IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 13, 2020, accepted March 17, 2020, date of publication March 23, 2020, date of current version March 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2982456

Deep Learning-Based Object Detection Improvement for Tomato Disease

YANG ZHANG, CHENGLONG SONG, AND DONGWEN ZHANG
School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China

Corresponding author: Yang Zhang (zhangyang@hebust.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61440012, in part by the Scientific Research Foundation of Hebei Educational Department under Grant ZD2019093, and in part by the Fundamental Research Foundation of Hebei Province under Grant 18960106D.

ABSTRACT To improve the recognition model accuracy of crop disease leaves and locating diseased leaves, this paper proposes an improved Faster RCNN to detect healthy tomato leaves and four diseases: powdery mildew, blight, leaf mold fungus and ToMV. First, we use a depth residual network to replace VGG16 for image feature extraction so we can obtain deeper disease features. Second, the k-means clustering algorithm is used to cluster the bounding boxes. We improve the anchoring according to the clustering results. The improved anchor frame tends toward the real bounding box of the dataset. Finally, we carry out a k-means experiment with three kinds of different feature extraction networks. The experimental results show that the improved method for crop leaf disease detection had 2.71% higher recognition accuracy and a faster detection speed than the original Faster RCNN.

INDEX TERMS Faster RCNN, disease recognition, deep residual network, K-means clustering, disease diagnosis.

I. INTRODUCTION

Crop disease detection is the basis of crop disease prevention to guarantee crop quality. Traditional detection methods for crop disease mainly depend on manual observation and consequently lead to low detection efficiency and poor reliability. Farmers lack professional knowledge, and agricultural experts cannot serve the field at all times so that they miss the best time for prevention. In recent years, image processing [1], pattern recognition [2], computer vision [3] and other technologies have developed rapidly. Computer automatic detection of diseases provides method for effectively solving agricultural problems.

The traditional machine vision [4] method for the detection of crop leaf diseases follows three steps: (1) image preprocessing, (2) researchers manually designing complex disease features for feature extraction [5], and (3) machine learning algorithms [6] for classifying crop diseases. Khirade and Patil [7] discussed how to detect crop diseases based on leaf images and some feature extraction algorithms. Sasaki *et al.* [8] proposed automatic recognition technology for cucumber anthracnose. According to the different spectral reflection characteristics and the influence of optical filtering on disease recognition. They used a genetic algorithm to establish recognition parameters from two angles of spectral reflection and shape characteristics to identify diseases. Traditional machine vision methods require complex preprocessing and design of image features, which is time consuming and labor intensive. In particular, the effectiveness of this method depends largely on the accuracy of the artificial design features and the learning algorithm.

Neural networks [9] contribute to image recognition. They have excellent nonlinear fitting capabilities so that they can achieve higher accuracy in some image recognition tasks. El-Helly *et al.* [10] used an artificial neural network to better recognize cucumber powdery mildew, downy mildew and leaves damaged by leaf dips. Sammany and Medhat [11] employed genetic algorithms to optimize neural networks and support vector machines for recognizing plant disease images. Baum *et al.* [12] conducted disease recognition on barley and used edge detection [13] to separate the diseased area from the background area. Their experimental results showed that the diseased area could be extracted. However, these shallow structures have great limitations. The generalization ability of some complex classification problems is limited.

The associate editor coordinating the review of this manuscript and approving it for publication was L. Zhang.

With the rapid development of deep learning [14], the accuracy of image classification [15] and object detection [16] has greatly improved, and it can accurately classify large datasets, even better than humans in many aspects. Deep learning has the advantage of directly extracting classification features. Additionally, the deep learning feature extraction method is suitable for classification on various occasions and has a strong generalization ability. With the continuous improvement of many researchers, deeper neural networks have been proposed, such as AlexNet [17] and VGGNet [18]. Lee and Kwon [19] and Lee *et al.* [20] developed a cross-domain convolutional neural network for a variety of hyperspectral images. After three sets of experiments, the results improved accuracy by 1% to 3% compared to the method using independent convolutional neural networks. Sun *et al.* [21] used a convolutional neural network to detect tea disease. They carried out data enhancement and image segmentation for tea images and achieved higher accuracy through frequently adjusting iteration times and learning rates. Therefore, the general trend has been to use deep learning to detect crop diseases in agriculture.

To improve the low recognition model accuracy of traditional crop leaf diseases and the lack of consideration of leaf positioning, this paper uses deep learning technology and the object detection algorithm faster regional convolutional neural networks (Faster RCNN) [22] to detect tomato disease. First, we use the depth residual network [23] to replace VGG16 for image feature extraction so we can obtain deeper disease features. Second, the k-means [24] clustering algorithm is used to cluster the bounding boxes. We improve the anchoring according to the clustering results. The improved anchor frame tends toward the real bounding box of the dataset. Finally, the two improved methods are combined to detect tomato diseases. The experiment shows that our approach can both improve the accuracy of tomato disease recognition and detect the position of diseased leaves.

This paper makes the following contributions:

- We use the residual network instead of VGG16 to extract features in Faster RCNN. ResNet can extract deeper tomato disease features than VGG16.
- We use the k-means algorithm to cluster the bounding boxes of datasets and improve the anchor in Fast-RCNN according to the clustering results.

Our method improves the feature extraction network and anchor. In feature extraction, we use resnet to extract features. It can both extract the deep disease characteristics of tomato disease and solve the problem of gradient disappearance caused by the increase of layers. In region proposal network (RPN), we use k-means to cluster the bounding boxes of tomato disease datasets. We improve anchor by clustering results.

The rest of this paper is organized as follows. Section II describes the methods and structure of this paper. Section III introduces the dataset and two parts of the experiments. Section IV is the conclusion and prospects.

II. MATERIALS AND METHODS

A. TOMATO DISEASE DETECTION METHOD BASED ON IMPROVED FASTER RCNN

The detection model of tomato disease based on improved Faster RCNN is depicted in Fig. 1. To better apply the algorithm model to disease detection, we perform several improvements. First, we use ResNet instead of the VGG16 network. The tomato disease image of any size is convoluted by ResNet to obtain a deep feature map, which is used as the input of a fast regional convolutional neural networks (Fast-RCNN) and RPN. Second, the k-means algorithm was used for clustering analysis on the bounding boxes of tomato disease objects to improve the anchor. In the RPN, we use the improved anchor frame to convolve the feature image, then use bbox regression to locate the coordinates of the convolution results, and use softmax [25] classification to classify the convolution results to distinguish whether they are foreground or background. Finally, the region of interest (ROI) [26] is obtained as the input of the Fast-RCNN. In the Fast-RCNN, we map to the feature map according to the candidate box coordinates provided by the ROIs and then conduct max-pooling. Finally, we use fully-connected layers and softmax to classify the specific diseases of the region proposals and then use bbox regression [27] to obtain the position offset of each region proposal, which is used to obtain a more accurate tomato disease object detection frame by regression.

B. FEATURE EXTRACTION NETWORK

Feature extraction aims to obtain the original image feature. First, images of the same size are obtained through the normalization process [28], and then the features of the image are extracted through convolution pooling of the CNN network. Finally, a feature map is generated as an input to the RPN and Fast-RCNN.

Faster RCNN generally uses a VGG16 network for feature extraction of images. Resnet has more convolution layers than VGG16, and can get more object features by convolution. Resnet has a skipping structure, which can directly skip one or more layers. It solves the problem of gradient disappearance caused by stacking of layers. However, the VGG16 network cannot extract features deeply of tomato leaf diseases. Therefore, the VGG16 network is replaced with ResNet. The input image is unified first and then input to the feature extraction network for unified output. We reduce the image by 16 times in the feature extraction stage, the same as VGG16, and the MobileNet in the experimental part is the same.

The deep residual network is implemented by a feedforward neural network with a skip connection. The skipping structure performs identity mapping so that the output of one layer can be direct across several layers as the input of the latter layer. The calculation method has the advantage that no other coefficients are imported, and the calculation does not increase significantly. By cross-layer operation and reusing intermediate features, gradient disappearance [29] caused by

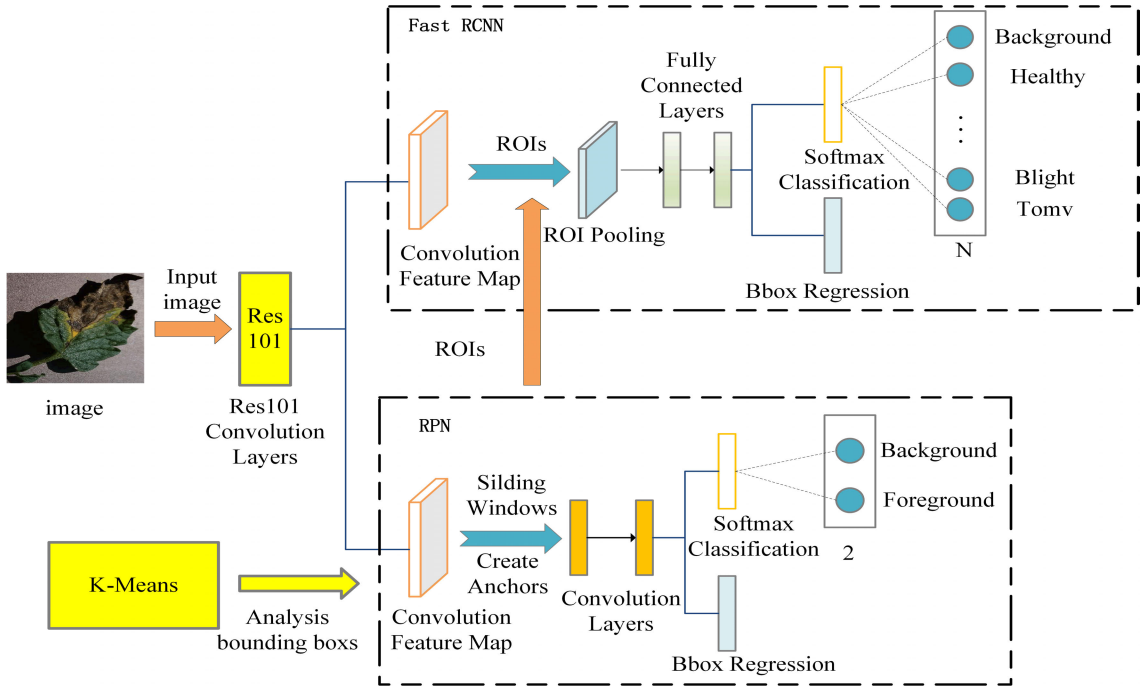


FIGURE 1. Tomato disease leaf detection model based on improved Faster RCNN.

increasing the number of neural network layers is avoided. Therefore, the deep residual network plays an important role in the field of recognition. Moreover, deep residual network is also widely applied on fault diagnosis and fault tolerant control [30], [31].

C. REGION PROPOSAL NETWORK

In the overall framework of Faster RCNN, the RPN network is an important network connected to the feature extraction network. It is the core network of Faster RCNN. The tomato disease leaf positioning task is performed by RPN. RPN generates region proposals with a sliding window [32] centered on each point over the convolutional feature map output by the last shared convolutional layer. The sliding window slides on each point of the feature map to generate k anchor points. Therefore, there are 4K outputs in the bbox regression layer, indicating the position coordinates of tomato diseased leaves, and 2K outputs in the classification layer, indicating whether the tomato leaves are diseased (foreground or background). Since the classification layer has only two deterministic outputs, we depicted two deterministic branches in Fig. 1.

Faster RCNN has 9 anchors computed by three scales (128^2, 256^2, 512^2) and three ratios (1: 1, 1: 2, 2: 1). These anchors can cover most of the objects in object detection, but they cannot perform the best on the dataset of tomato disease leaves in our experiment. It is just a universal size. To improve the size of anchors, this paper introduces the k-means algorithm so that the anchors in the algorithm tend to the bounding box of our dataset.

K-means clustering is a typical unsupervised learning algorithm [33]. It can classify data without the need to label the data. The k-means algorithm regards distance as the criterion for the similarity between data objects. If two data objects are similar to each other, they are of the same category.

In our network framework, k-means is not connected to the feature extraction network. It is an independent module. K-means is not directly applied to RPN. K-means is used to obtain the scale and size of the border that tends to the real value by clustering the border boxes of the dataset. Then, we input the size of the analyzed bounding box into RPN to improve the anchor.

The tomato disease bounding box is clustered by k-means into the following steps:

1) We take the size of the tomato disease bounding box as the original input of the k-means algorithm. Assume there are n samples of the tomato disease bounding box, we set a value k as the number of categories for the bounding box samples.
2) k is randomly selected from the tomato sample as cluster points. The distance between each disease boundary box sample and cluster point is calculated. Each sample is classified to the cluster point with the smallest distance. Finally, a total number of all bounding box samples are classified into k categories according to the smallest distance, and a cluster is completed.
3) Based on the bounding box samples by each cluster point, the positions of the cluster points are recalculated.
4) We set a threshold t and calculate the distance between the new cluster point and the original cluster point. If the distance is greater than t, we repeat steps (2) and (3). If the distance is less than t, the result tends

to stabilize and converge. We obtained the clustering results of the tomato disease leaf bounding box.

D. ROI POOLING AND CLASSIFICATION REGRESSION

Based on the mapping relationship between the feature maps and the original image, bounding boxes are transformed into the corresponding regions of interest on the feature maps. Max-pooling is employed to obtain a tomato disease feature map with a fixed size and to complete the normalization operation. The ROI pooling layer not only greatly improves the speed of tomato disease diagnosis but also achieves end-to-end training [34].

The classification task is performed by softmax. After the fully-connected operation, the feature map generates a feature vector. The feature vector is used by softmax to calculate the probable disease to which the tomato belongs.

The regression task is performed simultaneously with the classification task. The regression operation fine-tunes the bounding box of the diseased leaves so that the experimentally generated bounding box of the tomato disease tends to the true object position.

III. RESULTS AND DISCUSSION

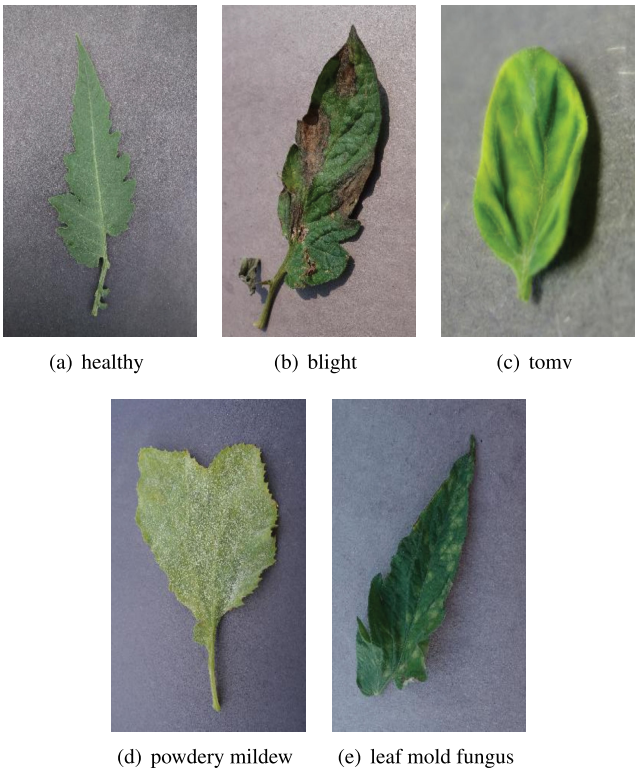
In this section, we introduce the experimental dataset and evaluation methods, and we describe the results from three aspects. First, the bounding box of tomato disease images was analyzed by the k-means clustering algorithm, and the anchor was improved based on the analysis results. Second, we selected different feature extraction networks to analyze tomato disease leaves. Finally, we used different feature extraction networks with improved anchors for experimental comparative analysis.

A. IMAGE DATASETS

The training dataset is from AIChallenger (https://challenger.ai/competition/pdr2018), which are laboratory data. This dataset contains 61 categories. We use the Python program to process annotation files. The annotation file contains the number of each disease category. We separated the pictures of all tomato diseases through the corresponding relationship between the number and the picture name. We selected four tomato disease images for the experiment. The selected image dataset contains 4,178 images. We used the random division method to divide the images of four tomato diseases. Sixty% of the dataset is selected as the training set, 30% of the dataset is selected as the validation set, and 10% of the dataset is selected as the test set. Fig. 2 shows samples of tomato health and four other diseases (powdery mildew, blight, leaf mold fungus, and ToMV), in which each leaf has different characteristics, such as different ground textures, color brightness, and shading effects. The image annotation tool LabelImg [35] is used to label images.

B. EXPERIMENTAL ENVIRONMENT AND ALGORITHM EVALUATION

Faster RCNN is based on TensorFlow, and the model has experimented on an NVIDIA GeForce GTX 1060 GPU with 12 GB of memory. The operating system is Ubuntu. In the parallel computing architecture, the CUDA version is v9.0 and the cuDNN version is v7.6.5.


(a) healthy (b) blight (c) tomv

(d) powdery mildew (e) leaf mold fungus

FIGURE 2. Images of tomato disease.

The algorithm evaluation standard used in this paper is mean average precision (mAP) [36], which consists of precision, recall and mean.

Precision and recall are important values for image processing and object detection. Precision refers to the proportion of correctly classified and located categories in the returned results to the total returned results. The recall rate refers to the proportion of correctly classified and located categories in the returned results to the total related categories.

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

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

TP (true positive) represents the number of positive samples classified correctly in the prediction result. FP (false positive) represents the number of negative samples for which the prediction result is incorrectly detected. FN (false negative) represents the number of undetected positive samples in the prediction result.

According to the values of precision and recall, the precision-recall curve can be drawn with the precision as the ordinate and the recall as the abscissa, and the AP is the area under the precision-recall curve. Therefore, the integral

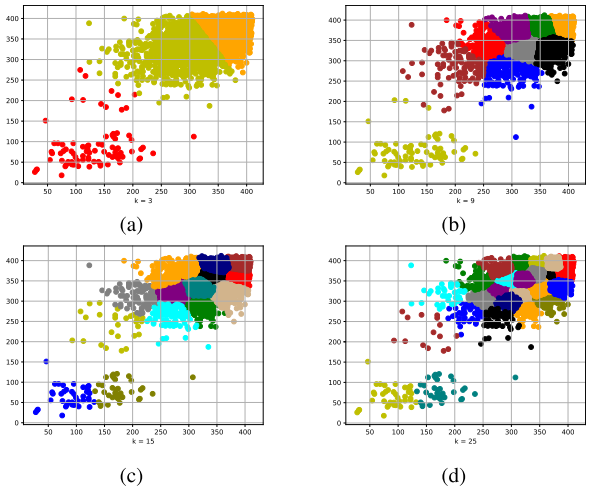


FIGURE 3. K-means cluster the results of the bounding box of this dataset.

of the precision-recall curve is the AP.

$$AP = \int_0^1 p(r)dr \tag{3}$$

The recall is divided into n blocks$[0, \frac{1}{n} \ldots, \frac{n-1}{n}, n]$, and then the mAP is determined by the average value. If there are N categories in total, the mAP is calculated using (4).

$$mAP = \frac{\sum_{n=1}^{N} AP(n)}{N} \tag{4}$$

C. RESULTS AND ANALYSIS

1) THE K-MEANS CLUSTERING ALGORITHM ANALYZES BOUNDING BOXES

We used a k-means clustering algorithm to select the proper size of tomato disease leaves. The clustering results (k=3, k=9, k=15, k=25) are shown in Fig. 3. By analyzing the clustering results, we find that the size of the bounding box of the tomato disease leaf object is clustered between [200, 400]. To increase the intersection over union (IOU) ratio, we need to subdivide the scale interval and ratio. Because IOU is set as a positive sample at 0.7, we added some scales and ratios based on the 9 types of anchors preset by Faster RCNN, so that the area ratio of each basic scale is approximately 0.7. Finally, the bounding box was $(128^2, 256^2, 320^2, 384^2, 448^2, 512^2)$ as the scale, and (0.5, 0.75, 1, 1.5, 2) as the ratio was used for experiments.

2) ANALYSIS OF DIFFERENT FEATURE EXTRACTION NETWORKS BASED ON THE FASTER RCNN ALGORITHM

In general, different feature extraction networks [37] have different capabilities for feature extraction. Therefore, we use 3 different CNN models (vgg16, mobile, res101) instead of Faster RCNN's feature extraction network for comparison experiments. Table 1 shows the results of 10,000 iterations using the pretrained model in the same environment. Time represents the time of each test.

TABLE 1. Comparative experiments of different feature extraction networks.

Model	mAP(%)	Time(ms)
Faster RCNN	95.83	731
Faster RCNN-mobile	94.67	142
Faster RCNN-res101	97.18	452

From the results in Table 1, it can be seen that Faster RCNN-res101 not only improves the detection speed but also increases mAP by 1.35% compared to Faster RCNN. The reason is that the characteristics of the residual network are easy to optimize and can improve the accuracy by adding deeper layers. ResNet's internal residual block uses skip connections [38], which alleviates the problem of gradient disappearance caused by increasing the number of layers in deep neural networks. In contrast, Faster RCNN-mobile greatly improved the detection speed, but it decreased by 1.16% on mAP compared to the original Faster RCNN. Because MobileNet [39] is based on a streamlined architecture, it uses deep separation convolution to build lightweight deep neural networks [40]. In summary, we selected the residual network to detect tomato diseases.

3) COMPARATIVE ANALYSIS OF FASTER RCNN FOR DIFFERENT FEATURE EXTRACTION NETWORKS AFTER THE IMPROVED ANCHOR

We used the k-means algorithm to perform cluster analysis on the bounding box of the dataset to obtain the approximate distribution of the bounding box size of the diseased tomato leaf. We improved the anchor to refine the size interval. After improving the algorithm, the detection results tended more to the ground truth, and the accuracy of detection improved. Table 2 shows the results of our experiments after improving the anchor where True represents that the network improved using k-means, and False is the opposite. Healthy, powdery mildew, blight, leaf mold fungus, and ToMV are the names of tomato diseases.

Table 2 reveals that the accuracy of the improved model using k-means was generally higher than the accuracy of ordinary models. It shows that the k-means clustering algorithm used in this paper can indeed improve accuracy. When k-means was false in Table 2, the times of the three models were 731 ms, 142 ms, and 452 ms. When k-means was true, it indicated the detection time after clustering. The times of the three models were 751 ms, 160 ms, and 470 ms. The experimental results show that we used cluster analysis to increase the number of anchors from 9 to 30, and the detection time increased but did not increase much. The reason is that we increased the number of anchors, resulting in an increase in the calculation of the RPN module, and the detection time of the whole Fast-RCNN also increased correspondingly. However, the detection time is the detection time of the Fast-RCNN overall framework, and the increased time of the

TABLE 2. Comparative experiment of different feature extraction networks after improved anchors.

Model	K-means	Healthy(%)	Powdery mildew(%)	Blight(%)	Leaf mold fungus(%)	Tomv(%)	mAP(%)	Time(ms)
Faster RCNN	False	90.91	90.91	98.23	99.73	99.39	95.83	731
	True	90.12	98.58	96.79	99.73	99.84	97.01	751
Faster RCNN-mobile	Fasle	97.12	94.30	97.84	90.34	93.76	94.67	142
	True	90.91	97.74	99.39	99.10	96.69	97.37	160
Faster RCNN-res101	Fasle	90.91	98.91	97.48	98.77	99.84	97.18	452
	True	99.12	96.62	98.03	99.59	99.36	98.54	470

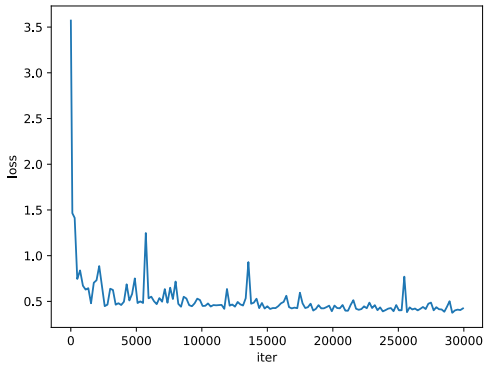


FIGURE 4. The graph of loss as a function of the number of iterations.

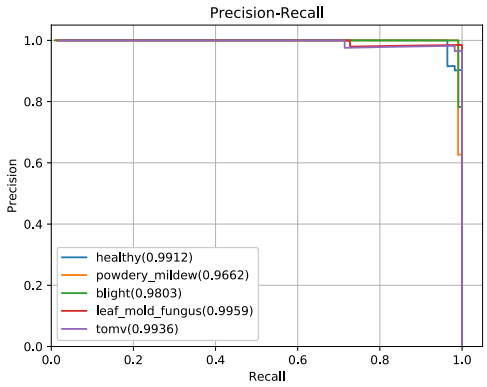
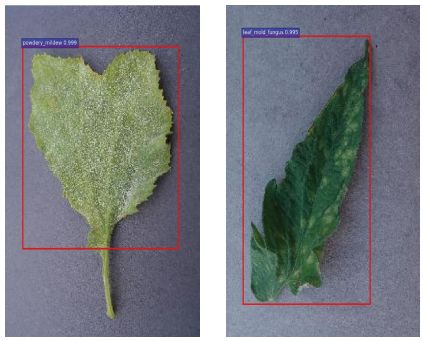


FIGURE 5. The relationship between precision and recall at 25,000 iterations.

(a) healthy (b) blight (c) tomv



(d) powdery mildew (e) leaf mold fungus

FIGURE 6. Improved Faster RCNN was used to detect the results.

anchor frame was only a small part of it. From the overall point of view, the increase in detection time was not much.

By comparing the data of the mAP and k-means, it can be seen that in Faster RCNN, the mAP analyzed by k-means increased by 1.18% to 97.01%. In Faster RCNN-mobile, the mAP after using k-means analysis increased by 2.7% to 97.31%. In Faster RCNN-res101, the mAP after using k-means analysis increased by 1.36% to 98.54%. The reason is that we use the k-means algorithm to make the anchor tend toward the real value of the data. The results show that mAP generally improved after k-means clustering analysis on the same network model, which demonstrates that our method is effective.

Through the above comparison, we found that the k-means improved anchor can improve the detection of mAP, and the detection time also increased correspondingly.

The loss represents the average of the cross-entropy of each iteration in the training process. In general, the lower the loss value is, the better the classification effect of the model is. The image of loss changes with the number of iterations of the algorithm in this paper, as depicted in Fig. 4. The loss value in the image decreased with the number of iterations. After 20,000 iterations, the curve in the image became flat and converged. The horizontal axis of the precision-recall curve was that recall reflects the ability of the classifier to cover positive cases, while the vertical axis of precision reflects the accuracy of the classifier in predicting positive cases. Therefore, the precision-recall curve can reflect the trade-off between the classifier's recognition accuracy and converge ability. From (1) to (4), it can be seen that mAP is the average value of the area under the precision-recall curve. Due to the lack of experimental data and the characteristics of the data, Fig. 5 shows the precision-recall curve of this method after 25,000 iterations on the tomato disease detection dataset. The image includes the precision-recall curves of five

experimental subjects: healthy, powdery mildew, blight, leaf mold fungus, and ToMV.

Fig. 6 shows the results of diagnostic methods for some tomato diseases. This shows that our method can detect tomato disease leaves correctly. In Fig. 6 (a) and (d), the detection probability of healthy and powdery mildew is 0.998 and 0.999, respectively. In Fig. 6 (b) and (c), the disease characteristics of blight and ToMV are obvious, so their detection probability is up to 1.000. In Fig. 6 (e), the disease characteristics of leaf mold fungus are easily confused with other diseases, so the lowest detection probability is 0.995.

IV. CONCLUSION

This paper improves the Faster RCNN algorithm to detect diseased tomato leaves, which can both recognize tomato diseases and detect tomato leaf locations. To make the anchors in the algorithm closer to the ground truth of our dataset, we use the k-means algorithm to cluster the bounding boxes of tomato disease images and improve the anchors based on the results. We selected ResNet101 to replace VGG16 for feature extraction, which can extract the deep features of tomato disease. The experimental results show that our method can effectively detect and recognize tomato diseases and has higher detection accuracy than the original Faster RCNN. However, the dataset is laboratory data, so only a single leaf disease in the image can be detected. Future works will include images of natural plants that should be collected for detection. In addition, tomato diseases have other disease characteristics, such as fruits and stems. Some diseases are a combination of these characteristics. Future works should combine these factors to carry out multifaceted diagnoses, which will help the development of smart agriculture.

ACKNOWLEDGMENT

The authors also gratefully acknowledge the insightful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] Z. Iqbal, M. A. Khan, M. Sharif, J. H. Shah, M. H. ur Rehman, and K. Javed, "An automated detection and classification of citrus plant diseases using image processing techniques: A review," *Comput. Electron. Agricult.*, vol. 153, pp. 12–32, Oct. 2018.

[2] Y. Saijo, E. P.-I. Loo, and S. Yasuda, "Pattern recognition receptors and signaling in plant-microbe interactions," *Plant J.*, vol. 93, no. 4, pp. 592–613, Feb. 2018.

[3] T. Pridmore, S. Tsaftaris, and H. Scharr, "Computer vision problems in plant phenotyping," in *Proc. CVPPP*, 2018, pp. 1–4.

[4] S. Cubero, N. Aleixos, E. Moltó, J. Gómez-Sanchis, and J. Blasco, "Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables," *Food Bioprocess Technol.*, vol. 4, no. 4, pp. 487–504, May 2011.

[5] L. Fei, G. Lu, W. Jia, S. Teng, and D. Zhang, "Feature extraction methods for palmprint recognition: A survey and evaluation," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 2, pp. 346–363, Feb. 2019.

[6] M. Abadi, "Tensorflow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Oper. Syst. Design Implement. (OSDI)*, 2016, pp. 265–283.

[7] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," in *Proc. Int. Conf. Comput. Commun. Control Autom.*, Feb. 2015, pp. 768–771.

[8] Y. Sasaki, T. Okamoto, K. IMOU, and T. TORII, "Automatic diagnosis of plant disease," *J. Jpn. Soc. Agricult. Mach.*, vol. 61, no. 2, pp. 119–126, 1999.

[9] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 391–407.

[10] M. El-Helly, S. El-Beltagy, and A. Rafea, "Image analysis based interface for diagnostic expert systems," in *Proc. Winter Int. Symp. Inf. Commun. Technol.*, Dublin, Ireland, 2004, pp. 1–6.

[11] M. Sammany and T. Medhat, "Dimensionality reduction using rough set approach for two neural networks-based applications," in *Proc. Int. Conf. Rough Sets Intell. Syst. Paradigms*. Cham, Switzerland: Springer, 2007, pp. 639–647.

[12] T. Baum, A. Navarro-Quezada, W. Knogge, D. Douchkov, P. Schweizer, and U. Seiffert, "HyphArea—Automated analysis of spatiotemporal fungal patterns," *J. Plant Physiol.*, vol. 168, no. 1, pp. 72–78, 2011.

[13] S. Ueda and M. Nakamura, "Edge detection device, image forming apparatus, and edge detection method," Google Patents 10 106 353, Oct. 23, 2018.

[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[15] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, Feb. 2017.

[16] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2117–2125.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556

[19] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.

[20] H. Lee, S. Eum, and H. Kwon, "Cross-domain CNN for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 3627–3630.

[21] X. Sun, S. Mu, Y. Xu, Z. Cao, and T. Su, "Image recognition of tea leaf diseases based on convolutional neural network," 2019, *arXiv:1901.02694*. [Online]. Available: http://arxiv.org/abs/1901.02694

[22] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[24] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm," *Procedia Comput. Sci.*, vol. 54, pp. 764–771, Jan. 2015.

[25] F. Zang and J.-s. Zhang, Kirchhain, "Softmax discriminant classifier," in *Proc. 3rd Int. Conf. Multimedia Inf. Netw. Secur.*, Nov. 2011, pp. 16–19.

[26] Y. Qin, S. He, Y. Zhao, and Y. Gong, "RoI pooling based fast multi-domain convolutional neural networks for visual tracking," in *Proc. 2nd Int. Conf. Artif. Intell. Ind. Eng. (AIIE)*, 2016, pp. 198–202.

[27] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 658–666.

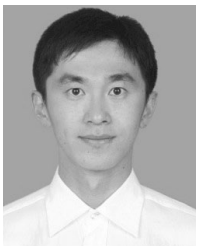
[28] M. S. Fisher and N. Arab, "Apparatus, system, and method for image normalization using a Gaussian residual of fit selection criteria," Google Patents 15 987 359, Sep. 20, 2018.

[29] M. Andrychowicz, "Learning to learn by gradient descent by gradient descent," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 3981–3989.

[30] Y. Wu, B. Jiang, and N. Lu, "A descriptor system approach for estimation of incipient faults with application to high-speed railway traction devices," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 10, pp. 2108–2118, Oct. 2019.

[31] Y. Wu, B. Jiang, and Y. Wang, "Incipient winding fault detection and diagnosis for squirrel-cage induction motors equipped on CRH trains," *ISA Trans.*, to be published, doi: 10.1016/j.isatra.2019.09.020.

[32] H. Nakahara, H. Yonekawa, and S. Sato, "An object detector based on multiscale sliding window search using a fully pipelined binarized CNN on an FPGA," in *Proc. Int. Conf. Field Program. Technol. (ICFPT)*, Dec. 2017, pp. 168–175.
[33] T. Zhou, M. Brown, N. Snavely, and D. G. Lowe, "Unsupervised learning of depth and ego-motion from video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1851–1858.
[34] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.
[35] A. V. Mityakov, V. K. Varankin, and Y. S. Tatarinov, "Application of modern architectures of deep neural networks for solving practical problems," in *Proc. IEEE Int. Conf. Soft Comput. Meas. (SCM)*, May 2017, pp. 389–390.
[36] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2009.
[37] Y. Liang, D. Niu, and W.-C. Hong, "Short term load forecasting based on feature extraction and improved general regression neural network model," *Energy*, vol. 166, pp. 653–663, Jan. 2019.
[38] H. Gao, X. Tao, X. Shen, and J. Jia, "Dynamic scene deblurring with parameter selective sharing and nested skip connections," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3848–3856.
[39] Y. Li, H. Huang, Q. Xie, L. Yao, and Q. Chen, "Research on a surface defect detection algorithm based on MobileNet-SSD," *Appl. Sci.*, vol. 8, no. 9, p. 1678, 2018.
[40] J. J. Almagro Armenteros, K. D. Tsirigos, C. K. Sønderby, T. N. Petersen, O. Winther, S. Brunak, G. von Heijne, and H. Nielsen, "Signal P 5.0 improves signal peptide predictions using deep neural networks," *Nature Biotechnol.*, vol. 37, no. 4, pp. 420–423, Apr. 2019.



YANG ZHANG received the Ph.D. degree from the School of Computer, Beijing Institute of Technology. He is currently an Associate Professor with the School of Information Science and Engineering, Hebei University of Science and Technology. His research interests focus on intelligent software, parallel programming model, and software refactoring for parallelism.

CHENGLONG SONG was born in 1995. He is currently pursuing the M.S. degree with the School of Information Science and Engineering, Hebei University of Science and Technology. His current research interests include computer vision, image processing, and deep learning.

DONGWEN ZHANG received the Ph.D. degree from the Beijing Institute of Technology. She is currently a Professor with the School of Information Science and Engineering, Hebei University of Science and Technology. Her research interests focus on parallel programming model and software refactoring for parallelism.