IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 30, 2020, accepted November 13, 2020, date of publication November 25, 2020, date of current version December 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3040423

Apple Detection in Natural Environment Using Deep Learning Algorithms

GUANTAO XUAN[1], CHONG GAO[1], YUANYUAN SHAO[1,2], MENG ZHANG[2], YONGXIAN WANG[1], JINGRUN ZHONG[3], QINGGUO LI[4], AND HONGXING PENG[3]

[1]College of Mechanical and Electrical Engineering, Shandong Agricultural University, Tai'an 271018, China
[2]Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing 210014, China
[3]College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China
[4]Tai'an Jiacheng Technology Company Ltd., Tai'an 271018, China

Corresponding authors: Yuanyuan Shao (syy007@sdau.edu.cn) and Hongxing Peng (xyphx@scau.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 31701325, Grant 31671632, and Grant 61863011; in part by the Green Farming and Mechanical Innovation Team of Fruit Harvesting under Soil, Ningbo Public Welfare Science and Technology Project, under Grant 202002N3011; and in part by the Applied Science and Technology Special Fund Project, Meizhou, in 2019, under Grant 2019B0201005.

ABSTRACT It is a challenging problem to detect the apple in natural environment using traditional object recognition algorithms due to occlusion, fluctuating illumination and complex backgrounds. Deep learning methods for object detection make impressive progress, which can automatically extract the number, pixel position, size and other features of apples from the images. In this paper, four deep learning recognition models, Faster RCNN based on AlexNet, Faster RCNN based on ResNet101, YOLOv3 based on DarkNet53 and improved YOLOv3 were employed to carry out recognition experiments on red and green apple under three illumination and two image sharpness conditions, with the transfer learning to accelerate the training process. The results showed that improved YOLOv3 model had the best recognition effect among the four detection models. F1 value of red apple recognition was 95.0%, 94.6% and 94.1% for normal, insufficient and excessive illumination, respectively, and F1 value of green apple recognition was 94.9%, 94.0% and 91.1%. There were F1 value of 92.8% and 92.1% for red and green apple recognition in blurred images, respectively. Moreover, improved YOLOv3 algorithm still had the better performance for occlusion, spot, overlap and incomplete apples, with a recognition recall rate higher than 88.5%. It can be concluded that improved YOLOv3 algorithm can provide a more efficient way for apple detection in natural environment.

INDEX TERMS Deep learning, image process, apple detection, faster RCNN, YOLOv3.

I. INTRODUCTION

Apple is a kind of fruit consumed and grown worldwide because of its delicacy and high nutrition [1], [2]. In 2017, China produced more than 49 percent of the world's apples, ranking first in the world [3]. However, apple harvest depends extensively on manual operation with the high labour intensity and the low efficiency, which has restricted the development of apple industry in China [4]–[6]. With the rapid development of smart farming and large-scale planting in recent years, apple harvesting robots have provided a mean to reduce harvest costs while improving labour productivity [7]–[9]. Therefore, it is of great significance to recognize the apples rapidly and accurately for real-time automatic harvesting [10].

Numerous studies have been reported on the apple detection. Peng *et al.* [11] used the SVM classifier to classify and identify fruits based on the extracted feature vectors. Sengupta and Lee [12] developed an algorithm including Hough circle detection, texture classification based on SVM and other false positive removal techniques to detect immature citrus in green canopy under natural conditions. Sun *et al.* [13] designed an apple target segmentation method by fusing fuzzy set theory and the manifold ranking algorithm to carry out the recognition of green apples in similar background areas. Lv *et al.* [14] conducted the detection of apple fruit by OTSU dynamic threshold segmentation method, edge detection and improved RHT transformation method. Ji *et al.* [15] developed a new classification algorithm based

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

on SVM to detect about 89% of fruits with an average recognition time of 352 ms. Although these methods can be used for apple recognition, their overall performances are still far from satisfactory due to changes in illumination, branches and leaves masking.

Recent developments in deep learning technique have proved its powerful ability for object recognition under the complicated background [16]–[19]. Dias *et al.* [20] proposed CNN + SVM method for apple flower detection, which could also achieve accurate identification even in scenarios of different flower species and illumination conditions. Fu *et al.* [21] developed a kiwifruit detection system for field images by using Faster RCNN with ZFNet, which had a good robustness against the subjectivity and limitation of the features selected manually. Xiong *et al.* [22] applied faster RCNN to detect the green citrus in different quantities, sizes and illumination angles. Peng *et al.* [23] proposed an improved SSD fruit detection model to accurately detect various types of fruits in the natural environment. Tian *et al.* [24] proposed an improved YOLO-V3 model for detecting apples during different growth stages in orchards with fluctuating illumination, complex backgrounds, overlapping apples, and branches and leaves. However, current research on apple recognition using deep learning technique is at its early stage, and there are few reports on apple classification detection under different algorithms and different conditions.

The main object of this study was to explore the feasibility of applying deep learning technique for apple recognition. Images were first acquired from two cultivars of apple in different illumination conditions. Then four deep learning methods were developed to detect the apples under various conditions. The ability of different models to recognize apples was analyzed and evaluated in the end.

II. MATERIALS AND METHODS

A. IMAGE ACQUISITION

There were two cultivars of apples, ‘Honeycrisp’ with red colour and ‘Fuji’ with green colour, growing at a commercial apple orchard in Prosser, Washington State, USA. Image acquisition was conducted in different illumination conditions covering the sunny, cloudy and different times of the day in October, 2017. All the images were acquired using smartphone camera (iphone 7 Plus, Apple Inc., Cupertino, CA, USA) with 4032 × 3024 pixels. There were 849 apple images collected from various viewing direction of camera including 609 red apple images and 240 green apple images.

B. DATASET PRODUCTION

Images were divided into training dataset, validation dataset and test dataset. Original images were selected to produce the test dataset. For red apples, the test dataset had 200 images selected equally from original images under normal, excessive, and insufficient illumination, and blurred images. Similarly, there were 80 images selected as the test dataset for green apples. Figure 1 showed some red and green apple images in different conditions. Then the remaining images were used to augment for training dataset and validation dataset.

FIGURE 1. Red and green apples in (a) normal, (b) excessive, (c) Insufficient, and (d) blurred image.

FIGURE 2. Incomplete/small fruit sample annotation.

C. IMAGE AUGMENTATION

To recognize the apples from the uneven sunlight and complex background, original images were augmented and expanded to get enough data for extracting the features effectively and avoiding overfitting, including brightness enhancement and reduction, contrast enhancement and reduction, saturation enhancement and attenuation, and noise points addition. Image augmentation was achieved in Python 2.7 by writing an image processor. As a result of the image augmentation, red and green apple images were expanded to 3096 and 1280 images, respectively. Table 1 illustrated the number of images created by various augmentation methods. Then these images were randomly split into training and validation sets in a ratio of 5:1.

D. MAN-MADE ANNOTATION OF SUPERVISED LEARNING

The categories and the location of samples were labeled manually in term of bounding boxes according to the principle of minimum tangential rectangle. In case of occlusion from leaves and other apples, annotation was applied to cover the whole apple, even for some incomplete or small fruit. Figure 2 showed an example of image annotation, including some incomplete or small fruit. The image annotation software used in this study was LabelImg, which saved the annotation file in the format of ‘xml’.

E. DEEP LEARNING ALGORITHM FOR APPLE DETECTION

Compared with traditional image recognition methods, target detection and recognition technology based on deep learning

TABLE 1. **Number of apple images created by various augmentation methods.**

Apple	OI	BE	BR	CE	CR	SE	SA	NI	Total
Red	387	387	387	387	387	387	387	387	3096
Green	160	160	160	160	160	160	160	160	1280

Note: OI: original image; BE: brightness enhancement; BR: brightness reduction; CE: contrast enhancement; CR: contrast reduction; SE: saturation enhancement; SA: saturation attenuation; NI: noise increase.

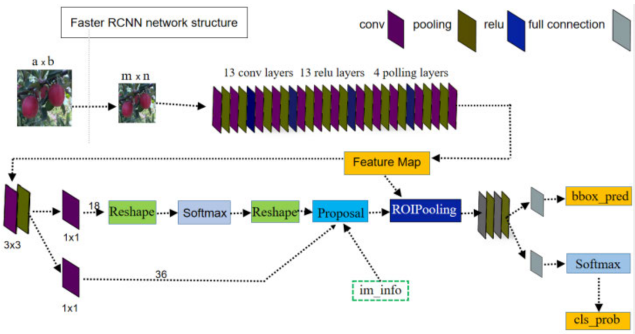


FIGURE 3. **Image processing procedures of Faster RCNN for apple recognition.**

has the advantage of learning features through deep convolution neural network, direct regression through end-to-end training, high accuracy and real-time performance, and is good at discovering complex structures in high-dimensional data. The more data is used to train deep learning model, the stronger robustness and generalization of the model will be.

1) FASTER RCNN ALGORITHM FOR APPLE DETECTION

Faster RCNN evolves from RCNN and Fast RCNN [25]–[27]. It innovatively uses the region proposal networks to integrate feature extraction, candidate box selection, border regression and classification into one network [28]. In this experiment, Faster RCNN structure was mainly composed of convolution layers, region proposal networks (RPN), ROI pooling layers, and classification and regression layers. A feature map was initially produced by AlexNet and ResNet101, respectively, and non-maximum suppression was applied to eliminate bounding box. Figure 3 illustrated image processing procedures of Faster RCNN for apple recognition.

The loss function of the RPN is defined as follows:

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & |x| \leq 1 \\ |x| - 0.5 & otherwise \end{cases} \tag{1}$$

$$L(\{P_i\}\{T_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(P_i, P_i^*) + \sigma \frac{1}{N_{reg}} \sum_i P_i^* L_{reg}(T_i, T_i^*) \tag{2}$$

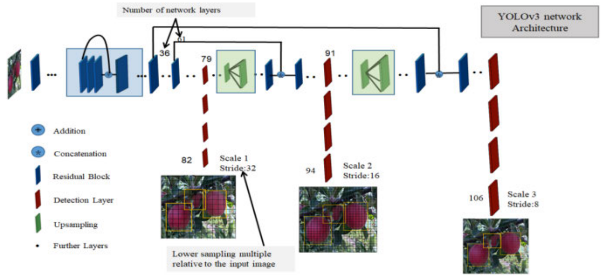


FIGURE 4. **Architecture of YOLOv3 for apple detection.**

where, i is the index of an anchor in a mini-batch and P_i is the predicted probability of anchor i being an object. The ground-truth label P_i^* is 1 if the anchor is positive, and is 0 if the anchor is negative. T_i is a vector representing the 4 parameterized coordinates of the predicted bounding box, T_i^* is that of the ground-truth box associated with a positive anchor. The outputs of the *cls* and *reg* layers consist of $\{P_i\}$ and $\{T_i\}$ respectively. The two terms are normalized with N_{cls} and N_{reg}, and a balancing weight σ. The classification loss L_{cls} is log loss over two classes (apple vs. not apple). For the regression loss, we use $L_{reg}(T_i, T_i^*) = R(T_i - T_i^*)$ where R is the robust loss function ($smooth_{L1}$) defined in equation (1).

2) YOLOV3 ALGORITHM FOR APPLE DETECTION

YOLO is the single-stage detector to deal with object detection by direct regression from input images to class probabilities and bounding box coordinates [29]. YOLOv3 detection algorithm discarded the pooling layer and the fully connected layer, and used DarkNet53 for multi-scale prediction of the apple feature instead of Softmax [30]. There were three feature maps adopted to detect large, small and medium-sized apples in Figure 4.

The loss calculation of YOLO during training is as follows:

$$\begin{aligned} \text{Loss Function} = {} & \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} \ell_{ij}^{\text{obj}} \times \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\ & + \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} \ell_{ij}^{\text{obj}} \times \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i}\right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i}\right)^2 \right] \\ & + \sum_{i=0}^{s^2} \sum_{j=0}^{B} \ell_{ij}^{\text{obj}} \left(C_i - \hat{C}_i\right)^2 \\ & + \lambda_{\text{noobj}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} \ell_{ij}^{\text{nobj}} \left(C_i - \hat{C}_i\right)^2 \\ & + \sum_{i=0}^{s^2} \ell_i^{\text{obj}} \sum_{c \epsilon \text{classes}} \left(p_i(c) - \hat{p}_i(c)\right)^2 \end{aligned} \tag{3}$$

YOLO network adopts square Loss, where S is the number of partitioned grids, B is the number of bounding boxes predicted by each grid cell, C is the number of categories of bounding boxes, λ is the loss coefficient, ℓ_{ij}^{obj} is the judgment

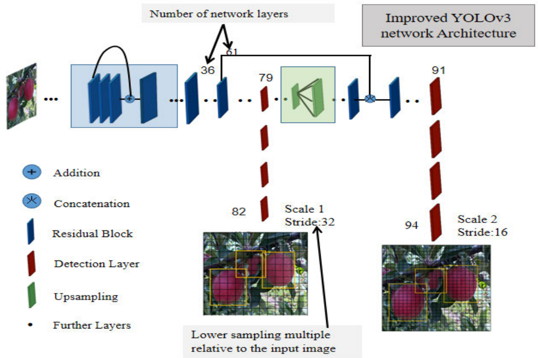


FIGURE 5. Architecture of Improved YOLOv3 for apple detection.

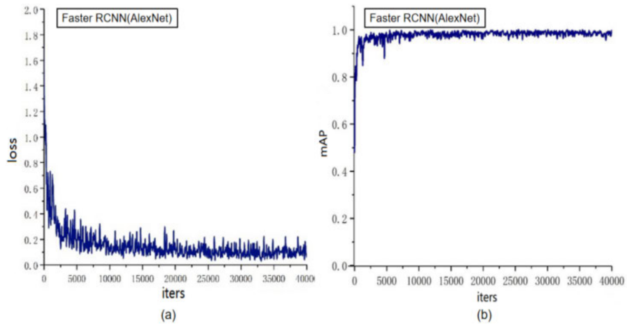


FIGURE 6. Training loss curve (a) and mAP score curve (b) of AlexNet+Faster RCNN algorithm.

of whether the jth bounding box in the ith grid is responsible for this object.

3) THE IMPROVED YOLOV3 ALGORITHM

The classic yolov3 detection network uses darknet-53 as a feature extraction network. It contains 53 convolution layers, which are powerful but have a deep network layer and a large computational cost. To improve the speed of training, reduce the redundancy of the network and misrecognition of small objects in the background, an improved yolov3 model was proposed by removing the full connected layer of darknet53 and the detection branch of 8 times down-sampling scale according to the typical requirements of apple identification. The improved yolov3 network framework was shown in figure 5.

III. RESULT AND ANALYSIS

A. EXPERIMENTAL ENVIRONMENT

Four detection models, AlexNet + Faster RCNN, ResNet101 + Faster RCNN, DarkNet53 + YOLOv3 and the improved YOLOv3 were used for apple recognition. They were trained and tested in NVIDIA GeForce GTX1080Ti platform, and performance of each model was improved greatly by CUDA toolkit 9.0 and CUDNN library 7.0. Images were analyzed using MATLAB 2018b and Python 2.7.

B. PARAMETER SETTING AND MODEL TRAINING

Because it is laborious to re-collect the training data and rebuild the models, transfer learning method is used to

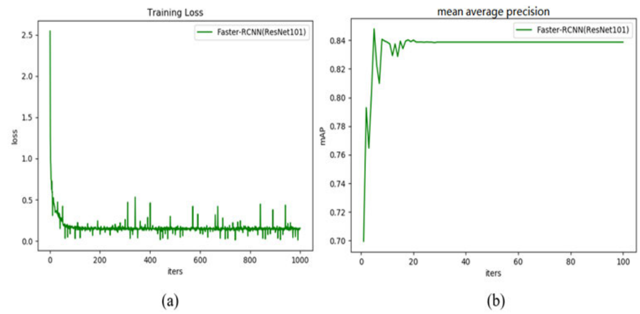


FIGURE 7. Training loss curve (a) and mAP score curve (b) of ResNet101+Faster RCNN algorithm.

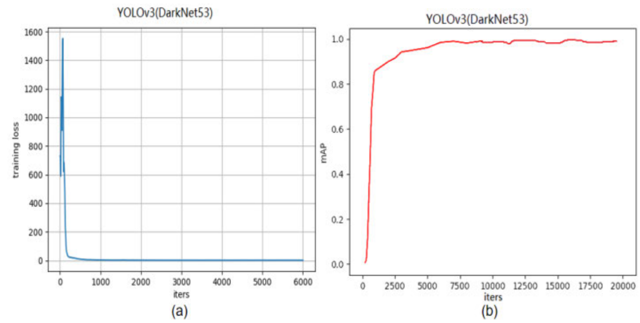


FIGURE 8. Training loss curve (a) and mAP score curve (b) of DarkNet53+YOLOv3 algorithm.

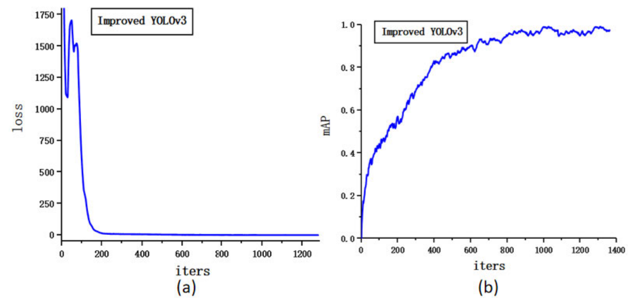


FIGURE 9. Training loss curve (a) and mAP score curve (b) of improved YOLOv3 algorithm.

accelerate the training process [31], [32]. Table 2 showed the initial parameter settings of the models. The performance of each model was assessed during training process in terms of mean average precision (mAP) and training loss.

Figure 6-9 showed the curves of training loss and mAP of the four models, respectively. AlexNet + Faster RCNN model fitted and was nearly stable when loss value dropped from 1.7319 to 0.2 after 3000 iterations, and the maximum value of mAP was 98%. ResNet101 + Faster RCNN model was nearly stable when loss value decreased from 2.5 to 0.15 after 200 iterations, and the maximum value of mAP was 85%. DarkNet53 + YOLOv3 model reached stability when loss value tended to 0 from 600 after 100 iterations, and the maximum value of mAP was 99%. The improved YOLOv3 model was nearly stable when loss value decreased from 1750 to 0 after 200 iterations, and the maximum value of mAP was 99.4%. Therefore, AlexNet + Faster RCNN model had the slowest fitting process and took a long time

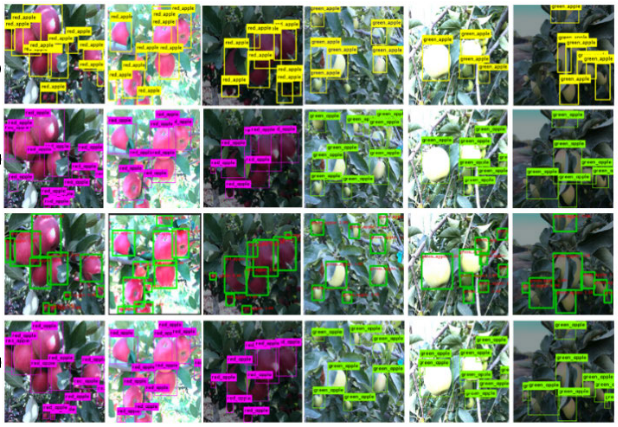
(a)
(b)
(c)
(d)


FIGURE 10. Examples of apple detection in different illumination conditions by (a) AlexNet + Faster RCNN, (b) ResNet101 + Faster RCNN, (c) DarkNet53 + YOLOv3, and (d) improved YOLOv3 algorithms.

to train. While ResNet101 + Faster RCNN had the faster fitting speed, but its mAP value was low, and the training effect was slightly worse than the other three models. The improved YOLOv3 model had the fastest fitting speed and good training effect, which performed well in the training process.

C. EVALUATION

Four indicators, i.e., precision (P), recall (R), error (E), F1 score were applied to evaluate apple recognition performance of these models. Among them, F1 score took both accuracy and recall into account. The larger F1 score was, the better performance of the model was. These indicators were defined as follows.

$$P = \frac{TP}{TP + FP} \tag{4}$$

$$R = \frac{TP}{TP + FN} \tag{5}$$

$$E = \frac{FP}{TP + FP} \tag{6}$$

$$F1 = \frac{2PR}{P + R} \tag{7}$$

where, TP, FP, and FN represent the true positives, false positives, and false negatives, respectively.

D. COMPARISON EXPERIMENTS OF DIFFERENT MODELS

1) APPLE DETECTION IN DIFFERENT ILLUMINATION CONDITIONS

Illumination was one of the key factors to capture objects in computer vision. Figure 10 presented a few test images for apple recognition in different illumination conditions using four algorithms. Illumination intensity in red and green apple images from left to right was normal, excessive, and insufficient, respectively. It could be observed from these images that red and green apples were recognized well.

The recognition results were further described in detail in Table 3. As shown in Table 3, improved YOLOv3 algorithms obtained the best performance of apple recognition

TABLE 2. Initial parameter setting of models.

Algorithm	Image size	Batch	Momentum	Learning Rate	Decay
Faster RCNN	416 x 416	1	0.9	0.0001	0.0005
YOLOv3	416 x 416	64	0.9	0.001	0.0001
Improved YOLOv3	416 x 416	64	0.9	0.001	0.0001

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


FIGURE 11. Examples of apple recognition in clear and blurred images by (a) AlexNet + Faster RCNN, (b) ResNet101 + Faster RCNN, (c) DarkNet53 + YOLOv3, and (d) improved YOLOv3 algorithms.

with higher precision, higher recall, lower error and higher F1 score in each illumination condition, followed by DarkNet53 + YOLOv3, ResNet101 + Faster RCNN and then AlexNet + Faster RCNN. Green apple recognition performance was slightly worse in comparison with red apple, with slightly lower accuracy rate and higher error. It might be because green apples had similarity in color to leaves and branches. As to all, the best recognition results occurred in the red apple in normal illumination conditions, and the performance of green apple recognition was worst in excessive illumination, with F1 scores of 83.4%, 88.4%, 89.9% and 91.1% for AlexNet + Faster RCNN,ResNet101 + Faster RCNN, DarkNet53 + YOLOv3 and improved YOLOv3 algorithms, respectively. This could be attributed to the fact that the colour or texture features were not obvious or even concealed due to the spot or reflection area formed on the apple surface by the excessive illumination.

2) APPLE DETECTION IN BLURRED IMAGES

There were some blurred images captured from long-range or in bad weather such as fog and strong wind. As a comparison with previous images in normal illumination, four methods trained were adopted to recognize the apples in blurred test images from the same condition, as shown in Figure 11. The left was clear and the right was blurred for red and green apple images.

TABLE 3. Apple recognition results in different illumination conditions.

Detection Model	Apple Type	Illumination Intensity	TP	FP	FN	P/%	R/%	E/%	F1/%
AlexNet + Faster RCNN	Red	Normal	297	15	42	95.2	87.6	4.8	91.2
		Insufficient	257	20	42	92.7	86	7.2	89.2
		Excessive	296	22	46	93.1	86.5	6.9	89.6
	Green	Normal	115	8	19	93.5	85.8	6.5	89.4
		Insufficient	107	11	23	90.6	82.3	9.3	86.2
		Excessive	119	10	37	92.2	76.3	7.7	83.4
ResNet101 + Faster RCNN	Red	Normal	307	7	32	97.8	90.6	2.2	94.1
		Insufficient	271	14	28	95.1	90.6	4.9	92.8
		Excessive	302	8	40	97.4	88.3	2.6	92.6
	Green	Normal	120	3	14	97.5	89.6	2.4	93.4
		Insufficient	115	6	15	95.0	88.4	5.0	91.6
			126	3	30	97.7	80.8	2.3	88.4
DarkNet53 + YOLOv3	Red	Normal	310	6	29	98.1	91.4	1.9	94.6
		Insufficient	278	11	21	96.2	92.9	3.8	94.5
		Excessive	309	7	33	97.8	90.4	2.2	94.0
	Green	Normal	123	3	11	97.6	91.8	2.4	94.6
		Insufficient	119	5	11	96.0	91.5	4.0	93.7
		Excessive	130	3	26	97.7	83.3	2.3	89.9
Improved YOLOv3	Red	Normal	312	6	27	98.1	92.0	1.9	95.0
		Insufficient	278	10	21	96.5	92.9	3.3	94.6
		Excessive	311	8	31	97.5	90.9	2.5	94.1
	Green	Normal	124	3	10	97.6	92.5	2.4	94.9
		Insufficient	119	4	11	96.7	91.5	3.3	94.0
		Excessive	133	3	23	97.8	85.3	2.2	91.1

TABLE 4. Apple recognition results in blurred images.

Detection Models	Apple Type	Correct	Error	Omitted	P/%	R/%	E/%	F1/%
AlexNet + Faster RCNN	red	226	22	47	91.1	82.8	8.8	86.7
	green	69	6	25	92	73.4	8.0	81.6
ResNet101 + Faster RCNN	red	232	9	41	96.2	84.0	3.7	89.7
	green	74	2	20	97.3	78.7	2.6	87.0
DarkNet53 + YOLOv3	red	240	7	33	97.2	87.9	2.8	92.3
	green	78	2	16	97.5	83.0	2.5	89.6
Improved YOLOv3	red	241	5	32	97.9	88.3	2.0	92.8
	green	82	2	12	97.6	87.2	2.4	92.1

The recognition results were further described in detail in Table 4. It was found in Table 4 that all tested methods could be applied to detect the red and green in blurred images, providing F1 score of 81.6% at least. Improved YOLOv3 method still achieved the best performance on both red and green apple recognition with higher precision, higher recall, lower error and higher F1 score, followed by DarkNet53 + YOLOv3, ResNet101 + Faster RCNN and AlexNet + Faster RCNN respectively. Compared F1 value in the normal illumination shown in Table 3, there was an obvious decrease in the F1 values in Table 4, which might indicate that the sharpness had a great influence on the recognition effect of apple targets.

3) APPLE RECOGNITION WITH VARIOUS INTERFERENCE CONDITIONS

Compared with the clear and complete apples in the images, it was relatively difficult for recognizing the apples with various interference conditions. There were occluded apples, light spotted apples, overlapped apples, and incomplete apples to be recognized by the four models (Figure 12), and performances of which were evaluated in term of the indicator R as shown in Table 5.

It could be found in Table 5 that four models still had the accurate recognition results for apple with various interference conditions, among which DarkNet53+YOLOv3 model and improved YOLOv3 model have better recognition effect.

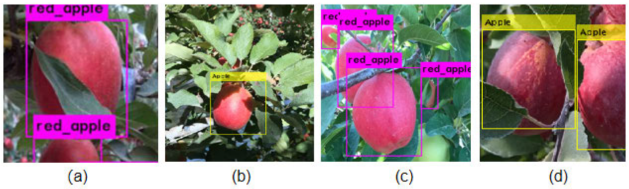


FIGURE 12. Examples of (a) occluded, (b) light spotted, (c) overlapped, and (d) incomplete apples recognition.

TABLE 5. Apple recognition results under various interference conditions.

Disturbance	Apple count	Detection model	TP	FN	R/%
overlap	127	AlexNet + Faster RCNN	112	15	88.1
		ResNet101 + Faster RCNN	119	8	93.7
		DarkNet53 + YOLOv3	122	5	96.0
		Improved YOLOv3	122	5	96.0
Occlusion	106	AlexNet + Faster RCNN	90	16	84.9
		ResNet101 + Faster RCNN	96	10	90.6
		DarkNet53+YOLOv3	98	8	92.4
		Improved YOLOv3	99	7	93.4
light spot	37	AlexNet + Faster RCNN	34	3	91.9
		ResNet101 + Faster RCNN	35	2	94.5
		DarkNet53 + YOLOv3	36	1	97.2
		Improved YOLOv3	35	2	94.5
Incomplete	52	AlexNet + Faster RCNN	41	11	78.8
		ResNet101 + Faster RCNN	45	7	86.5
		DarkNet53 + YOLOv3	46	6	88.5
		Improved YOLOv3	46	6	88.5

In the case of incomplete apples in the images, recognition recall rate decreased significantly for the four recognition models due to appearance features missing, up to 88.5%. Generally speaking, the improved YOLOv3 model performs better than the other three algorithms under most interference conditions due to model simplification and computational cost reduction.

IV. CONCLUSION

In this study, four different detection methods, AlexNet + Faster RCNN, ResNet101 + Faster RCNN, DarkNet53 + YOLOv3 and improved YOLOv3 were used to detect red and green apples in natural environment. Illumination conditions and image sharpness, occlusion and so on were used as control variables, and recognition accuracy, recall rate, false recognition rate and F1 value were used as evaluation indexes to compare the recognition effect of the four detection models in various complex natural environments. Through the analysis of experimental data, the following conclusions could be drawn:

1) The four recognition models based on deep learning algorithms used in this paper had good recognition effect on apples under various complex environments, which could provide technical support for automatic fruit picking in orchard.

2) The F1 value of improved YOLOv3 model was not less than 91.1% in different illumination conditions, which indicated that the improved YOLOv3 model was more suitable for apple detection in the natural environment than Faster RCNN algorithm and the original YOLOv3 algorithm.

3) Compared with different illumination conditions, the F1 values of four models for apple detection in blurred image had a decline at different levels. This showed that image sharpness had the greater impact on the recognition effect.

4) Future work will focus on developing the models to further improve the detection accuracy. A larger dataset and effective sample preprocessing are beneficial for establishing a robust model.

REFERENCES

[1] D. Zhang, Y. Xu, W. Huang, X. Tian, Y. Xia, L. Xu, and S. Fan, "Nondestructive measurement of soluble solids content in apple using near infrared hyperspectral imaging coupled with wavelength selection algorithm," *Infr. Phys. Technol.*, vol. 98, pp. 297–304, May 2019, doi: 10.1016/j.infrared.2019.03.026.

[2] J.-P. An, X.-W. Zhang, R.-R. Xu, C.-X. You, X.-F. Wang, and Y.-J. Hao, "Apple MdERF4 negatively regulates salt tolerance by inhibiting MdERF3 transcription," *Plant Sci.*, vol. 276, pp. 181–188, Nov. 2018, doi: 10.1016/j.plantsci.2018.08.017.

[3] S. Sun, M. Jiang, D. He, Y. Long, and H. Song, "Recognition of green apples in an orchard environment by combining the GrabCut model and ncut algorithm," *Biosyst. Eng.*, vol. 187, pp. 201–213, Nov. 2019.

[4] J. Song, T. Zhang, L. Xu, and X. Tang, "Research actuality and prospect of picking robot for fruits and vegetables," *Trans. Chin. Soc. Agricult. Machinery*, vol. 37, no. 5, pp. 158–162, 2006, doi: 10.1360/jos172537.

[5] A. Gongal, S. Amatya, M. Karkee, Q. Zhang, and K. Lewis, "Sensors and systems for fruit detection and localization: A review," *Comput. Electron. Agricult.*, vol. 116, pp. 8–19, Aug. 2015, doi: 10.1016/j.compag.2015.05.021.

[6] D. Wang, H. Song, and D. He, "Research advance on vision system of apple picking robot," *Trans. Chin. Soc. Agricult. Eng.*, vol. 33, no. 10, pp. 59–69, May 2017, doi: 10.11975/j.issn.1002-6819.2017.10.008.

[7] A. C. Tyagi, "Towards a second green revolution," *Irrigation Drainage*, vol. 65, no. 4, pp. 388–389, Oct. 2016, doi: 10.1002/ird.2076.

[8] Z. De-An, L. Jidong, J. Wei, Z. Ying, and C. Yu, "Design and control of an apple harvesting robot," *Biosystems Eng.*, vol. 110, no. 2, pp. 112–122, Oct. 2011, doi: 10.1016/j.biosystemseng.2011.07.005.

[9] J. R. Davidson, C. J. Hohimer, and C. Mo, "Preliminary design of a robotic system for catching and storing fresh market apples," *IFAC-PapersOnLine*, vol. 49, no. 16, pp. 149–154, 2016, doi: 10.1016/j.ifacol.2016.10.028.

[10] H. A. M. Williams, M. H. Jones, M. Nejati, M. J. Seabright, J. Bell, N. D. Penhall, J. J. Barnett, M. D. Duke, A. J. Scarfe, H. S. Ahn, J. Lim, and B. A. MacDonald, "Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms," *Biosyst. Eng.*, vol. 181, pp. 140–156, May 2019, doi: 10.1016/j.biosystemseng.2019.03.007.

[11] H. Peng, Y. Shao, K. Chen, Y. Deng, and C. Xue, "Research on multi-class fruits recognition based on machine vision and SVM," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 817–821, 2018, doi: 10.1016/j.ifacol.2018.08.094.

[12] S. Sengupta and W. S. Lee, "Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions," *Biosyst. Eng.*, vol. 117, pp. 51–61, Jan. 2014, doi: 10.1016/j.biosystemseng.2013.07.007.

[13] S. Sun, Q. Wu, L. Jiao, Y. Long, D. He, and H. Song, "Recognition of green apples based on fuzzy set theory and manifold ranking algorithm," *Optik*, vol. 165, pp. 395–407, Jul. 2018, doi: 10.1016/j.ijleo.2018.03.085.

[14] L. Jidong, Z. De-An, J. Wei, and D. Shihong, "Recognition of apple fruit in natural environment," *Optik*, vol. 127, no. 3, pp. 1354–1362, Feb. 2016, doi: 10.1016/j.ijleo.2015.10.177.

[15] W. Ji, D. Zhao, F. Cheng, B. Xu, Y. Zhang, and J. Wang, "Automatic recognition vision system guided for apple harvesting robot," *Comput. Electr. Eng.*, vol. 38, no. 5, pp. 1186–1195, Sep. 2012, doi: 10.1016/j.compeleceng.2011.11.005.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[17] R. Wang, Y. Shi, and W. Cao, "GA-SURF: A new speeded-up robust feature extraction algorithm for multispectral images based on geometric algebra," *Pattern Recognit. Lett.*, vol. 127, pp. 11–17, Nov. 2019, doi: 10.1016/j.patrec.2018.11.001.

[18] R. Wang, M. Shen, T. Wang, and W. Cao, "L1-norm minimization for multi-dimensional signals based on geometric algebra," *Adv. Appl. Clifford Algebras*, vol. 29, no. 2, p. 33, Apr. 2019, doi: 10.1007/s00006-019-0950-7.

[19] W. Cao, Q. Liu, and Z. He, "Review of pavement defect detection methods," *IEEE Access*, vol. 15, pp. 14531–14544, 2020, doi: 10.1109/ACCESS.2020.2966881.

[20] P. A. Dias, A. Tabb, and H. Medeiros, "Apple flower detection using deep convolutional networks," *Comput. Ind.*, vol. 99, pp. 17–28, Aug. 2018, doi: 10.1016/j.compind.2018.03.010.

[21] L. Fu, Y. Feng, Y. Majeed, X. Zhang, J. Zhang, M. Karkee, and Q. Zhang, "Kiwifruit detection in field images using faster R-CNN with ZFNet," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 45–50, 2018, doi: 10.1016/j.ifacol.2018.08.059.

[22] J. Xiong, Z. Liu, L. Tang, R. Lin, and H. Peng, "Visual detection technology of green citrus under natural environment," *Trans. Chin. Soc. Agricult. Machinery*, vol. 49, no. 4, pp. 45–52, 2019, doi: 10.6041/j.issn.1000-1298.2018.04.005.

[23] H. Peng, B. Huang, Y. Shao, Z. Li, and J. Xiong, "General improved SSD model for picking object recognition of multiple fruits in natural environment," *Trans. Chin. Soc. Agricult. Machinery*, vol. 34, no. 16, pp. 155–162, 2018, doi: 10.11975/j.issn.1002-6819.2018.16.020.

[24] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, "Apple detection during different growth stages in orchards using the improved YOLO-V3 model," *Comput. Electron. Agricult.*, vol. 157, pp. 417–426, Feb. 2019, doi: 10.1016/j.compag.2019.01.012.

[25] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

[26] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.

[27] D. Zhao, R. Wu, X. Liu, and Y. Zhao, "Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background," *Trans. Chin. Soc. Agricult. Eng.*, vol. 35, no. 3, pp. 164–173, 2019, doi: 10.11975/j.issn.1002-6819.2019.03.021.

[28] 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, doi: 10.1109/TPAMI.2016.2577031.

[29] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[30] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*. [Online]. Available: http://arxiv.org/abs/1804.02767

[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.

[32] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

GUANTAO XUAN received the M.S. degree in mechanical and electronic engineering from Shandong University, Jinan, Shandong, China, in 2004, and the Ph.D. degree in agricultural mechanization engineering from Shandong Agricultural University, Tai'an, Shandong, China, in 2014. From 2017 to 2019, he was a Postdoctoral Research Fellow with the University of Missouri, Columbia, MO, USA. He is currently an Associate Professor with Shandong Agricultural University. He has published over 20 articles in reputed journals/conferences. His current research interests include agricultural automation, plant phenotyping, computer vision, and spectroscopy technology.

CHONG GAO was born in Jining, Shandong, China, in 1997. He received the B.S. degree in mechanical engineering from Shandong Agricultural University, Tai'an, Shandong, China, in 2019, where he is currently pursuing the M.S. degree in mechanical engineering.

He began scientific research with the School of Mechanical and Electrical Engineering, Shandong Agricultural University, in 2019. He is the author of several patents and articles. He holds four software copyrights. His research interests include deep learning, image recognition, hyperspectral nondestructive testing, spectral diagnosis, embedded system development, software development, and so on.

Mr. Gao received awards and honors, include the Scholarship of Shandong Agricultural University and the National Encouragement Scholarship.

YUANYUAN SHAO was born in Jining, Shandong, China, in 1980. She received the Ph.D. degree in mechanical and electronic engineering from the Shandong University of Science and Technology, Qingdao, Shandong, China, in 2012.

Since 2016, she has been an Associate Professor with the Mechanical and Electronic Engineering College, Shandong Agricultural University, Tai'an, Shandong, China. She is the author of two books, more than 20 articles, and more than four inventions. Her research interests include smart agriculture, nondestructive testing of agricultural products quality, imaging process, deep learning, hyperspectral technology, mechanical design, and so on.

MENG ZHANG was born in Zouping, Shandong, China, in 1982. He received the B.S. degree in mechanical engineering from Jinan University, China, in 2004, and the M.S. degree in industrial economics from the Hebei University of Science and Technology, China, in 2007. He is currently pursuing the Ph.D. degree in applied economics with Nanjing University, Nanjing, China.

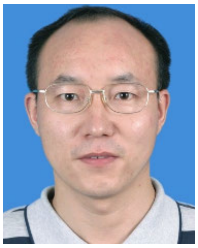
Since 2016, he has been an Associate Professor with the Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing. He is the author of two books and more than 20 articles. His research interests include agricultural mechanization developing strategies, smart agriculture, and so on.

YONGXIAN WANG was born in Tai'an, Shandong, China, in 1995. He received the B.S. degree in mechanical engineering from Zaozhuang University, Shandong, China, in 2018, and the M.S. degree in mechanical engineering from Shandong Agricultural University, Shandong, China, in 2020. His research interests include research on hyperspectral agriculture application and agricultural intelligent equipment.

JINGRUN ZHONG received the M.S. degree in computer technology from South China Agricultural University, in 2020. His current research interests include computer vision and image processing.

QINGGUO LI was born in Zaozhuang, Shandong, China, in 1982. He received the M.S. degree in mechanical engineering from the Shandong University of Science and Technology, Qingdao, Shandong, China, in 2007.

From 2007 to 2013, he was a Mechanical Design Engineer with the Taiyuan Research Institute, China Coal Technology and Engineering Group Corp., Taiyuan, Shanxi, China. From 2013 to 2018, he was a Deputy General Manager with Shanxi Kelong Energy Technology Company Ltd., Xianyang, Shanxi, China. Since 2018, he has been a Deputy General Manager with Tai'an Jiacheng Technology Company Ltd., Tai'an, Shandong, China. He is the author of two provincial and ministerial awards. He holds more than ten inventions. His research interests include new energy mine special vehicle, machine vision, underground driverless vehicle, and underground disaster relief robot.



HONGXING PENG received the B.S. degree from Jiangxi Normal University, Nanchang, China, in 1998, the M.S. degree from Tianjin Normal University, Tianjin, China, in 2004, and the Ph.D. degree from South China Agricultural University, Guangzhou, China, in 2014. From 2004 to 2014, he was a Lecturer with South China Agricultural University. Since 2015, he has been an Associate Professor with the Computer Science and Engineering Department, South China Agricultural University. His research interests include computer vision, image processing, artificial intelligence, robot, big data, and virtual reality.