Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-17 (2018) 45–50

Kiwifruit detection in field images using Faster R-CNN with ZFNet

Longsheng Fu*. Yali Feng. Yaqoob Majeed***. Xin Zhang****. Jing Zhang*****. Manoj Karkee******. Qin Zhang*******.**

* College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China
Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture, Yangling 712100, China
Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, 712100, China
Center for Precision&Automated Agricultural Systems, Washington State University, Prosser, WA 99350, USA
(**Corresponding Author**, Tel: +86-2987092391; e-mail: fulsh@nwafu.edu.cn, longsheng.fu@wsu.edu)
** College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China
(e-mail: lzhdelightsome@foxmail.com)
*** Center for Precision and Automated Agricultural Systems, Washington State University, Prosser, 99350, USA
(e-mail: yaqoob.majeed@wsu.edu)
**** Center for Precision and Automated Agricultural Systems, Washington State University, Prosser, 99350, USA
(e-mail: xin.zhang4@wsu.edu)
***** Center for Precision and Automated Agricultural Systems, Washington State University, Prosser, 99350, USA
(e-mail: j.zhang@wsu.edu)
****** Center for Precision and Automated Agricultural Systems, Washington State University, Prosser, 99350, USA
(e-mail: manoj.karkee@wsu.edu)
******* Center for Precision and Automated Agricultural Systems, Washington State University, Prosser, 99350, USA
(e-mail: qinzhang@wsu.edu)

Abstract: A kiwifruit detection system for field images was developed based on the deep convolutional neural network, which has a good robustness against the subjectivity and limitation of the features selected artificially. Under different lighting conditions, 2,100 sub-images with 784×784 pixels were prepared by random sub-sampling from 700 field captured images with a pixel resolution of 2352×1568 pixels. Sub-images were used as network training and validation samples. A faster R-CNN was trained end-to-end by using back-propagation and stochastic gradient descent techniques with Zeiler and Fergus network (ZFNet). The average precision of the Faster R-CNN-based kiwifruit detector was 89.3%. Finally, another 100 images of kiwifruit canopies in the field environment (including 5,918 fruits) were used for testing the network. The test results showed that the recognition ratio of occluded fruit, overlapping fruit, adjacent fruit and separated fruit were 82.5%, 85.6%, 94.3% and 96.7%, respectively. Overall, the model reached a recognition rate of 92.3%. The technique took 0.274 s to process each image (for images with 2352×1568 pixels) and only 5.0 ms on average to detect a fruit. Comparing against the conventional methods, it suggested that the proposed method has higher recognition rate and faster speed. Especially, the proposed technique was able to simultaneously detect individual kiwifruit in clusters, which provides a promise for accurate yield mapping and multi-arm robotic harvesting.

Keywords: image recognition; kiwifruit detection; Faster R-CNN; ZFNet; multi clusters

1. INTRODUCTION

China is the largest cultivator of kiwifruits with 1.45×10^5 ha (a yield of 1.84×10^6 t) planted in 2014. Shaanxi Province has the largest contribution to this industry with approximately 70% of the total production in China, and 33% of the global production coming from this province (Sun, 2013). Harvesting kiwifruits in this area rely mainly on manual picking, which is labor-intensive (Fu et al., 2015a). Therefore, there is a strong desire to introduce mechanical/robotic harvesting for kiwifruit.

Kiwifruits are commercially grown on sturdy support structures, such as T-bar and pergola systems (Huang and Ferguson, 2001; Miller et al., 2001). T-bar trellis is common in China because of its low cost. It consists of a 1.7 m high post and an approximately 1.7 m wide cross arm, which may vary slightly in width according to the shape and size of the orchard. Wires run on the top of the cross arms connecting them to each other in the middle and on both edges of the cross arms. The upper stems of the kiwifruit are tied to the top wires so that the egg-sized kiwifruits would be hanging downward, which makes them visible and accessible for manual picking (Chen et al., 2012; Fu et al., 2015b). This canopy structure also provides relatively simpler and structured workspace for mechanized or automated field operations such as robotic picking compared to other fruit trees such as apples.

As discussed before, it is highly desirable to investigate robotic harvesting systems for this crop due to a large labor force required in the manual picking operation. Fast and effective detection of kiwifruit in the field under natural environment is the foundational step for robotic harvesting.

10.1016/j.ifacol.2018.08.059

In the past, various studies were conducted in China and New Zealand (another country with a large plantation of kiwifruit). Ding et al. (2009) used R-B color parameters for fruit segmentation on natural backgrounds, but this approach was limited in identifying all fruit, particularly single fruit in clusters. Wu et al. (2012) applied Haar training to identify fruit and watershed algorithm to separate adjacent fruit. Cui et al. (2013) studied fruit recognition and feature extraction based on color and shape of kiwifruit in outdoor environment in daytime. Zhan et al. (2013) developed an image processing technique based on an Adaboost algorithm to segment kiwifruits from the background. Mu et al. (2014) employed L*a*b* color space to extract characteristic parameters of kiwifruits by using Canny operator for edge detection and elliptical Hough transformation for fruit recognition. Those research efforts suggested the scenario of harvesting in daytime taking advantage of the sunlight and captured kiwifruit images by locating the cameras near the underside of the canopy and its central axis approximately parallel to the canopy. This technique, however, caused the background of the images to contain pendulous foliage or remote non-target fruit, which would add excessive noise in image segmentation, and thus influence accuracy. Moreover, target fruit in the images might overlap since kiwifruit grows in clusters. This issue also caused difficulties to identify individual target fruit and decreased recognition accuracy.

Another image acquisition method for kiwifruits was to place the camera underneath the fruits, so that its central axis would be perpendicular to the canopy since kiwifruits are commercially grown on sturdy support structures as discussed before (Fu et al., 2015a; Scarfe et al., 2009). From bottom-up view, most target fruits and their calyxes (each fruit has one calyx) would appear in the camera field of view, and the fruits in cluster would appear adjacent to each other instead of overlapping to each other, as shown in Fig. 1. Scarfe (2012) segmented images acquired with sensors looking up into the horizontal canopy from underneath (called *bottom-viewed* images in this paper). They subtracted a predefined reference color range and used Sobel filter to detect fruit and calyx edges, then applied a diameter of 28 pixels circular template to scan over from extract calyx area and assigned center of mass as each fruit position. This approach identified 83.6% of kiwifruit with the processing time of 15.2 to 28.7 s for each image. Fu et al. (2015b) segmented bottom-viewed kiwifruit images using Otsu threshold in 1.1R-G color component and eliminated noises by a morphological operation and an area thresholding. Then, fruit boundaries were extracted using the Canny filter, and each target fruit area was detected using minimal bounding rectangle and elliptical Hough transform. This algorithm, on average, achieved 88.3% recognized rate and 0.8 s processing time to segment an image and 1.5 s to detect per fruit. These studies primarily used color and shape of calyx and fruit to detect the kiwifruit, which was limited to detect fruit in a single cluster with few fruit and were less effective for multi-clusters in the field. To overcome these limitations, a well-generalized model that is invariant and robust to brightness and viewpoint changes and highly discriminative feature representations is desired.

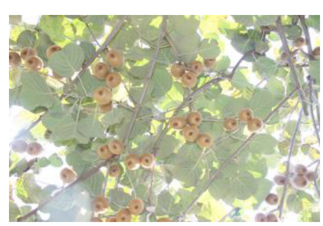
Recent years have witnessed enormous advancement in the computer vision research based on deep learning. A variety of vision-based tasks, such as object recognition, classification, and counting, can be performed with a deep learning technique to achieve a high level of accuracy and robustness. Faster R-CNN (Ren et al., 2015) was one of those deep learning network widely used for increasing computational speed by introducing Region Proposal Network (RPN), which shares convolutional features with the classification network, and two networks are concatenated as one network that can be trained and tested for a complete or through end-to-end process. Some pre-trained models (e.g. Zeiler and Fergus network, ZFNet) (Zeiler and Fergus, 2013) were developed for object detection based on Faster R-CNN.

In this work, a kiwifruit detection system was developed for field images based on the Faster R-CNN implemented by ZFNet, which has a good robustness avoiding the subjectivity and limitation of the features selected artificially in the conventional image processing methods.

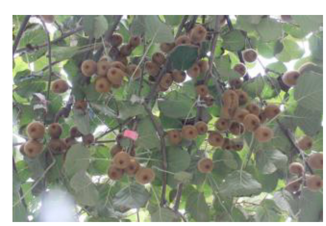
2. MATERIALS AND METHODS

2.1 Image Acquisition

In this study, the canopy images were captured using a camera placed underneath the fruits, so that its central axis would be perpendicular to the canopy. A common single-lens reflex camera (Canon S110, Canon Inc., Tokyo, Japan) on "AUTO" mode with a resolution of 2352×1568 was placed at around 100 cm below the fruit surface to optimize the number of kiwifruits included in the field of view of the camera. RGB (Red, Green, and Blue) images were taken in late October 2017 during the harvesting season. An experimental plot of ‘Hayward’ kiwifruit, which is one of the most common cultivars in China, planted in Shaanxi at the Meixian Kiwifruit Experimental Station (34°07'39"N, 107°59'50"E, and 648 m in altitude), the Northwest A&F University was used. The images were taken in different times of the day (400 images each in the morning and afternoon) with varied lighting conditions, as shown in Fig. 1.



a. Morning



b. Afternoon

Fig. 1 Kiwifruit images under the natural lighting conditions in field environment

2.2 Image Datasets

In total, the image dataset consisted of 700 images with 2352×1568 pixel resolution. Each image included around 30 to 100 target kiwifruits. It would be a challenging task to manually label a large number of small objects in high resolution images (Bargoti and Underwood, 2017). Therefore, this problem was alleviated by randomly sampling smaller subset from the larger set of images. Each sample image was then divided into 6 sub-images with 784×784 pixels, and a sub-image was randomly sampled from each column and maximize the variability in the sample dataset. A total of 2,100 sub-images were collected and manually annotated for the ground truth using rectangular annotations. Each sub-image included 10 kiwifruits in average. The labelled dataset for kiwifruit was divided into training, validation and testing groups. The training images were randomly obtained from the independent and uniform sampling of the whole dataset. The validation images and test images were mutually exclusive, which ensured the reliability of the later evaluation standards. The data configuration was summarized in Table 1.

Table 1 Image dataset used to training the Faster R-CNN for kiwifruit detection

Sub-image size (pixel)	Total no. of images	No. of training images	No. of validation images	No. of test images
784×784	2100	1176	504	420

2.3 Experimental environment

The training platform included a desktop computer with Intel i5, 6400 (2.70 GHz) quad-core CPU, and a GeForce GTX 960M 4 GB GPU (1536 CUDA cores) and 16 GB of memory, running on a Windows 7 64 bit system. The software tools used included CUDA 7.5, CUDNN 4.0, Matlab R2016a and Microsoft Visual Studio 12.0. The experiments were implemented in the Caffe framework (Jia et al., 2014).

2.4 Construction of Faster R-CNN in ZFNet

The Faster R-CNN model merges region proposals and object classification and localization into one unified deep object detection network, and two networks (RPN Network and Faste R-CNN Network) are concatenated as one network that can be trained and tested through an end-to-end process (Bargoti and Underwood, 2016). It can be applied in most state of the art CNN models.

ZFNet (Zeiler and Fergus, 2013) implementation was used in this study to test the Faster R-CNN for kiwifruit detection. ZFNet was the winner of the ILSVRC (ImageNet Large-Scale Visual Recognition Challenge) in 2013. It provided great intuition on to how CNNs work and illustrated more ways to improve the performance. Intermediate outputs from an input image resulted in by the ZFNet are illustrated in Fig. 2.

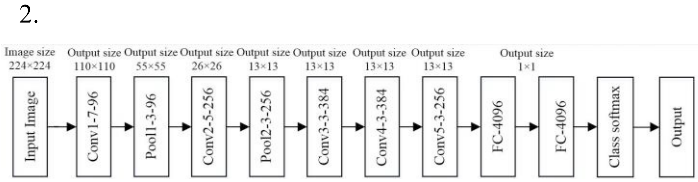


Fig. 2 Illustration of the ZFNet architecture. It consisted of 5 shareable convolutional layers, max-pooling layers, dropout layers, and 3 fully connected layers. It used a 7×7 size filter and a decreased stride value in the first layer. The last layer of ZFNet is the softmax layer.

The input to ZFNet was an RGB image was quantized to 224×224 pixels. The image was processed by the network resulting in the probability that the image belongs to each category. The network consisted of 5 shareable convolutional layers, max-pooling layers, dropout layers, and 3 fully connected layers. It used a 7×7 size filter and a decreased stride value in the first layer. The last layer of ZFNet is the softmax layer, which was used to convert the score into the probability that the image belongs to each category. In ZFNet, in addition to the pooling layer, a modified linear unit (ReLU) was set behind each hidden layer for nonlinear transformation, which helps to increase the speed of network convergence.

2.5 Network Training

Kiwifruit detection can be achieved by segmenting and locating the rectangular boundaries (candidate areas) of individual fruit (objects to be detected) in an image, and then performing feature extraction and classification of candidate regions using the recognition network.

The Faster R-CNN uses color (RGB) images to perform general object detection. The object detection system consists of two parts: i) a Region Proposal Network (RPN) for detecting the region of interests (RoIs) in the image, and ii) a classification module, which classifies the individual regions and regresses a bounding box around the object (Sa et al., 2016). The training architecture for kiwifruit image detection based on Faster R-CNN with ZFNet was illustrated in Fig. 3. The RPN is implemented as a full convolutional network, which was optimized through end-to-end using back-propagation and SGD. Non-maximum suppression was applied on the proposal regions to reduce redundancy. The Faster R-CNN framework is capable of multi-class detection. Our work only considered a binary classification problem of kiwifruit images acquired in field environment. In this case, the output layer resulted in only "background" and "kiwi" regions and marked the fruit position with a rectangular box.

The random initialization of the weights takes more time to converge the model to a stable value or may even fall into a local minimum. Transfer learning (Zhen et al., 2017) can quickly adapt to new tasks in the event of insufficient data set. Using pre-trained model with a large dataset, the underlying structure weight parameters can be shared, and the model can be fine-tuned to overcome the differences between the pre-training and new datasets. In this work, the shared convolutional layers of Faster R-CNN were initialized by pre-training a model for ImageNet classification. All other layers were randomly initialized by drawing weights from a zero-mean Gaussian distribution with a standard deviation of 0.01.

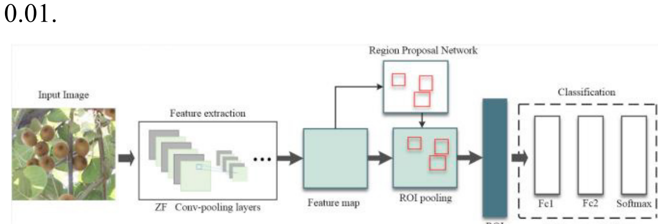


Fig. 3 Training architecture for kiwifruit detection based on Faster R-CNN with ZFNet

2.6 Fruit categories

The fruits in the images taken in the field were not all independent (or singulated) to each other. In this work, fruits were categorized into four groups according to the degree of completeness of fruit contours in the image. The first category (called occluded fruit, Fig. 4a) included fruit partially occluded by leaves or branches causing incomplete outline of the fruit. The second category refers to the fruit with some overlap between each other (called overlapping fruit; rectangular boxes in Fig. 4b). The third category refers to the fruits where contours of two or more fruits were adjacent to each other (called adjacent fruit, Fig. 4c). The fourth category was separated fruit, in which fruit contours were completely independent and separated from each other (Fig. 4d).

a. Occluded b. Overlapped c. Adjacent d. Separated

Fig. 4 Categories of individual kiwifruit in canopy images

3. RESULTS AND DISCUSSION

3.1 Precision and recall of the network

This paper follow the "image-centric" sampling strategy to train the Faster R-CNN described above. The momentum of the network was set to a fixed value of 0.9 and a weight decay of 0.0005 used. Mini-batches of 128 examples were used trained our model. The same learning rate of 0.001 was used for all layers in the network. During the training process, when the increase in classification accuracy over the validation set was below certain threshold, the learning rate was reduced by changing it to 10% of the current rate until the classification accuracy no longer improved (or saturated). It took about 10 hours to perform a total of 40,000 iterations over the training set. Trajectory of Precision-Recall (PR) curve of the network during the training process was shown in Fig. 5.

Average Precision (AP), which is the area under the PR Curve, was used to evaluate the performance of the network in detecting kiwifruit. AP is a standard measure for measuring the sensitivity of the network to a target object, and is an indicator that reflects the global performance of the network. The higher the AP value, the better the detection accuracy of the convolutional neural network is. AP of the network at the end of the training process was 89.3%, indicating that the convolutional neural network has the potential to achieve expected training effect.

3.2 Validation of the model on the original images

The trained network/model was then tested for its reliability and stability with 100 kiwifruit images (including 5,918 kiwifruits) acquired in the field environment in the morning and afternoon under different natural lighting conditions. The detection results were shown in Table 2, and some sample images of Faster R-CNN detection were shown in Fig. 6.

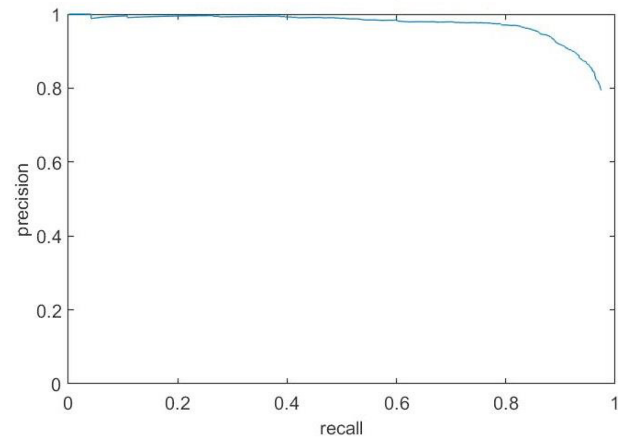


Fig. 5 Precision-Recall curve of the network

Table 2 Results of kiwifruit detection with a faster R-CNN

	Occluded Fruit	Overlapped Fruit	Adjacent Fruit	Separated Fruit	Overall
Recognized Fruits	671	451	3378	964	5464
Total Fruits	813	527	3581	997	5918
Recognition Percentage	82.5%	85.6%	94.3%	96.7%	92.3%

The detection rate of the separated/individual fruit was the highest (96.7%), followed by the adjacent fruit (94.3%) and then the overlapped fruit (85.6%). The lowest detection rate was achieved for occluded fruits (82.5%). The trained Faster R-CNN model performed an overall prediction of 92.3%, with detection time averaging at 0.274 s per image (for images with 2352×1568 pixels).

The test results of Faster R-CNN (Fig. 6) showed that the detector has good detection results for these images acquired with multiple clusters of kiwifruits from natural conditions in the morning and afternoon under different natural lighting conditions. As can be seen from the Fig. 6, this algorithm has better generalization and good robustness for kiwifruits with different occlusion ratios, overlap ratios, different forms of separated and adjacent in the image. In addition, the kiwifruit candidate detection area (red boxes in Fig. 6) covered the fruit very well in the images. From the test results, the detection probability about 80% of the fruits was greater than 0.9, indicating that the Faster R-CNN used in this work can effectively classify and identify kiwifruit within the candidate image areas. The detector can effectively detect different categories of fruit in the field environment.

3.3 Comparison with conventional methods

The performance of the proposed fruit detection technique was compared with four conventional methods used by Scarfe (2012), Fu et al. (2015b), Fu et al. (2017), and Fu et al. (2018). The results were shown in Table 3.

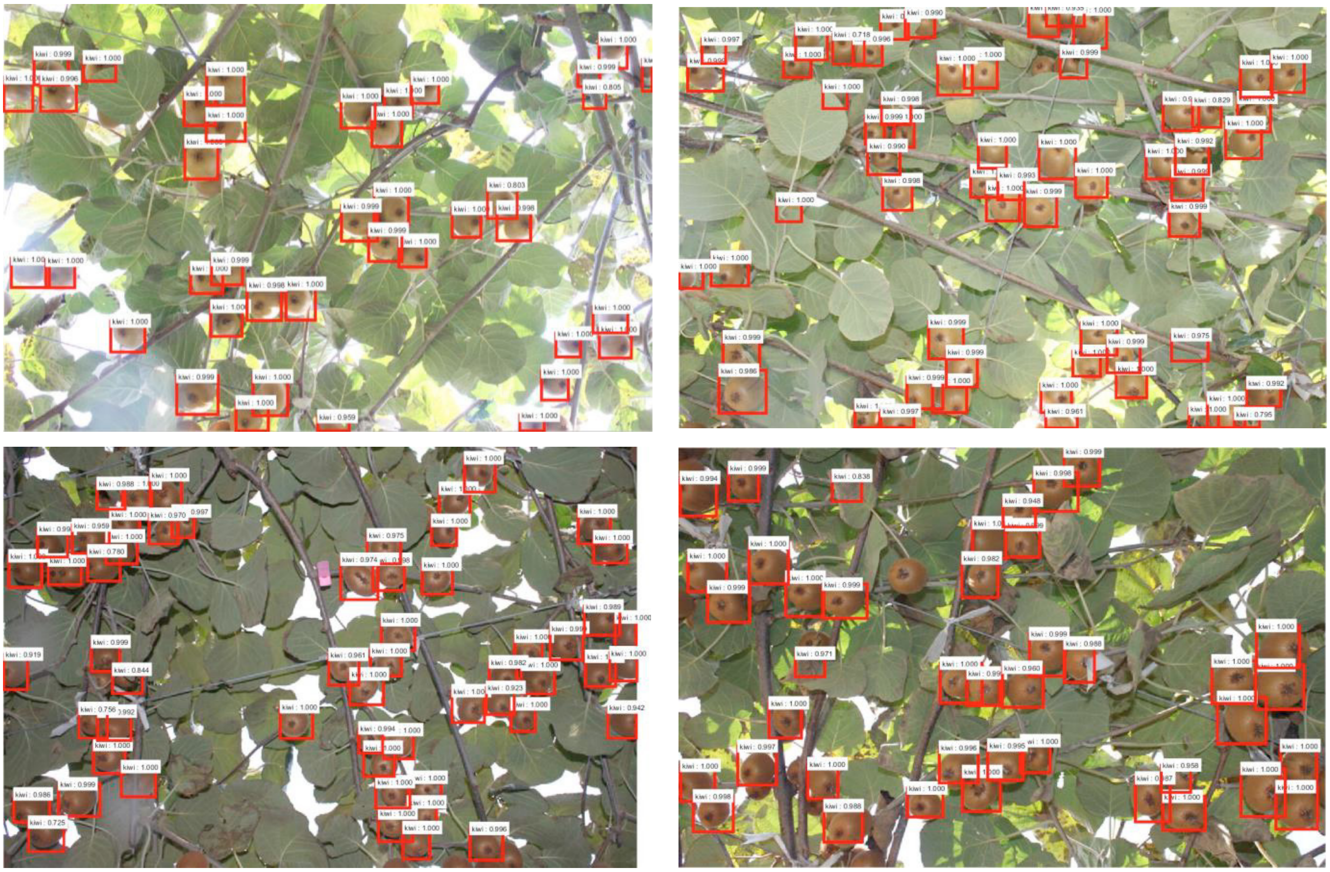


Fig. 6 Examples of kiwifruit images recognized by Faster R-CNN with ZFNet

The proposed technique for kiwifruit detection in images from field environment obtained by a regular RGB color camera provided a foundation for accurately performing higher-level tasks such as yield mapping and estimation. Each of the test images used in the experiment contains the four categories of the fruit, and each image contains at least 30 fruits. Under the varied lighting conditions, the overall recognition rate of this algorithm was 92.3%, which was higher than other three algorithms (Scarfe, 2012; Fu et al., 2015b; Fu et al., 2018), but lower than that of Fu et al (2017). The images of kiwifruit taken at close distance by Fu et al. (2015b) and Fu et al. (2017), and the algorithm used only focused on the fruit of single clusters with separated and adjacent, and had poor recognition effect on the cluster that with 5 or more fruits due to the lack of high-level image expression, and it was difficult to reflect the spatial relationships among the selected underlying feature. The recognition method of Scarfe (2012) needed to extract artificially selected bottom layer features from kiwifruit images, perform a large number of preprocessing on the images, and have complex operations. Under the same experimental field, the recognition rate of this algorithm was 92.3% higher than 89.3% of Fu et al. (2017) and 83.6% of Scarfe (2012). It showed that this technique performed well in images acquired from natural conditions with multiple clusters of kiwifruits. In addition, the recognition speed of this work has greatly improved to 5.0 ms per fruit. The algorithm of this work was more suitable for picking requirements in the practical application of multi-arm kiwifruit picking robot.

Table 3 Comparing of CNNs and other methods for kiwifruit image recognition

Method	Shooting distance	Image features	Fruit categories	Recognition rates (%)	Recognition speed (s/ fruit)
Ref. algorithm (Scarfe. 2012)	Long distance	Multiple clusters	Occlusion, Overlap, Adjacency, Separated	83.6	0.28
Ref. algorithm (Fu et al. 2015b)	Close distance	Single cluster	Adjacency, Separated	88.3	1.64
Ref. algorithm (Fu et al. 2017)	Close distance	Single cluster	Adjacency, Separated	94.3	0.5
Ref. algorithm (Fu et al. 2018)	Long distance	Multiple clusters	Occlusion, Overlap, Adjacency, Separated	89.3	0.27
Proposed method	Long distance	Multiple clusters	Occlusion, Overlap, Adjacency, Separated	92.3	0.50×10^{-2}

4. CONCLUSIONS

In this work, a Faster R-CNN-based kiwifruit recognition model was developed and evaluated in images collected from field environment. ZFNet framework was used to implement the Faster R-CNN. The results showed that the model used in this work obtained a good detection accuracy on kiwifruit images that captured in the day time. This technique performed well in images acquired from natural conditions with multiple clusters of kiwifruits, which has been the limitation of most of the conventional image processing algorithms. The model proposed in this work took a substantially short computational time to processes images and has relatively good robustness to light variance and foliage blocking, which provides strong support for the research on crop-load estimation and robotic picking kiwifruit with multi-arm operations.

ACKNOWLEDGEMENTS

The work was supported by the Key Research and Development Program in Shaanxi Province of China (Project No. 2017NY-164), and the Science and Technology Innovation Project in Shaanxi Province of China (grant number 2015KTCQ02-12), and the International Scientific and Technological Cooperation Foundation of Northwest A&F University (A213021505). The "Young Faculty Study Abroad Program" of Northwest A&F University Scholarship sponsors Dr. Longsheng Fu in conducting post-doctoral research at Centre for Precision and Automated Agricultural Systems (CPAAS), Washington State University (WSU). The China Scholarship Council (CSC) sponsors Xin Zhang and the University of Agriculture, Faisalabad (UAF) of Pakistan sponsors Yaqoob Majeed in conducting Ph.D. dissertation studies at WSU CPAAS. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of WSU.

REFERENCES

Bargoti, S., & Underwood, J. (2016). Deep fruit detection in orchards. *Robotics and Automation (ICRA), 2017 IEEE International Conference on*.

Bargoti, S., & Underwood, J. (2017). Image segmentation for fruit detection and rield estimation in apple orchards. *Journal of Field Robotics, 34*(6), 1039-1060.

Chen, Y, Chen X, & Liu Y. (2012). Study on frame type of kiwifruit. *North Hortic,*14, 56-57.

Cui, Y., Su, S., Wang, X., Tian, Y., Li, P., & Zhang, F. (2013). Recognition and feature extraction of kiwifruit in natural environment based on machine vision. *Transactions of the Chinese Society for Agricultural Machinery, 44*(5), 247-252.

Ding, Y., Geng, N, & Zhou., Q. (2009). Research on the object extraction of kiwifruit based on images. *Microcomputer Information, 25*(18), 294-295.

Fu, L., Feng, Y., Elkamil T, Liu, Z., Li, R., & Cui, Y. (2018). Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks. *Transactions of the Chinese Society of Agricultural Engineering*, *34*(2), 205-211.

Fu, L., Sun, S., Vázquezarellano, M., Li, S., Li, R., & Cui, Y. (2017). Kiwifruit recognition method at night based on fruit calyx image. *Transactions of the Chinese Society of Agricultural Engineering*, *33*(2), 199-20.

Fu, L., Zhang, F., Yoshinori, G., Zhen, L., Wang, B., & Cui, Y. (2015a). Development and experiment of end-effector for kiwifruit harvesting robot. *Transactions of the Chinese Society for Agricultural Machinery, 46*(3), 1-8.

Fu, L., Wang, B., Cui, Y. , Shuai, S., Gejima, Y., & Kobayashi, T. (2015b). Kiwifruit recognition at nighttime using artificial lighting based on machine vision. *International Journal of Agricultural & Biological Engineering, 8*(4), 52-59.

Huang, H., & Ferguson, A. R. (2001). Review: kiwifruit in China. *New Zealand Journal of Experimental Agriculture, 29*(1), 1-14.

Inkyu, S., Ge, Z., Feras, D., Ben, U., Tristan, P., & Chris, M. C. (2016). DeepFruits: A fruit detection system using deep neural networks. *Sensors, 16*(8), 1222.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., & T. Darrell. (2014). Caffe: convolutional architecture for fast feature embedding. Proceedings of the 22nd ACM international conference on Multimedia, 675-678.

Miller, S., Broom, F., Thorp, T., & Barnett, A. (2001). Effects of leader pruning on vine architecture, productivity and fruit quality in kiwifruit (Actinidia deliciosa cv. Hayward). *Scientia Horticulturae, 91*(3–4), 189-199.

Mu, J., Chen, J., Sun, G., Liu F., Ma, Y., & Wang, F. (2014). Characteristic parameters extraction of kiwifruit based on machine vision. *Journal of Agricultural Mechanization Research, 36*(6), 138-142.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: towards real-time object detection with region proposal networks. *International Conference on Neural Information Processing Systems.* MIT Press*, 39*(6): 91-99.

Scarfe, A. (2012). Development of an autonomous kiwifruit harvester, *Massey University: Manawatu, New Zealand*.

Scarfe, A., Flemmer, R., Bakker, H., & Flemmer, C. (2009). Development of an autonomous kiwifruit picking robot. *Proceedings of the 4th International Conference on Autonomous Robots and Agents*, 380-384, Wellington, New Zealand

Sun, Z. (2013). Fruit area and yield was increased 12 years in shannxi. *China Fruit News, 30*(1): 44-45.

Wu, T., Yuan, C., & Chen, J. (2012). Research on the object extraction of kiwifruit based on machine vision. *Journal of Agricultural Mechanization Research, 34*(12), 21-26.

Zhen, X., Chen, J., Zhong, Z., Hrycushko, B., Jiang, S., Albuquerque, K., & Gu, X. (2017). Exploring Deep Convolution Neural Networks with Transfer Learning for Rectum Toxicity. *Prediction in Cervical Cancer Radiotherapy. Brachytherapy*, 16(3), S33-34.

Zeiler, M., & Fergus, R. (2013). Visualizing and understanding convolutional networks. *European Conference on Computer Vision, 8689*, 818-833.

Zhan, W., He, D., & Shi, S. (2013). Recognition of kiwifruit in field based on Adaboost algorithm. *Transactions of the Chinese Society of Agricultural Engineering*, 29(23), 140-146.