RESEARCH ON RECOGNITION OF OCCLUDED ORANGE FRUIT ON TREES BASED ON YOLOv4
/
基于 YOLOv4 模型的树上遮挡橙果的识别研究

LI Yan[1)], XIAO Liming[1)], LI Weiqi[1)], LI Hao[1)], LIU Jie[*1,2,3)] [1]
[1)] College of Engineering, Huazhong Agricultural University, Wuhan, Hubei/ China;
[2)] Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River, Ministry of Agriculture and Rural Affairs, Wuhan, Hubei/ China;
[3)] Citrus Mechanization Research Base, Ministry of Agriculture and Rural Affairs, Wuhan, Hubei/ China
Tel: 15623263208; E-mail: liujie@mail.hzau.edu.cn
Corresponding author: Jie Liu
DOI: https://doi.org/10.35633/inmateh-67-13

Keywords: *orange detection, image processing, deep learning, machine vision, YOLOv4*

ABSTRACT
For accurate recognition of orange fruit targets, a detection algorithm based on YOLOv4 was applied in this research. The results showed that AP (average precision) of YOLOv4 had reached 98.17%, 2.14% and 2.67% respectively higher than SSD and Faster RCNN while recognition rate of traditional image processing algorithms was merely 54.94%. Additionally, the extent of occlusion was proved to have obvious influences on the accuracy of orange detection. The accuracy on slight occlusion conditions appeared to be higher than that on serious occlusion conditions. Generally, YOLOv4 detection algorithm showed its feasibility and superiority on fruit detection in the complex natural environment.

摘要
本研究采用了一种基于 YOLOv4 的检测算法来准确识别橙果。结果表明，YOLOv4 的检测平均精度达到了 98.17%，分别比 SSD 和 Faster RCNN 提升了 2.14%和 2.67%，而传统图像处理算法的识别率仅为 54.94%。此外，研究表明遮挡程度对橙果检测的精度有明显影响。处于轻度遮挡条件下的橙果检测精度高于重度遮挡条件下。总体而言，YOLOv4 检测算法在复杂自然环境下的水果检测中表现出了可行性和优越性。

INTRODUCTION
As important cash crops, the production, processing and sales of fruits is of great significance to economic development (*Rocha et al., 2010*). With the continuous expansion of orchard planting area, the tasks of harvesting in orchards are becoming increasingly arduous, and cannot be completed relying only on manpower and simple tools (*Zhang et al., 2016*). On this occasion, picking robot is vigorously promoted and applied in agricultural production for its long working time and low cost in recent years, and the recognition and location of target fruits is the first step for picking robots to realize automatic harvest task (*Naranjo-Torres et al., 2020*). For picking robots, the speed, accuracy and adaptability to the surrounding environment of the visual recognition system have a great impact on the working efficiency and stability. However, complex natural environment with interference of light, occlusion of branches and leaves and overlap between fruits brings huge challenges to stable target recognition of robots (*Jin, 2020*). Machine vision technology offers the possibility of identifying fruits against the nature background, and mainly focus on traditional image processing algorithms, machine learning algorithms, and in recent years, on the fast-developing deep learning technology (*Rehman et al., 2018*).

Traditional image processing algorithms mostly make use of colour differences between target fruits and the background combined with different segmentation algorithms and edge detection methods to achieve target detection (*Scharr et al., 2016*). Based on the colour and shape of fruit and calyx, an image processing algorithm (*Al-Mallahi, 2019*) was developed to segment kiwifruits from the background, count the number of fruits in each cluster, and identify the edges of each target. Instead of drawing a separating line to determine borders of clustered fruits, a method (*Fu et al., 2015*) which used Canny operator for edge detection and ellipse Hough transformation for fruit recognition combined with artificial lighting were applied to harvest kiwifruits at night and this method had achieved the successful recognition rate of 88.3%.

---
[1] *Yan Li, M.S. Stud. Eng.; Liming Xiao, M.S. Stud. Eng.; Weiqi Li, M.S. Stud. Eng.; Hao Li, M.S. Stud. Eng.; Jie Liu, Prof. Ph.D. Eng.*

Machine learning is a common research hot spot in the field of artificial intelligence and pattern recognition applied in different fields (*Zhao et al., 2016*). In recent years, feasible algorithms are proposed combining the ordinary image processing algorithms with machine learning to achieve target information in agriculture (*Xu et al., 2021*). In order to achieve accurate detection of litchi fruits in the natural environment, a ripe litchi recognition method (*Yu et al., 2021*) based on RGB-D images was proposed. After excluding redundant image information outside the effective picking range by depth image segmentation, the random forest binary classification model was trained employing colour and texture features to recognize litchi fruits, which achieved a recognition accuracy of 89.92% for green litchis and 94.50% for red litchis. *Lu et al.* had proposed an innovative method to detect green citrus fruit in images of trees only by using texture and intensity distribution (*Lu et al., 2018*). Choosing local binary pattern (LBP) features around local intensity maxima of the green component of images taken in low natural light conditions with a flashlight as an input of an ensemble classifier-RUSBoost, the positive predictions were considered as candidates and fitted with Circular Hough Transform. The results showed that the proposed methods had achieved 80.4% true positive rate and 82.3% precision rate, and F-measure was 81.3%.

With the continuous development of research, deep learning technology has become increasingly popular in the research of image recognition. *Zhao et al.* used the YOLO model to locate apples under complex background (*Zhao et al., 2019*). The mAP (mean average precision) was 87.71%, and the frame rate of the detected video reached 60 frames/s. Sun et al. had applied improved Feature Pyramid Network to traditional Faster RCNN model to fuse the detailed bottom features and high-level semantic features to detect small-sized tomato organs (*Sun et al., 2020*).

As a kind of citrus fruit, orange has an extremely high field and economic values (*Li et al., 2019*). In this paper, the methods of fruit recognition were emphatically explored, including traditional image processing methods and deep learning algorithms, to search for the best way to detect orange fruits.

MATERIALS AND METHODS

Sample selection and data collection

In this experiment, image collecting was carried out in orchard of Huazhong agricultural university, WuHan, Hubei Province in China by iPhone 11.

Fully considering the complexity of the actual orchard in the image acquisition process, the images were collected in different environments including sunny, cloudy days and forward, reverse light. Additionally, the images were shot in multiple angles and in different distances. Through the above shooting method, a total of 723 images were collected which contained representative scenes such as overlap and occlusion during the actual growth of fruits.

In order to further improve the accuracy of the detection model, the sample data set needed to be expanded. Common image data enhancement methods including rotating, cropping, horizontally flipping, adding noise, changing brightness and enlarging were used to solve the problem of insufficient number of samples and promote the convergence and fitting of the model during training. Through data enhancement methods, 723 orange images were increased to 1445. After manually labelled through LabelImg software, the data set was divided into training set, validation set and test set randomly. Furthermore, in the test set, the average occlusion extent was used as a criterion to further divide the data set into two parts. Image set whose orange occlusion extent was less than 25% was defined as slight occlusion dataset represented by the letter A while image set whose orange occlusion was in the range of 25% to 50% was defined as serious occlusion dataset represented by the letter B.

Table 1

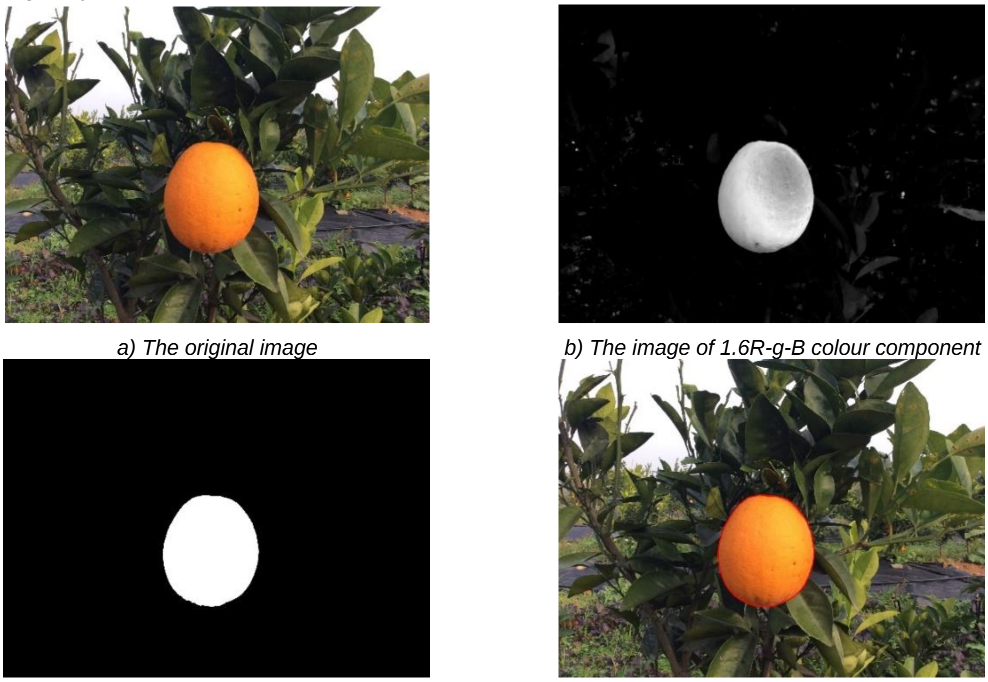
Category and quantity of orange data sets

Data set	**Training set**	**Validation set**	**Test set**		**Total**
			A (Slight occlusion)	**B (Serious occlusion)**	
Number of images	1036	116	206	87	1445

Traditional fruit recognition and segmentation algorithm

For the sake of making comparisons with YOLOv4, traditional image processing methods were applied in the whole test set to explore the detection effects. Traditional target recognition includes preprocessing, threshold segmentation, morphological operations and edge detection. After removing noise by median filtering to improve quality of original images, the next step was to distinguish target from background. At present, traditional fruit recognition is mainly realized through image processing methods on the basis of colour features, for the reason that existence of colour differences between the target and the background make it easy to extract object from the natural environment (*Benavides et al., 2020*).

The first step to extract colour features was to select an appropriate colour component. Chosen RGB model as the colour space, 1.6R-G-B colour component was found in which the gray level of orange was quite different from that of the background (Fig. 1b). Then, Otsu threshold segmentation method and binarization were performed on the obtained colour component map. Segmentation result after the above steps was that holes and burrs still remained. To solve the problem, 5 * 5 rectangular structural element was chosen as the kernel, 3 opening operations and 3 closing operations were carried out. In addition, hole filling method was used to remove small connected domains (Fig. 1c). That is, an area threshold was set, and the value of pixels in the small connected domain whose area was lower than this threshold were set to 0. Through the above methods, the outline of orange fruit could be segmented effectively. Finally, the object contours would be found and drawn (Fig. 1d).



a) The original image *b) The image of 1.6R-g-B colour component*

c) The image after morphological processing *d) The final result image*

Fig. 1 – The whole process of image processing

By the proposed image processing methods, a single fruit, multiple non-overlapping fruits and fruits slightly obscured by branches and leaves could be recognized well which could be seen in Fig. 2. However, for overlapping fruit, the overlapping connected domain was easy to be misjudged as a single fruit and could not be segmented accurately (Fig. 2f).

a) A single fruit *b) Recognition result of a single fruit*

c) Multiple non-overlapping fruits

d) Recognition result of multiple non-overlapping fruits

e) Overlapping fruits

f) Recognition result of overlapping fruits

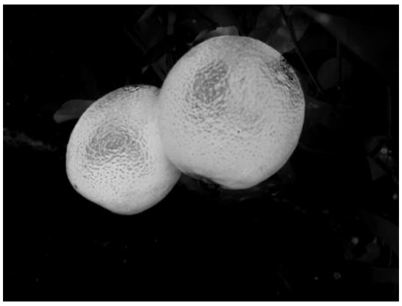
Fig. 2 – The effects of orange detection for oranges in different conditions

It could be seen from Fig. 3c that the overlapping oranges in the binary image were bonded together, if the two oranges were not separated, they would be automatically considered as the only one object in the image. Therefore, before the search of orange contours, the overlapped parts had to be separated. Watershed segmentation algorithm is one of the algorithms to find the segmentation line of overlapping objects (*Miao et al., 2016*).

In the process of watershed segmentation, taking the similarity with adjacent pixels as an important reference basis, the pixels with similar spatial position and similar gray value are connected to form a closed contour (*Zhang et al., 2018*). In this proposed orange recognition algorithm, Euclidean distance transformation was selected to achieve the foreground object (Fig. 3d). Euclidean distance transformation generally refers to the shortest distance from pixels to non-zero pixels in a binary image. The remaining area was obtained by subtracting the corrosion map from the expansion map, which was called as unknown area because it was uncertain whether it was the foreground or background (Fig. 3e). After obtaining these areas, seeds would be obtained. Watershed algorithm would take the contour passed in by the markers as the seed (that was, the so-called water injection point) to judge other pixels on the image according to the watershed algorithm rules, and delimit the regional ownership of each pixel until all pixels on the image were processed. The value at the boundary between regions was set to "- 1" to distinguish.

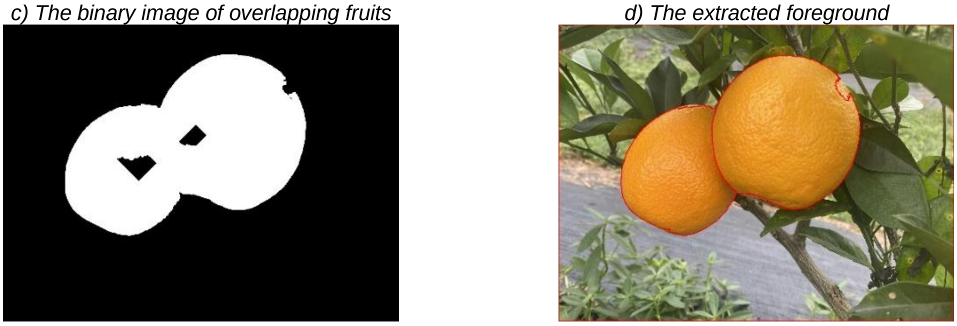
From the Fig. 3f, watershed algorithm was proved to be able to effectively segment overlapping fruits, but it also could discover that outline drawn was not completely consistent with the actual contour as a result of the shadows caused by light.

a) The original image of overlapping fruits



b) 1.6R-G-B colour component of overlapping fruits

c) The binary image of overlapping fruits d) The extracted foreground



e) The unknown area f) recognition result by watershed algorithm

Fig. 3 – Separating overlapping fruits based on watershed algorithm

Orange detection on the basis of deep Convolutional neural network

Nowadays, convolutional neural network has been increasingly popular in fruit detection due to its high accuracy and fast detection speed (*Fountsop et al., 2020*). Additionally, the neural network model established by training and learning had strong robustness and generalization ability, and had good experimental results in different complex field environments.

Deep convolution neural networks are mainly divided into two types: one is the method based on region proposal among which the most common algorithms are R-CNN, Fast RCNN, Faster RCNN (*Girshick et al., 2014; Girshick, 2015; Ren et al., 2017*), and the other is the method without region proposal, whose representative algorithms concludes YOLO and SSD (*Redmon et al., 2016; Liu et al., 2016*). In the method based on region proposal, the corresponding region proposal algorithm is used to generate the proposed target candidate region from the input image, and then all the candidate regions are sent to the classifier for classification. The whole process includes two steps: generation of candidate regions and classification. Therefore, this method is also called two-stage target detection. At the same time, the other method without region recommendations is called one-stage, for the reason that the achievement of target position and classification labels is completed by a single network.

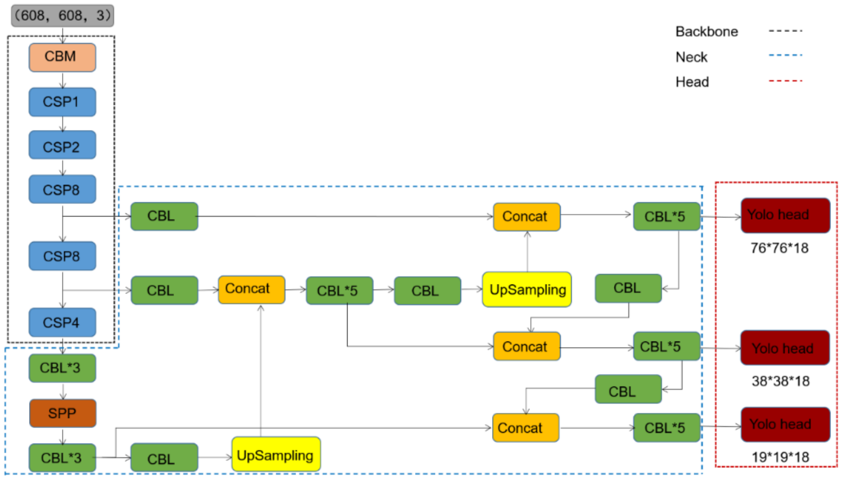
In this work, YOLOv4 convolution neural network was chosen to be applied into recognizing and locating oranges in the unstructured environment. By observation and analysis of the indexes of precision, recall, F1 and average precision, it was found that YOLOv4 had really outstanding performance, even under the condition of occlusion and overlap.

- ***Structure of YOLOv4 network***

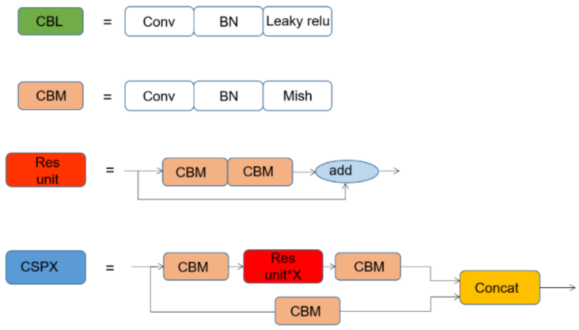
The network structure diagram of YOLOv4 is shown in Fig. 4. It is mainly composed of backbone, neck and head. The backbone of YOLOv4 is CSPDarknet53, playing function in feature extraction, and is proposed on the basis of Darknet53 with reference to CSPNet (*Cai et al., 2021*). Retaining the framework of Darknet53, the backbone of YOLOv3, YOLOv4 uses CSP mechanism to optimize the gradient back propagation path, and greatly reduces the amount of calculation on the premise of ensuring the accuracy. CSPDarknet is mainly composed of CSPX module. CSPX module means CSP module with X residual components inside. Its function is to divide the feature map into two parts. The first part is convoluted by CBM module and X residual units, and the second part is directly CONCAT combined with the first part.

The neck of YOLOv4 is a feature transfer network added between the backbone network and the prediction output layer. Through PANet and SPP, the eigenvalues extracted from the backbone network is sampled and aggregated to aggregation features at different scales.

The head of YOLOv4 stays the same with YOLOv3. The three branches of the neck output pass through convolution layers and then predict different sizes of targets concluding large, medium and small ones.



a) The network structure diagram of YOLOv4



b) The structure diagram of different modules in YOLOv4

Fig. 4 - The architecture of YOLOv4 network

- ***Training strategy based on Transfer Learning***

The experiment was based on Windows 10 operating system, with NVIDIA Ge-Force RTX 3070 Ti GPU and inter(R) Xeon(R) Platinum 8222CL CPU @ 3.00GHz. The model construction and training verification were implemented in Python language based on PyTorch deep learning framework, and the parallel computing framework was CUDA version 11.1.

In order to reduce the training time and obtain favourable training effect, the proposed approach adopted the idea of transfer learning. The weight achieved from training in VOC2007 and VOC2012 datasets was loaded as a pretrained weight to help the new model adjust parameters.

The whole training process contained two parts: freeze training and unfreeze training. Totally, the network was trained 100 epochs, and 50 epochs of freezing training and 50 epochs of unfreeze training were carried out successively. In the freezing stage, the backbone of the model was frozen, in other words, the feature extraction network did not change, only the network was fine-tuned, consequently the occupied memory was small. On the contrary, in the unfreeze stage, the backbone and all parameters of the network would change, resulting in large memory consumption. The picture size sent to the model was set as 608*608. The batch size and learning rate were set as 16 and 0.001 in freeze stage while those in unfreeze stage are 2 and 0.0001.

RESULTS

Detection results of the traditional image processing methods

The results showed that the recognition rate of the traditional image processing algorithms on the whole test set was 54.94%, and results on test A with slightly sheltered oranges and test B with seriously sheltered oranges were 59.72% and 46.78% respectively. The low recognition rate could be explained by the fact that most of the pictures in the test set were not in a single scene, but in a variety of scenes with combinations of occlusion, overlap and light, and any of them would bring difficulties to accurate recognition of orange.

To some extent, traditional object recognition and segmentation methods could distinguish oranges from complex field environment. However, there was denying that its poor universality and susceptibility would bring extremely great difficulties for picking robots to complete field picking task.

Detection results of the deep learning algorithms

In order to test the performance of the proposed YOLOv4 algorithm in orange detection, fruit detection experiments were carried out on test set A with slightly sheltered oranges, test set B with heavily sheltered and test set A + B respectively.

Through performance in different test sets, the effects of extent of occlusion and overlap to the trained model would be exposed. At the same time, commonly used detection networks like SSD and Faster RCNN were also applied into orange detection to make a comparison.

In the process of detection, test images were input into the trained network for orange target position regression. Precision (P), recall (R), F1 and average precision (AP) were selected as evaluation criteria.

- ***Effects of YOLOv4 on different test sets***

The occlusion of branches and leaves to oranges and the overlap between oranges would bring a great impact on the detection accuracy of the model. In this paper, the test set was divided into two parts according to the degree of occlusion, and different test sets were used as variables to verify the performance of the trained model. The results were shown in Table 2. From Table 2, it was easy to find that YOLOv4 generally worked extremely well in whatever test sets. AP had reach 98.17% in the A+B image set. In slight occlusion images, AP had reached 99.04 %, more than this, P, R and F1 all had passed 97.60%, which convincingly proved that YOLOv4 model could extract oranges not only accurately but also comprehensively from complicated natural environments. In serious occlusion and overlap circumstances, the performance of the model was obviously reduced, all evaluation indicators had decreased by 2% to 6%. Nevertheless, P, R and AP still remained over 92%. These results demonstrated convincingly that YOLOv4 still worked well in occlusion and overlap situations.

Table 2

Experimental results of the YOLOv4 algorithm in different test sets

Evaluation index	Test set		
	A	B	A+B
Precision (%)	98.40	92.72	96.26
Recall (%)	97.62	94.92	96.62
F1 (%)	98	94	96
Average Precision (%)	99.04	96.61	98.17

a) The original image in test set A

b) Result of the image in test set A

c) The original image in test set B

d) Result of the image in test set B

Fig. 5 – Detection effect of YOLOv4 in different test sets

- ***Results of detection based on different networks***

Selected as contrast models, SSD and Faster RCNN were applied to orange detection in this paper. The results of detection were represented in Table 3.

From Table 3, the performances of Faster RCNN and SSD have a significant decline on all test sets compared with that of YOLOv4, especially P of Faster RCNN and R of SSD. In test set A, P of Faster RCNN has declined nearly 17%, and R of SSD has decreased approximately 6%. This downward trend is more obvious in dataset B. P of Faster RCNN has declined to less than 70% and R of SSD has declined to less than 80%. Through the analysis of the evaluation indexes in the table, both the precision and recall of YOLOv4 were higher than the other two models, especially in the occlusion and overlap scenes. Consequently, generally speaking, YOLOv4 was the best of the three models to effectively identify orange fruits in natural environment.

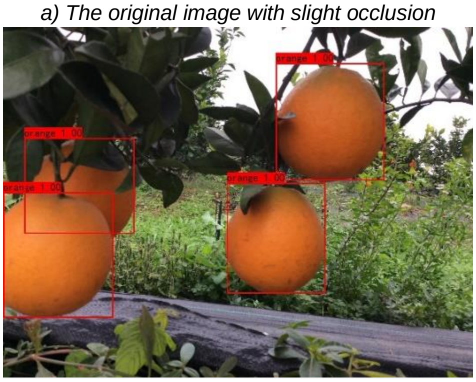
Table 3

Comparison of detection results with different algorithms

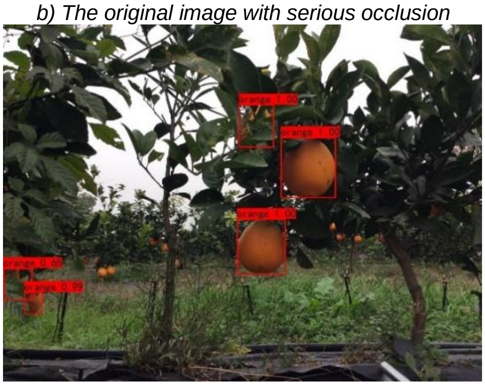
Evaluation index	YOLOv4			SSD			Faster RCNN		
	A	B	A+B	A	B	A+B	A	B	A+B
Precision (%)	98.40	92.72	96.26	96.45	95.00	95.97	81.50	67.39	75.74
Recall (%)	97.62	94.92	96.62	91.67	77.29	86.36	97.02	94.58	96.12
F1(%)	98	94	96	94	85	91	89	79	85
Average Precision (%)	99.04	96.61	98.17	96.90	94.25	96.03	96.75	93.04	95.50

a) The original image with slight occlusion

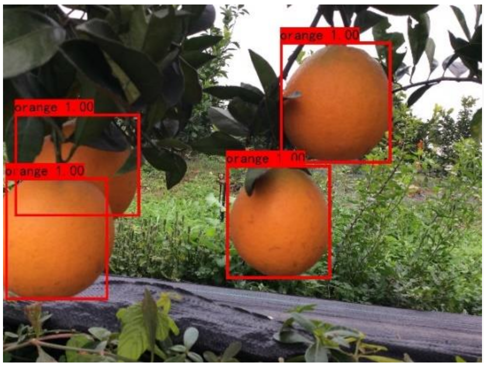
b) The original image with serious occlusion



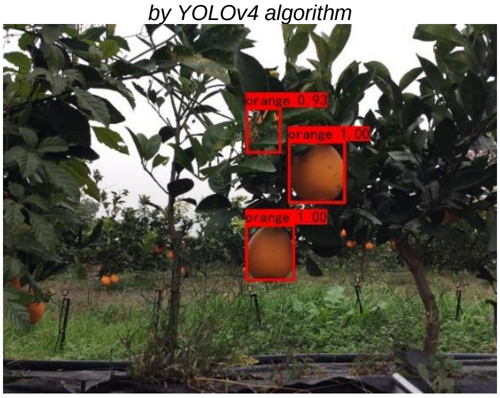
c) Detection result in slight occlusion image by YOLOv4 algorithm



d) Detection result in serious occlusion image by YOLOv4 algorithm



e) Detection result in slight occlusion image by SSD algorithm



f) Detection result in serious occlusion image *by SSD algorithm*

g) Detection result in slight occlusion image by Faster -RCNN algorithm

h) Detection result in serious occlusion image by Faster- RCNN algorithm

Fig. 6 – Detection effect of 4 algorithms on oranges with different occlusion

CONCLUSIONS

Occlusion of branches and leaves and the overlap between fruits during the natural growth of oranges brings great difficulties to the recognition of target fruits. To solve the problem, in this paper, traditional image processing algorithms and deep convolution neural network were used to segment and extract orange fruit in natural environment.

(1) Image processing methods including the selection of RGB model, application of Otsu threshold algorithm, implementation of morphological operation and detection of edge could effectively distinguish the slightly sheltered and non-overlapping fruits. For overlapping fruits, the watershed method was used to segment the overlapping fruit boundary accurately. However, this method was greatly affected by illumination and occlusion and has poor stability, which has great limitations in practical orchard application.

(2) Detection model was trained by using deep convolution neural network of YOLOv4. The detection effect of YOLOv4 was analysed under different occlusion conditions and it was compared with other neural networks like Faster RCNN and SSD. Although occlusion had a certain effect on detection accuracy, YOLOv4 had still been proven to be the relatively appropriate model to detect orange fruit in complex natural environments and could effectively distinguish target from the orchard and provide accurate position information for subsequent picking task.

ACKNOWLEDGEMENT

The authors appreciate the support from Fundamental Research Funds for the Central Universities of China (2662020GXPY011), China Agriculture (Citrus) Research System (CARS-27) and National Key R&D Program of China (2018YFD0701105-2).

REFERENCES

[1] Al-Mallahi, A. (2019). A novel image processing algorithm to separate linearly clustered kiwifruits. *Biosystems Engineering*, Vol. 183, pp. 184-195, England.

[2] Benavides, M., Cantón-Garbín, M., Sánchez-Molina, J., Rodríguez, F. (2020). Automatic Tomato and Peduncle Location System Based on Computer Vision for Use in Robotized Harvesting. *Applied Sciences*, 10(17): 5887, Switzerland.

[3] Cai, S., Sun, Z., Liu, H., Wu, H., Zhuang, Z. (2021). Real-time detection methodology for obstacles in orchards using improved YOLOv4 (基于改进型 YOLOv4 的果园障碍物实时检测方法). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 37 (2), pp: 36-43, Beijing/China.

[4] Fountsop, A. N., Fendji, J., Atemkeng, M. (2020). Deep Learning Models Compression for Agricultural Plants. *Applied Sciences*, 10 (19). 6866, Switzerland.

[5] Fu, L., Wang, B., Cui, Y., Su, S., Gejima, Y., Kobayashi, T. (2015). Kiwifruit recognition at nighttime using artificial lighting based on machine vision. *International Journal of Agricultural and Biological Engineering*, Vol. 4, pp. 52-59, Beijing/ China.

[6] Girshick, R., Donahue, J., Darrell, T., Malik J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation tech report. *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2014, pp. 580-587, Columbus/ USA.

[7] Girshick, R. (2015). Fast R-CNN. *IEEE International Conference on Computer Vision*, Vol. 2015, pp. 1440-1448, USA.
[8] Ren, S., He, K., Girshick, R., Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 39(6), pp.1137-1149, USA.
[9] Jin, Y. J. (2020). Recognition Technology of Agricultural Picking Robot Based on Image Detection Technology. *INMATEH - Agricultural Engineering*, Vol. 62(3), pp. 191-200, Romania.
[10] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. *European Conference on Computer Vision*, Vol. 2016, pp. 21-37, Amsterdam/ Netherlands.
[11] Li, S. J., Hu, D. Y., Gao, S. M., Lin, J. H., An, X. S., Zhu, M. (2019). Real-time classification and detection of citrus based on improved single short MultiBox detector (基于改进 SSD 的柑橘实时分类检测). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 35 (24), pp. 307-313, Beijing/ China.
[12] Lu, J., Lee, W. S., Hao, G., Hu, X. W. (2018). Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis. *Biosystems Engineering*, Vol. 171, pp. 78–90, England.
[13] Miao, Z. H., Shen, Y. C., Wang, X. H., Zhou, X. F., Liu, C. L. (2016). Image Recognition Algorithm and Experiment of Overlapped Fruits in Natural Environment (自然环境下重叠果实图像识别算法与试验). *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 47 (006), pp. 21-26, Beijing/ China.
[14] Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R. J., Valenzuela, A. (2020). A Review of Convolutional Neural Network Applied to Fruit Image Processing. *Applied Sciences*, Vol. 10(10). 3443, Switzerland.
[15] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2016, pp.779-788, Las Vegas/USA.
[16] Rocha, A., Hauagge, D. C., Wainer, J., Goldenstein, S. (2010). Automatic fruit and vegetable classification from images. *Computers and Electronics in Agriculture*, Vol. 70 (1), pp. 96-104, England.
[17] Rehman, T. U., Mahmud, M. S., Chang, Y. K., Jin, J., Shin, J. (2018). Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Computers and Electronics in Agriculture*, Vol. 156, pp. 585-605, England.
[18] Scharr, H., Minervini, M., French, A. P., Klukas, C., Kramer, D. M., Liu, X. M., Luengo, I., Pape, J., Polder G., Vukadinovic, D., Yin, X., Tsaftaris, S. A. (2016). Leaf segmentation in plant phenotyping: a collation study. *Machine Vision and Applications*, Vol. 27(4), pp. 585-606, USA.
[19] Sun, J., He, X. F., Wu, M. M., Wu, X. H., Shen, J. F., Lu, B. (2020). Detection of tomato organs based on convolutional neural network under the overlap and occlusion backgrounds. *Machine Vision and Applications*, Vol. 31 (5). https://doi.org/10.1007/s00138-020-01081-6. USA.
[20] Xu, Z. H., Song, H. Y., Wu, Z. M., Xu, Z. F., Wang, S. F. (2021). Research on crop information extraction of agricultural UAV images based on blind image deblurring technology and SVM. *INMATEH - Agricultural Engineering*, Vol. 64, pp. 33-42, Romania.
[21] Yu, L. Y., Xiong, J. T., Fang, X. Q., Yang, Z. G., Chen, Y. Q., Lin, X. Y., Chen, S. F. (2021). A litchi fruit recognition method in a natural environment using RGB-D images. *Biosystems Engineering*, Vol. 204 (1), pp. 50-63, England.
[22] Zhao, Y. S., Gong, L., Huang, Y. X., Liu, C. L. (2016). A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture*, Vol. 127, pp. 311-323, England.
[23] Zhang, Z., Heinemann, P. H., Liu, J., Baugher, T. A., Schupp, J. R. (2016). The Development of Mechanical Apple Harvesting Technology: A Review. *Transactions of the ASABE*, Vol. 59 (5), pp. 1165-1180, USA.
[24] Zhang, J. H., Han, S. Q., Zhai, Z. F., Kong, F. T., Feng, X., Wu, J. Z. (2018). Improved adaptive watershed method for segmentation of cotton leaf adhesion lesions (改进自适应分水岭方法分割棉花叶部粘连病斑). *Transactions of the ASABE*, Vol. 034 (024), pp. 165-174, USA.
[25] Zhao, D. A., Wu, R. D., Liu, X. Y., Zhao, Y. Y. (2019). Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background (基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位). *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 35 (03), pp. 172-181, Beijing/China.