[26] 王前程. 基于深度学习的水果图像识别算法研究[D].保定：河北大学，2016.
Wang Qiancheng, The Algorithm Research of Fruit Image Recognition Based on Deep Learning[D]. Baoding: Hebei University，2016. (in Chinese with English abstract)

[27] Sa I, Ge Z, Dayoub F, et al. Deepfruits: A fruit detection system using deep neural networks[J]. Sensors, 2016, 16(8): 1□23.

[28] 杨国国，鲍一丹，刘子毅. 基于图像显著性分析与卷积神经网络的茶园害虫定位与识别[J]. 农业工程学报，2017，33(6)：156－162.
Yang Guoguo, Bao Yidan, Liu Ziyi. Localization and recognition of pests in tea plantation based on image saliency analysis and convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(6): 156－162. (in Chinese with English abstract)

[29] Vedaldi A, Lenc K. MatConvNet: Convolutional neural networks for MATLAB[C]// 23rd ACM International Conference on Multimedia, Brisbane, Australia, 2015: 689－692.

[30] Lécun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278－2324.

[31] 李辉，石波. 基于卷积神经网络的人脸识别算法[J].软件导刊, 2017,16(3):26－29.

[32] 宋怀波，张卫园，张欣欣，等. 基于模糊集理论的苹果表面阴影去除方法[J]. 农业工程学报，2014，30(3)：135－141.
Song Huaibo, Zhang Weiyuan, Zhang Xinxin, et al. Shadow removal method of apples based on fuzzy set theory[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(3): 135－141. (in Chinese with English abstract)

Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks

Fu Longsheng[1,2], Feng Yali[1], Elkamil Tola[3], Liu Zhihao[1], Li Rui[1], Cui Yongjie[1,2]

(1. *College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling* 712100*, China*;
2. *Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture, Yangling* 712100*, China*;
3. *Precision Agriculture Research Chair, King Saud University, Riyadh* 11451*, Saudi Arabia*)

Abstract: China is the largest country for cultivating kiwifruit, and Shaanxi Province provides the largest production, which accounts for approximately 70% of the production in China and 33% of the global production. Harvesting kiwifruit in this region relies mainly on manual picking which is labor-intensive. Therefore, the introduction of robotic harvesting is highly desirable and suitable. The fast and effective recognition of kiwifruit in the field under natural scenes is one of the key technologies for robotic harvesting. Recently, the study on kiwifruit recognition has been limited to a single cluster and multi clusters in the field have seldom been considered. In this paper, according to growth characteristics of kiwifruit grown on sturdy support structures, an RGB (red, green, blue) camera was placed around 100 cm underneath the canopy so that kiwifruit clusters could be included in the images. We proposed a kiwifruit image recognition system based on the convolutional neural network (CNN), which has a good robustness avoiding the subjectivity and limitation of the features selection by artificial means. The CNN could be trained end to end, from raw pixels to ultimate categories, and we optimized the critical structure parameters and the training strategy. Ultimately, the network was made up of 1 input layer, 3 convolutional layers, 2 sub-sampling layers, 1 full convolutional layer, and 1 output layer. The CNN architecture was optimized by using batch normalization (BN) method, which normalized the data distribution of the middle layer and the output data, accelerating the training convergence and reducing the training time. Therefore, the BN layers were added after the 1, 3 and 5th convolutional layer (Conv1, Conv3, and Conv5 layer) of the original LeNet network. The size of all convolutional kernels was 5×5, and that of all the sub-sampling layers was 2×2. The feature map numbers of Conv1, Conv3, and Conv5 were 6, 16 and 120, respectively. After manual selection and normalizing, the RGB image of kiwifruit was transferred into a matrix with the size of 32×32 as the input of the network, stochastic gradient descent was used to train our models with mini-batch size of 100 examples, and momentum was set as 0.9. In addition, the CNN took advantages of the part connections, the weight sharing and Max pooling techniques to lower complexity and improve the training performance of the model simultaneously. The network used rectified linear units (ReLU) as activation function, which could greatly accelerate network convergence. The proposed model for training kiwifruit was represented as 32×32-6C-2S-16C-2S-120C-2. Finally, 100 images of kiwifruit in the field (including 5918 fruits) were used to test the model, and the results showed that the recognition ratios of occluded fruit, overlapped fruit, adjacent fruit and separated fruit were 78.97%, 83.11%, 91.01% and 94.78%, respectively. The overall recognition rate of the model reached 89.29%, and it only took 0.27 s in average to recognize a fruit. There was no overlap between the testing samples and the training samples, which indicated that the network had a high generalization performance, and the testing images were captured from 9 a.m. to 5 p.m., which indicated the network had a good robustness to lightness variations. However, some fruits were wrongly detected and undetected, which included the fruits occluded by branches or leaves, overlapped to each other and the ones under extremely strong sunlight. Particularly, 2 or more fruits overlapped were recognized as one fruit, which was the main reason to the success rate not very high. This phenomenon demands a further research. By comparing with the conventional methods, it suggested that the method proposed obtained a higher recognition rate and better speed, and especially it could simultaneously identify multi-cluster kiwifruit in the field, which provided significant support for multi-arm operation of harvesting robotic. It proves that the CNN has a great potential for recognition of fruits in the field.

Keywords: image processing; image recognition; algorithms; deep learning; convolutional neural network; kiwifruit