. Jin et al. (2019) extended DeepLabV3+ to a light-weighted Bayesian version single-shot image parser for detecting defect on rail surface, which obtained a PA of 91.46% with processing speed of 5.6 fps on 250 × 160 images resolution using a GTX Titan XP GPU. Wu et al. (2020) employed DeepLabV3+ with ResNet-101 to segment cows from a dairy farm, which reported an IoU of 0.9870. Altini et al. (2020) utilized DeepLabV3+ with ResNet-18 and MFW assignation method to segment and classify glomerular from whole slide images (WSIs) stained sections, which achieved F_1 scores of 0.9240 and 0.7300 for non-sclerotic glomeruli and sclerotic glomeruli, respectively. Ayhan and Kwan (2020) applied DeepLabV3+ with Xception to classify three vegetation land covers in remote sensing images, which obtained mIoU of 0.6070 and 0.5820 for median frequency weights (MFW) and uniform weights (UW) assignation methods, respectively. Those studies showed good promising of DeepLabV3+ for semantic segmentation.

This study aimed to segment abnormal leaves, i.e., yellow, withered, and decay leaves, in pixel-level by DeepLabV3+. In addition to background, yellow leaves, withered leaves, and decay leaves, the whole lettuce was also labelled. Hydroponic lettuce image was labelled as five classes of 'background', 'yellow', 'withered', 'decay', and 'lettuce'. Pixels of 'background' and 'lettuce' in labelled images covered almost whole image, which led to a serious imbalance between the two classes and other three classes. MFW assignation method was thus employed to conduct data imbalance issue and was compared with UW assignation method. Four backbones, i.e., ResNet-50, ResNet-101, Xception-65, and Xception-71, were applied to reach an optimal backbone with DeepLabV3+.

2. Materials and methods

2.1. Hydroponic lettuce sorting sub-system

A hydroponic lettuce production system has been developed by our group to harvest, sort and pack this vegetable (Ma et al., 2019a, 2019b). Sorting sub-system, as a part of the hydroponic lettuce production system, was designed to remove hydroponic lettuce containing abnormal leaves before packing, as illustrated in Fig. 2.

The sorting sub-system consisted of image recognition part and sorting execution part. The image recognition part was used to collect and classify hydroponic lettuce that placed on a support pole, which will be described detailly late. If a hydroponic lettuce was recognized as abnormal, hydroponic lettuce bracket would be pushed under the hydroponic lettuce by the air cylinder. Then, the hydroponic lettuce was lifted by the hydroponic lettuce bracket which would be driven by stepping motor to rotate 90° or 270° to send the abnormal hydroponic lettuce to the further processing area. Otherwise, the sorting execution part would not operate and the hydroponic lettuce was transported to the packing sub-system when the hydroponic lettuce was recognized as normal.

The image recognition part included three high-definition cameras (LifeCam Studio, Microsoft, Beijing, China). Camera lens may be contaminated by debris if cameras were placed directly under the hydroponic lettuce. Therefore, cameras were arranged below the side of hydroponic lettuce, where the angle between the lens and the horizontal plane was 60° to ensure that whole hydroponic lettuce was in the field of view of the camera (74°), as shown in Fig. 2b. Considering that condensation on the camera in a humid environment will not only affect the camera's imaging, but also reduce the usage of the camera. We have waterproofed the camera and put a desiccant next to the camera to prevent condensation on the camera. RGB (Red, Green, and Blue) images were captured in 2020 at Yangling, Shaanxi Province, China. The images were taken indoors at different times during the day without artificial lighting and saved in JPEG format.

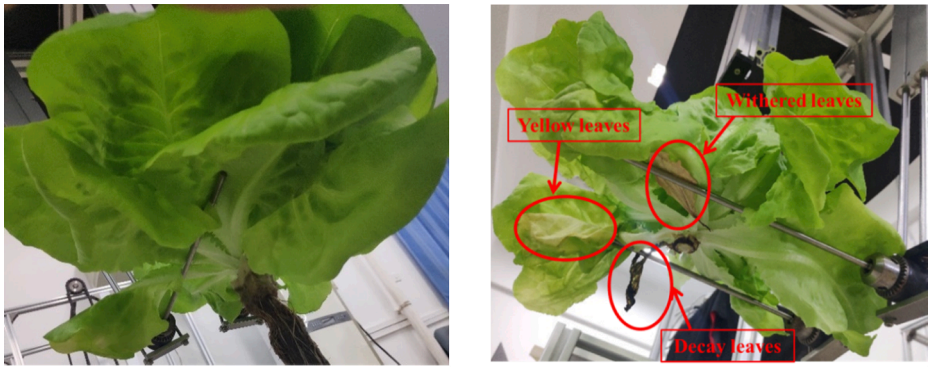


Fig. 1. Example of hydroponic lettuce with healthy leaves (a) and abnormal leaves (b).

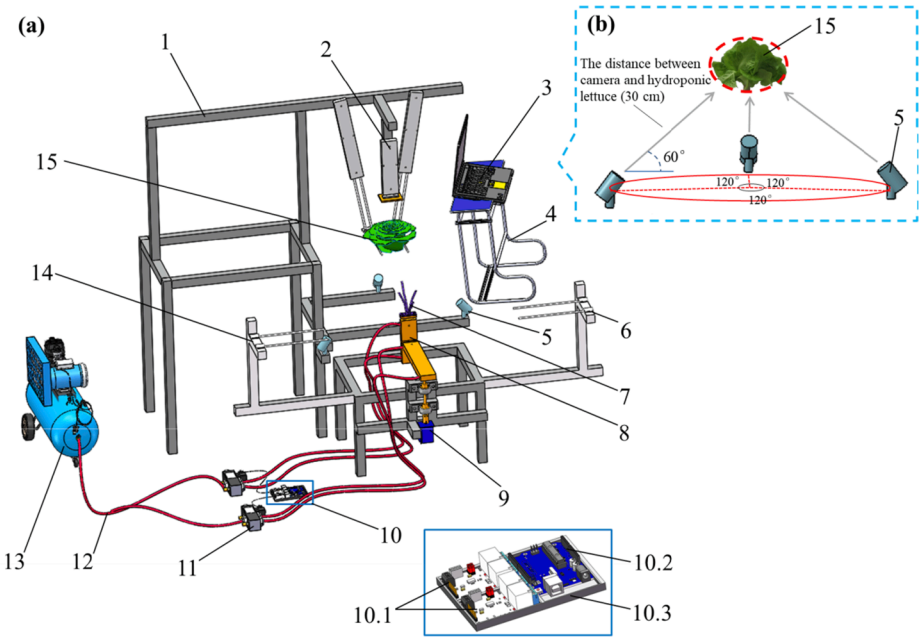


Fig. 2. Hydroponic lettuce sorting sub-system. (a) Structure of hydroponic lettuce sorting sub-system (1 frame, 2 harvesting and transportation device, 3 laptop, 4 laptop table, 5 camera, 6 further processing area A, 7 hydroponic lettuce bracket, 8 air cylinder, 9 stepping motor, 10 control components, 10.1 relay module, 10.2 single-chip microcomputer, 10.3 component housing, 11 electromagnetic valve, 12 air hose, 13 air pump, 14 further processing area B, 15 lettuce waiting to be sorted); (b) image recognition part.

2.2. Data annotation and augmentation

In total, there were 500 images with the resolution of 3024 × 4032 collected as original dataset. Considering that the camera could achieve triple digital zoom and autofocus at distances more than 10 cm, the distance between the camera and the lettuce was set to 30 cm. The size of original image was large and lead to the exhaustion of GPU memory. Shrinking images could not only save computational costs, but also serve as the basis for low pixel resolution vision sensor research. Therefore, each image was scaled to 378 × 504 pixels by bilinear interpolation. LabelMe 4.5.0 (an image labeler toolbox developed by MIT's Computer Science and Artificial Intelligence Laboratory) was used to label images. Abnormal leaves and whole hydroponic lettuce were marked with polygons and then given labels. Annotation files were saved in JSON format and then converted to PNG format images as ground truth image. Fig. 3a and Fig. 3b showed examples of an RGB image and its corresponding ground truth image, respectively. To avoid imbalance in the performance evaluation of the test set, the dataset was randomly selected as training set (400 images) and testing set (100 images) according to a ratio of 4 to 1.

Data augmentation plays a vital role in training of deep learning model. A small number of training set may lead to overfitting or non-convergence of deep learning algorithm (Deng et al., 2021). Data augmentation was performed on each image and its corresponding ground truth image of training set (Cotrim et al., 2020). The data augmentation methods in this study included image rotation in 90°, 180°, and 270° and image mirroring in horizontal and vertical axis on the training set using Matlab R2016a (MathWorks, Natick, Massachusetts, USA). For the image rotation, Matlab function '*imrotate*' was applied to get the images of 90°, 180°, and 270° by changing the function parameter '*angle*' (Gao et al., 2020). For the image mirroring, Matlab function '*flip*' was used to obtain horizontal and vertical mirroring images by setting the function parameter '*dim*' to '1' and '2' (Huang et al., 2020), respectively. After that, the training set was increased to 2400 images by the five above mentioned methods.

2.3. Network architecture of DeepLabV3+

DeepLabV3+ has encoder and decoder modules. The encoder module consists of a set of multi-scale semantic information processing module that reduce feature maps and capture semantic information, and the decoder module recovers spatial information and results in sharper segmentations. Block diagram of DeepLabV3+ is shown in Fig. 4. Atrous Spatial Pyramid Pooling (ASPP) mechanism with different atrous rates of atrous convolution is used in DeepLabV3+. It consists of a 1 × 1 atrous convolution, three 3 × 3 atrous convolutions (with atrous rates as 6, 12,



(a) (b)

Fig. 3. Example of an RGB image of hydroponic lettuce on support pole (a) and its ground truth image with 'polygon' (b). The black polygon, red polygon, green polygon, beige polygon, and blue polygon represented the areas that were labelled as 'background', 'lettuce', 'yellow leaves', 'withered leaves', and 'decay leaves', respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

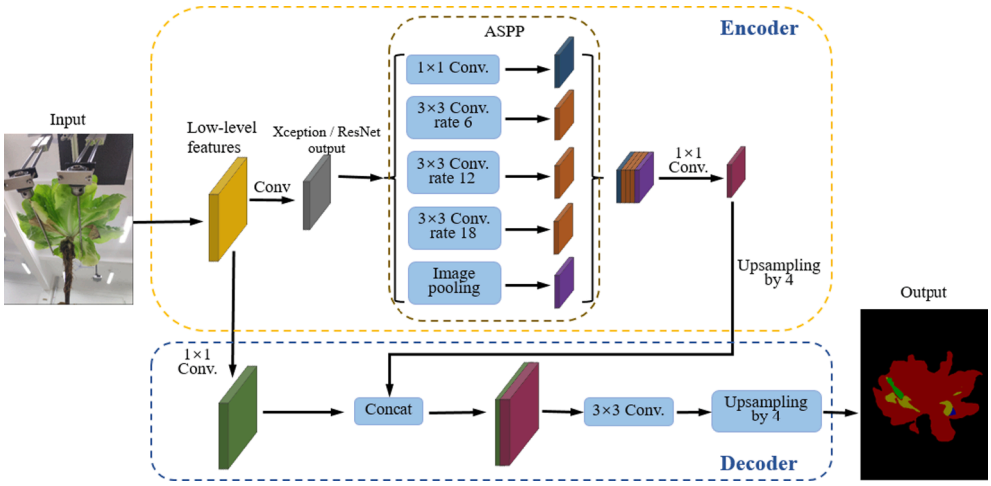


Fig. 4. Block diagram of DeepLabV3+ on hydroponic lettuce images. Note: different colors means different features maps. Conv., Convolution.

and 18), and an image pooling layer. All feature maps of ASPP are concatenated to generate combined feature map. A 1 × 1 atrous convolution is used in the combined feature map to fuse information and reduce dimensionality. Then, the processed combined feature map is input to the decoder module. The output of the encoder module is upsampled 4 times and then connected with corresponding low-level features. After that, a 3 × 3 convolution layer is used to extract features, and another simple bilinear upsampling with a factor of 4 is followed by.

There are four backbones, i.e., Xception-65, Xception-71, ResNet-50, and ResNet-101, were trained in DeepLabV3+ to obtain a better segmentation model for abnormal leaves of hydroponic lettuce. Xception, a convolutional neural network architecture based entirely on depthwise separable convolution layers (Chollet, 2017), which introduces Entry flow, Middle flow, Exit flow and Residual learning structures (65 and 71 represented the number of convolutional layers). ResNet, a residual neural network proposed to combat vanishing gradient problems during training deep convolutional networks (He et al., 2016), is modified by replacing the first 7 × 7 convolution with three 3 × 3 convolutions (50 and 101 represented the number of convolutional layers). Xception-65, Xception-71, ResNet-50, and ResNet-101 have been commonly used as backbones in DeeplabV3+ and achieved promising results (Chen et al., 2018; M.Roy and Ameer, 2021; Wu et al., 2020). Therefore, Xception-65, Xception-71, ResNet-50, and ResNet-101 were selected and compared to select to find the optimal backbone.

2.4. Weights assignation

In this investigation, DeepLabV3+ was trained in two assignation methods, i.e., UW and MFW. In UW assignation method, all of the five class weights were set as 1.0000. In MFW assignation method, weights were set according to pixel proportion of the classes (Abdalla et al., 2019). As shown in Eq. (1), f_n was frequency of class n, which was calculated by dividing the pixel number of class n (PM_n) by the total pixel number of images containing class n (PN_n). The value of n represented each class on hydroponic lettuce images in this study, background ($n = 1$), lettuce ($n = 2$), yellow ($n = 3$), withered ($n = 4$), and decay ($n = 5$). And $median(f_n)$ was the median of the f_n. The MFW of class n (denoted as w_n) was computed by $median(f_n)$ divided by f_n, as shown in Eq. (2).

$$f_n = \frac{PM_n}{PN_n} \tag{1}$$

$$\omega_n = \frac{median(f_n)}{f_n} \tag{2}$$

MFW was used to mitigate data imbalance by increasing the weights of classes with fewer pixels. The pixel number of each class (denoted by pixel count), f_n and MFW (w_n) for hydroponic lettuce dataset were shown in Table 1. A serious imbalance could be seen in the five classes. Decay pixel count was the least, with 152,410. Yellow and withered pixel counts were close, with 990,662 and 828,727, respectively. The background and lettuce had more pixels with 64,088,237 and 29,691,295, respectively. The pixel numbers of yellow/withered, background and lettuce were about 14, 1,294 and 599 times more than that of decay, respectively. The difference between f_n of each class was also obvious. After median frequency weighting, the weights of background, lettuce, yellow, withered, and decay were set to 0.0155, 0.0334, 1.0000, 1.2013, and 6.6718, respectively. The classes of yellow, decay, and withered had heavier weighting.

2.5. Training with DeepLabV3+

An Ubuntu 16.04 machine (shgentai SP16HDIET, www.Shgentai.com) with a GPU (NVIDIA Titan XP 12 GB GPU), six-core CPU (Intel Xeon E5-1650 v4, 3.60 GHz), and 32 GB memory was used for DeepLabV3+ model training and testing, which used TensorFlow 1.10.0 framework to run. To train DeepLabV3+ models for hydroponic lettuce dataset in this study, weights of a pre-trained model were used for initialization, which were fine-tuned with further training. These initial weights were obtained from pre-trained model of the PASCAL VOC 2012 dataset. Because the number of classes in this study was different from that in the PASCAL VOC 2012 dataset, the logit weights in pre-trained model were excluded. The training parameters of DeepLabV3+ used in this study were set after 18 trials, as shown in Table 2.

2.6. Performance evaluations

Evaluation indicators were used to measure performance of each trained model on testing set. In this study, model performance was evaluated by PA, IoU, mIoU, and speed. PA was percentage of correctly

Table 1
Pixel numbers for each class and the computed MFW used in DeepLabV3+ training.

	Background	Lettuce	Yellow	Withered	Decay
Pixel count	64,088,237	29,691,295	990,662	828,727	152,410
f_n	0.6728	0.3117	0.0104	0.0087	0.0016
MFW (w_n)	0.0155	0.0334	1.0000	1.2013	6.6718

Note: n is 1, 2, 3, 4, and 5, and f_1, f_2, f_3, f_4, and f_5 represent background, lettuce, yellow, withered, and decay.

Table 2
Training parameters used in DeepLabV3+.

Training parameters	Values
Learning policy	Poly
Base learning rate	0.005
Learning rate decay factor	0.1
Learning rate decay step	2000
Learning power	0.9
Training number of steps	20,000
Momentum	0.9
Train batch size	4
Weight decay	0.00004
Train crop size	[513, 513]
Train output stride	16
Fine tune batch norm	False
Initialize last layer	False
Tf initial checkpoint	deeplabv3_pascal_train_aug
Atrous rates	[6, 12, 18]

classified pixel number to total pixel number in testing set, as shown in Eq. (3). For each class, IoU was ratio of overlapped pixel number to total pixel number of ground truth images and segmentation maps, as shown in Eq. (4), and mIoU was the average value of the IoUs of all classes. Speed was time needed to segment an image, which was calculated by dividing time used to segment all images in the testing set by the total image number of testing set.

$$PA = \frac{\sum_{i=0}^{K} p_{ii}}{\sum_{i=0}^{K}\sum_{j=0}^{K} p_{ij}} \tag{3}$$

$$IoU = J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{4}$$

where p_{ij} was the number of pixels of class i in ground truth images predicted as class j on segmentation maps; p_{ii} was the number of pixels of class i in ground truth images predicted as class i on segmentation maps; K was the number of classes; A and B denoted ground truth images and segmentation maps, respectively. The value of IoU ranged between 0.0000 and 1.0000.

3. Results and discussion

3.1. Training assessment and performance

Training results of deep learning models were affected by the number of iterations. Fig. 5 illustrated the results of loss curves of eight DeepLabV3+ models based on Xception-65, Xception-71, ResNet-50, and ResNet-101 with two weights assignation methods with 20,000 iterations. Different line types and colors represented different weight assignation methods and different backbones, respectively. The loss values decreased as the number of iterations increased but were generally stable when the number of iterations reached 12,500 iterations. In UW assignation method, the lowest loss values of Xception-65, Xception-71, ResNet-50, and ResNet-101 were 0.13, 0.38, 0.22, and 0.22, respectively. But in MFW assignation method, the lowest values become smaller except for Xception-65 with 0.15. These values reached 0.35, 0.20, and 0.20 for Xception-71, ResNet-50, and ResNet-101, respectively. The low loss values demonstrated that these models adopted in this study may learned features efficiently and had good convergence ability, which had a potential to achieve desired results.

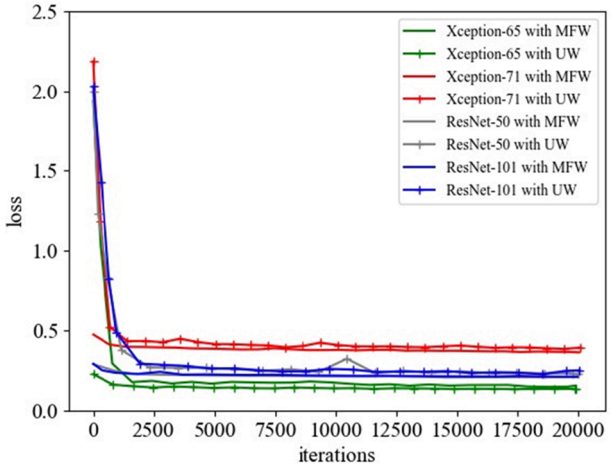


Fig. 5. Training loss curves of DeepLabV3+ models based on Xception-65, Xception-71, ResNet-50, and ResNet-101 with UW and MFW assignation method, respectively.

3.2. Comparison of different weights assignation methods

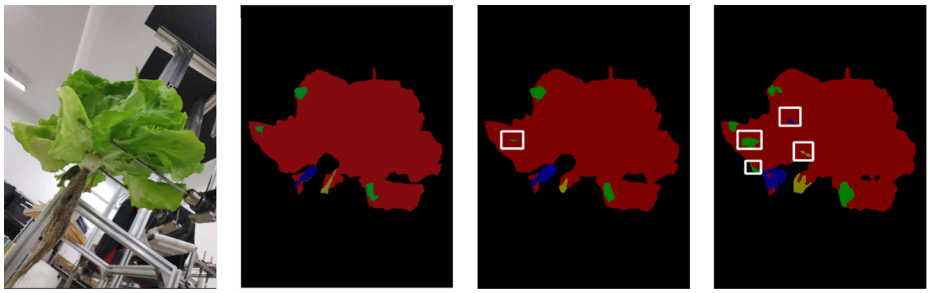
DeepLabV3+ models trained by UW assignation method outperformed those trained by MFW assignation method in evaluation indicators. From Table 3, it can be noticed that in UW assignation method, mIoUs of Xception-65, Xception-71, ResNet-50, and ResNet-101 were 0.4803, 0.7894, 0.7998, and 0.8326, respectively. Compared with the four backbones in MFW assignation method, mIoUs of same backbones in UW assignation method increased by 0.050, 0.213, 0.162 and, 0.181, respectively. Except for Xception-65, PAs of Xception-71, ResNet-50, and ResNet-101 were improved by 2.68%, 1.74%, and 1.57%, respectively.

DeepLabV3+ models in UW assignation method showed better performance in segmenting hydroponic lettuce images, clearer in border area, and fewer misclassified pixels. Segmentation maps of ResNet-101 with two weights assignation methods were presented in Fig. 6c and Fig. 6d. Hydroponic lettuce images were divided into five classes, i.e., background (black), lettuce (red), yellow (green), withered (beige), and decay (blue). Compared with ground truth image (Fig. 6b), it could be seen that areas of yellow, withered and decay leaves were larger when using MFW assignation method. There were larger false segmentation areas in the segmentation maps which were segmented by model trained with MFW assignation method, as shown by white rectangles in Fig. 6c and Fig. 6d. In addition, both of them had poor performance in area of withered leaves, which was due to the withered leaves had no obvious edge.

In general, UW assignation method outperformed MFW for yellow, withered, and decay leaves segmentation on the imbalanced hydroponic lettuce dataset. Song et al. (2021) applied DeepLabV3+ with ResNet-50, ResNet-101, Xception-65, and Xception-71 to segment calyx, branch, and wire on kiwifruit canopy images using MFW and UW assignation methods. They reported that mIoUs of models trained by UW assignation method outperformed that of MFW assignation method. Although it was not a report about hydroponic lettuce, it still gave some insights when the results of weights assignation methods were compared with that of achieved in this study.

Table 3
Performance of DeepLabV3+ assigned with UW and MFW assignation methods.

Weights assignation methods	Backbone	mIoU	PA
UW	Xception-65	0.4803	95.10%
	Xception-71	0.7894	99.06%
	ResNet-50	0.7998	99.20%
	ResNet-101	0.8326	99.24%
MFW	Xception-65	0.4303	97.32%
	Xception-71	0.5764	96.38%
	ResNet-50	0.6374	97.46%
	ResNet-101	0.6516	97.67%



(a) (b) (c) (d)

Fig. 6. Comparison of RGB image (a) and segmentation maps of DeepLabV3+ based on ResNet-101 with UW assignation method (c) and MFW assignation method (d) to ground truth image (b). The black, red, green, beige, and blue pixels represented the areas of background, lettuce, yellow, withered, and decay, respectively. Blue, green and yellow in white rectangles referred to false segmentation areas. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Comparison of different backbones

DeepLabV3+ models with ResNet-50 and ResNet-101 outperformed that with Xception-65 and Xception-71 whether in UW or MFW assignation methods. The backbones in UW were analyzed because they performed better than that of MFW assignation method. Xception-65, Xception-71, ResNet-50, and ResNet-101 obtained mIoUs of 0.4803, 0.7894, 0.7998, and 0.8326, respectively, as shown in Table 4. The mIoU of ResNet-101 (0.8326) was 0.3520, 0.0430, and 0.0328 higher than that of Xception-65, Xception-71, ResNet-50, respectively. With respect to different backbones, ResNet-101, ResNet-50, and Xception-71 achieved similar PAs of 99.24%, 99.20%, and 99.06%. But Xception-65 only achieved PA of 95.10%, which was about 4.00% lower than other three backbones.

In terms of speed, ResNet-50 took about 154.0 ms to segment lettuce, yellow, withered, and decay in one image with a resolution of 378 × 504 pixels, which was the fastest in the four backbones. Image segmentation speed of ResNet-50 was at least 20% faster than that of ResNet-101, Xception-65 and Xception-71. There was no significant difference in speed between Xception-71 and Xception-65. But there was a significant difference between Xception-65/Xception-71, ResNet-50, and ResNet-101. The speeds of models trained by the four backbones in UW assignation method were less than 250 ms/image. For the current hydroponic lettuce harvesting – packing device, it took an average of 7 s to harvest a hydroponic lettuce. In addition, although ResNet-101 was not the fastest (193.4 ms/image), it is still the optimal choice considering the PA.

In terms of IoU, DeepLabV3+ model based on ResNet-101 with UW assignation method had better performance than other three models, especially in the decay leaves, as shown in Table 4. IoUs of models trained by Xception-65, Xception-71, ResNet-50, and ResNet-101 on the decay leaves were obtained as 0.0449, 0.4947, 0.5115, and 0.6138, respectively. IoUs of the decay leaves in ResNet-101 were 0.1023, 0.1191, and 0.5689 higher than that of ResNet-50, Xception-71, and Xception-65, respectively. The prediction images were visualized as shown in Fig. 7c (Xception-65), Fig. 7d (Xception-71), Fig. 7e (ResNet-50) and Fig. 7f (ResNet-101). White circle areas in Fig. 7 showed difference on segmentation results using the four backbones when segmenting the decay leaves. Xception-65 only segmented some irrelevant background without any useful information. Xception-71 was performed better than that of Xception-65, but still could not segment the abnormal leaves. ResNet-50 and ResNet-101 could segment most of the yellow, withered, and decay leaves, as shown in Fig. 7c and Fig. 7d, where ResNet-101 performed better than ResNet-50 on the decay leaves.

Some studies reported that PA of ResNet-101 was higher than that of ResNet-50 Xception-71, and Xception-65 in different objects, but lower in speed (Fuentes et al., 2017; Kang and Chen, 2019; Takikawa et al., 2019), which were also obtained in our study. This indicated that the results in this study were expected to provide guidance for automatic sorting device of hydroponic lettuce containing abnormal leaves for indoor or outdoor agricultural environments.

4. Conclusions

To remove the hydroponic lettuce containing abnormal leaves and ensure the quality of hydroponic lettuces, a hydroponic lettuce sorting sub-system was designed based on deep learning. Abnormal leaves of hydroponic lettuce were divided into yellow, withered, and decay leaves. In this study, detection models of the hydroponic lettuce containing abnormal leaves were developed based on DeepLabV3+ with four backbones and two weight assignation methods. Results demonstrated that DeepLabV3+ model based on ResNet-101 achieved mIoU of 0.8326 and PA of 99.24% in UW assignation method. In terms of speed, it cost about 193.4 ms for ResNet-101 to segment a resolution of 378 × 504 pixels image using a GTX Titan XP 12G GPU. Therefore, detection model of the hydroponic lettuce containing abnormal leaves could meet requirements of actual production. This study could be used to guide automatic sorting device to handle hydroponic lettuce containing abnormal leaves.

CRediT authorship contribution statement

Zhenchao Wu: Data curation, Investigation, Writing – original draft. **Ruizhe Yang:** Data curation, Writing – review & editing. **Fangfang Gao:** Data curation, Writing – review & editing. **Wenqi Wang:** Investigation, Methodology, Writing – review & editing. **Longsheng Fu:** Conceptualization, Data curation, Methodology, Supervision, Writing – review & editing. **Rui Li:** Methodology, Writing – review & editing.

Table 4
Performance of DeepLabV3+ models trained by different backbones with UW.

Backbone	IoU					mIoU	PA	Speed (ms/image)
	Background	Lettuce	Yellow	Withered	Decay			
Xception-65	0.9425	0.8711	0.0715	0.4717	0.0449	0.4803	95.10%	243.4 ± 4.8[a]
Xception-71	0.9926	0.9725	0.7302	0.7570	0.4947	0.7894	99.06%	248.9 ± 4.1[a]
ResNet-50	0.9937	0.9752	0.7392	0.7793	0.5115	0.7998	99.20%	154.0 ± 3.8[c]
ResNet-101	0.9933	0.9760	0.8052	0.7749	**0.6138**	0.8326	99.24%	193.4 ± 4.0[b]

Note: Same letters in the 'speed' column represent no significant difference at the 0.05 level.

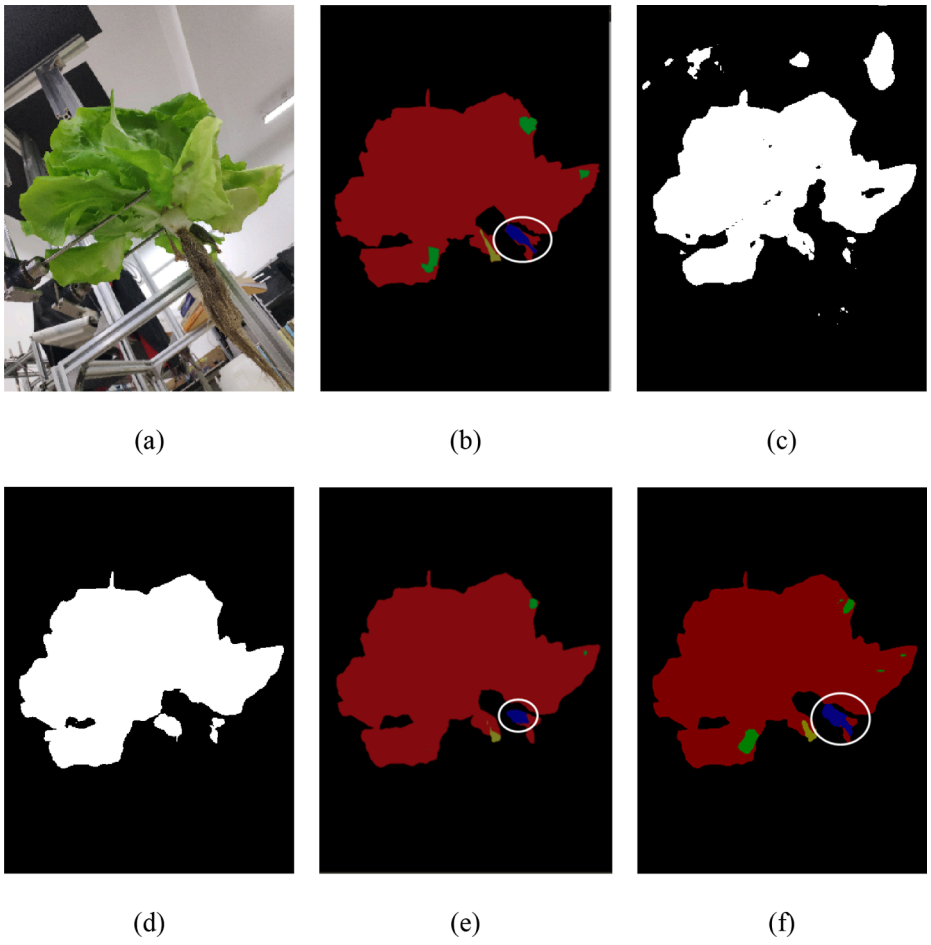


Fig. 7. Examples of RGB image (a) and its corresponding ground truth image (b) of hydroponic lettuce segmented by DeepLabV3+ based on four backbones of Xception-65 (c), Xception-71 (d), ResNet-50 (e), and ResNet-101 (f) with the UW assignation method. The black, red, green, beige, and blue represented the class of background, lettuce, yellow, withered, and decay. Blue area in white circles referred to decay leaves. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (32171897); Science and Technology Program in Yulin City of China (CXY-2020-076), Youth Science and Technology Nova Program in Shaanxi Province of China (2021KJXX-94); Recruitment Program of High-End Foreign Experts of the State Administration of Foreign Experts Affairs, Ministry of Science and Technology, China (G20200027075).

References

Abdalla, A., Cen, H., Wan, L., Rashid, R., Weng, H., Zhou, W., He, Y., 2019. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. Comput. Electron. Agric. 167 https://doi.org/10.1016/j.compag.2019.105091.

Altini, N., Cascarano, G.D., Brunetti, A., Marino, F., Rocchetti, M.T., Matino, S., Venere, U., Rossini, M., Pesce, F., Gesualdo, L., Bevilacqua, V., 2020. Semantic segmentation framework for glomeruli detection and classification in kidney histological sections. Electronics 9, 1–15. https://doi.org/10.3390/electronics9030503.

Ayhan, B., Kwan, C., 2020. Tree, shrub, and grass classification using only RGB images. Remote Sens. 12, 1333. https://doi.org/10.3390/RS12081333.

Cavallo, D.P., Cefola, M., Pace, B., Logrieco, A.F., Attolico, G., 2018. Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material. J. Food Eng. 223, 46–52. https://doi.org/10.1016/j.jfoodeng.2017.11.042.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, Munich, pp. 833–851. https://doi.org/10.1007/978-3-030-01234-2_49.

Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017 2017-Janua, pp. 1800–1807. https://doi.org/10.1109/CVPR.2017.195.

Cotrim, W. da S., Minim, V.P.R., Felix, L.B., Minim, L.A., 2020. Short convolutional neural networks applied to the recognition of the browning stages of bread crust. J. Food Eng. 277, 109916. https://doi.org/10.1016/j.jfoodeng.2020.109916.

Cui, Y., Wang, W., Wang, M., Ma, Y., Fu, L., 2021. Effects of cutter parameters on shearing stress for lettuce harvesting using a specially developed fixture. Int. J. Agric. Biol. Eng. 14, 152–158. https://doi.org/10.25165/j.ijabe.20211404.6346.

Deng, L., Li, J., Han, Z., 2021. Online defect detection and automatic grading of carrots using computer vision combined with deep learning methods. LWT - Food Sci. Technol. 149, 111832. https://doi.org/10.1016/j.lwt.2021.111832.

Eshkabilov, S., Lee, A., Sun, X., Lee, C.W., Simsek, H., 2021. Hyperspectral imaging techniques for rapid detection of nutrient content of hydroponically grown lettuce cultivars. Comput. Electron. Agric. 181, 105968. https://doi.org/10.1016/j.compag.2020.105968.

Fuentes, A., Yoon, S., Kim, S.C., Park, D.S., 2017. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors 17, 2022. https://doi.org/10.3390/s17092022.

Gao, F., Fu, L., Zhang, X., Majeed, Y., Li, R., Karkee, M., Zhang, Q., 2020. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R- CNN. Comput. Electron. Agric. 176, 105634. https://doi.org/10.1016/j.compag.2020.105634.

Hao, X., Jia, J., Gao, W., Guo, X., Zhang, W., Zheng, L., Wang, M., 2020. MFC-CNN: An automatic grading scheme for light stress levels of lettuce (Lactuca sativa L.) leaves. Comput. Electron. Agric. 179, 105847. https://doi.org/10.1016/j.compag.2020.105847.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

Herrmann, S.S., Granby, K., Duedahl-Olesen, L., 2015. Formation and mitigation of N-nitrosamines in nitrite preserved cooked sausages. Food Chem. 174, 516–526. https://doi.org/10.1016/j.foodchem.2014.11.101.

Huang, L., Pan, W., Zhang, Y., Qian, L., Gao, N., Wu, Y., 2020. Data augmentation for deep learning-based radio modulation classification. IEEE Access 8, 1498–1506. https://doi.org/10.1109/ACCESS.2019.2960775.

Jin, X., Wang, Y., Zhang, H., Liu, L., Zhong, H., He, Z., 2019. DeepRail: automatic visual detection system for railway surface defect using bayesian CNN and attention network. Acta Autom. Sin. 45, 2312–2327. https://doi.org/10.16383/j.aas.c190143.

Kang, H., Chen, C., 2019. Fruit detection and segmentation for apple harvesting using visual sensor in orchards. Sensors 19, 4599. https://doi.org/10.3390/s19204599.

Lin, K., Gong, L., Huang, Y., Liu, C., Pan, J., 2019. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Front. Plant Sci. 10, 155. https://doi.org/10.3389/fpls.2019.00155.

M.Roy, R., Ameer, P.M., 2021. Segmentation of leukocyte by semantic segmentation model: A deep learning approach. Biomed. Signal Process. Control 65. https://doi.org/10.1016/j.bspc.2020.102385.

Ma, Y., Wang, M., Cui, Y., Fu, L., Wang, W., Xu, C., 2019a. Design and test of automatic machine for hydroponic lettuce longitudinal packaging. Trans. Chinese Soc. Agric. Mach. 50, 113–121. https://doi.org/10.6041/j.issn.1000-1298.2019.09.013.

Ma, Y., Xu, C., Cui, Y., Fu, L., Liu, H., Yang, C., 2019b. Design and test of harvester for whole hydroponic lettuce with low damage. Trans. Chinese Soc. Agric. Mach. 50, 162–169. https://doi.org/10.6041/j.issn.1000-1298.2019.01.017.

Soltani Firouz, M., Alimardani, R., Mobli, H., Mohtasebi, S.S., 2021. Effect of modified atmosphere packaging on the mechanical properties of lettuce during shelf life in cold storage. Inf. Process. Agric. https://doi.org/10.1016/j.inpa.2020.12.005.

Song, Z., Zhou, Z., Wang, W., Gao, F., Fu, L., Li, R., Cui, Y., 2021. Canopy segmentation and wire reconstruction for kiwifruit robotic harvesting. Comput. Electron. Agric. 181 https://doi.org/10.1016/j.compag.2020.105933.

Stewart, E.L., Wiesner-Hanks, T., Kaczmar, N., DeChant, C., Wu, H., Lipson, H., Nelson, R.J., Gore, M.A., 2019. Quantitative phenotyping of northern leaf blight in UAV images using deep learning. Remote Sens. 11, 2209. https://doi.org/10.3390/rs11192209.

Takikawa, T., Acuna, D., Jampani, V., Fidler, S., 2019. Gated-SCNN: Gated shape CNNs for semantic segmentation. In: Proc. IEEE Int. Conf. Comput. Vis., pp. 5228–5237. https://doi.org/10.1109/ICCV.2019.00533.

Tassis, L.M., Tozzi de Souza, J.E., Krohling, R.A., 2021. A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images. Comput. Electron. Agric. 186 https://doi.org/10.1016/j.compag.2021.106191.

Wang, Q., Qi, F., Sun, M., Qu, J., Xue, J., 2019. Identification of tomato disease types and detection of infected areas based on deep convolutional neural networks and object detection techniques. Comput. Intell. Neurosci. 1–15 https://doi.org/10.1155/2019/9142753.

Wang, W., Ma, Y., Fu, L., Cui, Y., Majeed, Y., 2021. Physical and mechanical properties of hydroponic lettuce for automatic harvesting. Process. Agric Inf. https://doi.org/10.1016/j.inpa.2020.11.005.

Wu, D., Yin, X., Jiang, B., Jiang, M., Li, Z., Song, H., 2020. Detection of the respiratory rate of standing cows by combining the DeepLab V3+ semantic segmentation model with the phase-based video magnification algorithm. Biosyst. Eng. 192, 72–89. https://doi.org/10.1016/j.biosystemseng.2020.01.012.

Zhang, D., Wang, D., Gu, C., Jin, N., Zhao, H., Chen, G., Liang, H., Liang, D., 2019. Using neural network to identify the severity of wheat fusarium head blight in the field environment. Remote Sens. 11 https://doi.org/10.3390/rs11202375.