Overlapping citrus segmentation and reconstruction based on Mask R-CNN model and concave region simplification and distance analysis

Xiong Longye¹, Wang Zhuo¹, Liao Haishen¹, Kang Xilong¹, Yang Changhui¹-²*
¹College of Mechanical Engineering, Chongqing University of Technology, Chongqing 400054, China
²College of Mechanical Engineering, Xi’an Jiaotong University, Xi’an 710049, China
*Corresponding author’s e-mail: yangchanghui@cqut.edu.cn

Abstract. Accurate segmentation and reconstruction of overlapping citrus target contours is the primary problem for picking robots. In view of the poor effect of existing research methods on the segmentation and reconstruction of overlapping citrus fruit target contours under complex background, a segmentation and reconstruction method based on region simplification and distance analysis is proposed. Firstly, the overlapping citrus region (Mask region) is obtained by using the previously trained Mask R-CNN model. Then, the convex hull curve of the region is obtained by the roll-wrapped convex shell algorithm, and the region enclosed by the convex curve and the Mask region are pixel-operated. The concave region is polygon-simplified; then the vertices of the polygon are extracted by the Shi-Tomasi corner detection algorithm, and the contour segmentation points are determined by analysing the distance from each vertex to the contour convex-hub curve. Finally, the segmentation contour is reconstructed by the least squares fitting method. The experimental results show that the average error of the proposed method for the reconstruction of overlapping citrus contours is 3.21%, the non-coincidence degree and time are 4.13% and 0.273s respectively, which superior to RANSAC algorithm and Hough transform algorithm. It can satisfy the recognition requirements of overlapping citrus in natural environment for citrus picking robots.

1. Introduction
Citrus is one of the fruits with a large annual output in China, and it is also a fruit with a large amount of trade in the world. At present, citrus picking is mainly based on artificial methods, with problems such as seasonality, labor intensity and high cost. According to incomplete statistics, the cost of artificial picking accounts for about 30% ~ 40% of the total cost of citrus cultivation [1]. The picking operation can effectively reduce the intensity of manual work and reduce the production cost by picking robots instead of manual picking. The accurate identification of overlapping fruits in the natural environment is still one of the problems faced by intelligent picking [2].

A lot of researches have been carried out on the segmentation and reconstruction of overlapping fruit target contours at home and abroad. Chinchuluun R used color features to segment citrus images, and used marker-controlled watershed algorithm to segment overlapping fruit targets[3]. Rizon M used texture analysis, morphological operations, random Hough transform to identify mango, and ellipse fitting and centering the overlapping fruit targets[4]. Xie Zhonghong proposed a method for rapid localization and overlapping fruit detection based on pit search[5]. Miao Zhonghua proposed a
combined optimization algorithm for image recognition and boundary segmentation of overlapping fruits in natural environment[6]. Peng Hui proposed an overlapping fruit segmentation algorithm based on parallax images to preserve the contour information of uncovered apples[7]. Xiang Rong proposed an overlapping tomato recognition method based on edge curvature analysis. By eliminating the point of curvature anomaly of the contour edge, the missing contour is fitted and reconstructed by the method of circular regression[8]. Through the variation law of chain code difference, Feng J et al. [8] used the principle of local optimum to find effective pits and adopted effective pits to achieve the segmentation of overlapping apples[9]. Song Huaibo used the convex hull-based method to segment overlapping apples, and used Spline interpolation algorithm to reconstruct the contour of overlapping occluded fruits[10]. Lu J fusion the color difference information and normalized RGB model, through the length, curvature and convexity and other three indicators to filter the effective segment of the contour, and use the ellipse to fit the effective segment to achieve the reconstruction of the occluded fruit contour[11]. Xu Yue used the corner detection algorithm to find the real segmentation point of overlapping apples, and realized the segmentation of overlapping fruits by using the real segmentation points[12]. Wang Dandan used the K-means clustering algorithm and Ncut spectral clustering algorithm segmented the overlapping apples, and reconstructed the occlusion contour by Spline interpolation algorithm[13]. Li Yang by extracting the center of the convex region of the overlapping citrus target region and the concave region after the region is subtracted, and then extracting the segmented single target gravity center by distance transformation theory and fitting by ellipse Fit the contour of the occlusion section[14]. The above research methods have many problems such as relatively simple segmentation background, relatively stable light changes, small number of overlapping fruits, and little or no occlusion in front of the fruit, and it is difficult to meet the actual needs of the picking robot for overlapping fruit recognition in the natural environment.

With the rapid development of deep learning, many research scholars began to use different network structures to train different recognition models for the identification of various fruits under different scenes, and achieved good results [15-19]. In view of this, an overlapping citrus segmentation and reconstruction method based on Mask R-CNN model and concave region simplification and distance analysis was proposed in this paper. The Mask R-CNN model was trained to identify overlapping citrus using the fabricated overlapping citrus samples, and the identified citrus regions were segmented by the concave region simplification and distance analysis method. The least squares fitting method was used to segment the citrus contours. The contours are fitted to reconstruct the overlapping missing contours.

2. Experimental Materials and Methods

2.1 Experimental materials
The equipment for collecting images in this paper is Kinect v2 as shown in Figure 1, the location of shooting is located in Jinguoyuan, Beibei, Chongqing, China. The citrus varieties are citrus and ponkan. The time of shooting is January 2016 and December 2017, which is the ripening period of citrus. Shooting with multiple angles (Natural light, Back light, Side light), as shown in Figure 2. Among them, 1794 images were acquired at the angle of the Natural light, 1692 images were acquired at the Side light angle, and 1709 images were acquired at the Back light angle.

![Figure 1. Hardware diagram of Kinect v2](image1.png)

![Figure 2. Diagram of shooting angle](image2.png)
2.2 Sample making

In this paper, using Labelme [20] to mark samples. Mark the target with a closed polygon to maximize the target, and add a category label to each polygon to save the marked image as a json format data file containing the marker target fit polygon. Coordinate position information and category information in the picture. The effect of marking the training sample is shown in Figure 3. Figure 3a is the original image, Figure 3b is the effect of fitting the mark by polygon, and Figure 3c is the three-channel mask automatically generated by the Labelme marking tool after the mark. Figure 3c shows a single-channel binarization mask by OpenCV binarization. As shown in Figure 3d, the figure contains the shape information of the object for pixel segmentation training in the Mask R-CNN training network.

![Sample marking effect](image)

2.3 Model training

This paper constructs the Mask R-CNN network structure with ResNet101 and FPN as the backbone network as shown in Figure 4. Accelerated training and testing using the GTX1080Ti under the TensorFlow platform configured in the Ubuntu 16.04 system. The network layer structure configuration of the ResNet101 module is shown in Table 5. The training positive and negative training sample sets are assigned by 3:1. The learning rate in the training parameters is 0.001, the weight attenuation is 0.0001, the batch size is 1, and the IOU threshold is 0.5. The scale of the anchor is designed to be 16, 32, 64, 128, 256. The configuration ratio is 1/2, 1, 2, and the stride is 1, and the ROI area of each picture is 32. The iteration of each step is 100 times, and the total training is 140 times. Iterative training was 14,000 times. The ROI multitasking loss function after sampling in training is:

\[ L = L_{cls} + L_{box} + L_{mask} \]  

In the formula, \( L_{cls} \) and \( L_{box} \) are the same as those defined in Faster R-CNN. For the mask branch, like the other classification branches, the full convolutional network is used to output the K class mask, the output function is the sigmoid function, and finally the output is compared with the pre-set threshold. Binary mask to avoid competition between classes.

![Network structure of Mask R-CNN](image)
indicates the loss rate. It can be seen from the graph that the loss function value tends to be stable 12000 times.

Figure 5. Relation curve between loss rate and iterations

2.4 Concave region acquisition and simplification
Before extracting the overlapping citrus contours, it is necessary to extract the convex hull of the contour. In this paper, the Mask R-CNN is used to obtain the citrus region and the convex hull is extracted by the roll-wrapped convex shell algorithm [21], as shown in Figure 6. The convex hull curve surrounds the citrus completely, which indicates that the convex hull extraction is effective.

The concave region is obtained by performing a pixel operation on the closed region and the citrus region surrounded by the convex shell curve according to the equation (2).

\[
\begin{cases}
    h(x,y) = 255 & \text{if } (f(x,y) = g(x,y)) \\
    h(x,y) = 0 & \text{if } (f(x,y) \neq g(x,y))
\end{cases}
\]  

(2)

In the formula, \( f(x,y) \), \( g(x,y) \), \( h(x,y) \) represent the pixel gradation values at \((x,y)\) of Figure 7a, Figure 7b respectively, and the resulting concave region is as shown in Figure 7c.

The result of polygon simplification and vertex detection of the obtained concave region as shown in Figure 8, Figure 8a is the convex shell curve extraction result, Figure 8b is the polygon simplified result of the concave region, and Figure 8c is the polygon extracted by the Shi-Tomasi [22] corner detection operator vertex.
2.5 *Distance analysis from point to citrus contour*

After extracting the vertices of the polygon, map them to the citrus contour curve and the convex shell curve. The result is shown in Figure 9. Among them, the black line is the contour curve, and the blue line is the convex shell curve.

![Figure 9. Result of polygon vertex mapping](image)

The algorithm flow for calculating the distance from the corner point to the citrus convex contour curve is as follows:

1) Get all the pixel coordinates \((x_i, y_i)\) on the contour and divide the contour into many infinitesimal fragments \(S_i\).

2) Find the point with the smallest horizontal coordinate value from all the pixels (when there are multiple pixel points with the same abscissa value, select the point with the smallest ordinate value) as the starting point of the contour \(Q(x_0, y_0)\), and the contour point sequence direction is set to be counter clockwise.

3) Starting from the starting point \(Q\), in the counter clockwise direction, all the points on the contour convex hull curve and the coordinates of the corner points are sequentially recorded as \(A(x_i, y_i)\) and \(P(x_p, y_p)\), and the corner points are calculated according to formula (3). The distance to the convex hull curve and recorded as \(d_i\)

\[
d_i = \frac{[(y_p - y_i)(x_{i+1} - x_i) - (x_p - x_i)(y_{i+1} - y_i)]}{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}
\]  

(3)

In order to obtain the segmentation points of overlapping citrus contours, the distance from the apex to the convex hull curve was analysed. A large number of experimental results show that in order to reduce the influence of noise caused by contour segmentation, the minimum threshold of \(d\) is 2 pixels, that is, when \(d > 2\), the point is the contour segmentation point. According to the above principle, with respect to Figure 9, the results after distance screening are shown in Figure 10, Figure 10a is the effect before the distance screening, Figure 10b is the result after the distance screening, and Figure 10c is the segmentation effect of the citrus contour at the dividing point.

![Figure 10. Overlapping citrus contour segmentation](image)
2.6 Least squares circle fitting based on maximum contour optimization

In order to restore the missing outline of overlapping citrus, this paper reconstructs the citrus contour based on the least square method of round fit interpolation algorithm [23] and optimizes it by the maximum contour. The results are shown in Figure 11. Figure 11a is the original image, Figure 11b is the citrus outline, and Figure 11c is the contour reconstruction effect.

3. Experimental results and analysis

In order to verify the effectiveness of the method, segmentation and reconstruction experiments of overlapping citrus were carried out in orchard environment. The results are shown in Figure 12. Figure 12a is the original image, Figure 12b is the citrus region identified by the Mask R-CNN, Figure 12c is the outline of the citrus region, and Figure 12d is the effect of the contour reconstruction.

It can be seen that the method of in this paper not only accurately segment the overlapping citrus, but also accurately reconstruct the contour of the overlapping region, which can fully satisfy the
intelligent identification requirement of the unshielded overlapping fruit in the natural environment. In order to compare with other methods, the area of the citrus target that is not obscured by other fruits and the region of the reconstructed citrus target are obtained respectively. Through comparing the contour reconstruction error, un superpose degree and time of reconstruction between the method of this paper and RANSAC algorithm and Hough transform algorithm to evaluate the effect of segmentation and reconstruction and efficiency for overlapping citrus.

The calculation formulas for the contour reconstruction error $e$ and the un superpose degree $\varepsilon$ are:

$$e = \frac{|S - S_i|}{S} \times 100\% \quad (4)$$

$$\varepsilon = \frac{|S \cup S_i - S \cap S_i|}{S} \times 100\% \quad (5)$$

$S \cup S_i$ represents the number of pixels in the original image citrus region or on the reconstructed citrus target; $S \cap S_i$ represents the number of pixels on both the original image citrus region and the reconstructed citrus target; the difference between the two pixels In order not to coincide with the area $S_i$ of the citrus region, the lower the degree of non-coincidence, the better the segmentation reconstruction effect.

20 groups of unobstructed overlapping mature citrus in the orchard environment were selected for segmentation and reconstruction. The method of this paper, RANSAC algorithm and Hough transform algorithm to achieve the contour reconstruction error, non-coincidence and reconstruction time of overlapping citrus are shown in Figure 13.
The average reconstruction error of the RANSAC algorithm for reconstructing overlapping citrus contours is 4.10%, the average non-coincidence is 5.17%, and the average reconstruction time is 0.512 s. The average reconstruction error of Hough transform algorithm for reconstructing overlapping citrus contours is 14.67%, and the average non-coincidence The average reconstruction time is 16.90% for the 16.74%, the average reconstruction error for the overlapping citrus contours is 3.21%, the average non-coincidence is 4.13%, and the average reconstruction time is 0.273s, which is significantly lower than the RANSAC algorithm and Hough transform algorithm. It shows that the method of contour segmentation and reconstruction in this paper for unobstructed overlapping citrus is effective.

4. Conclusion
This paper proposes a method of overlapping citrus segmentation and reconstruction based on Mask R-CNN model and concave region simplification and distance analysis. Accurate segmentation and reconstruction of overlapping citrus is achieved through the combination of deep learning, image processing and planar geometry.

1) The recognition model of overlapping citrus using deep learning training with ResNet101 and FPN as the backbone network is beneficial to overcome the influence of light changes and leaf occlusion in the natural environment, and can accurately identify different types of overlapping citrus.

2) A segmentation method of overlapping citrus based on concave region simplification and distance analysis is proposed. The segmentation points of overlapping citrus target contours can be accurately found to complete the effective segmentation of overlapping citrus contours.

3) Compared with the traditional RANSAC algorithm and Hough transform algorithm, the method of this paper has higher precision and faster speed for the reconstruction of overlapping citrus contours.

Acknowledgments
This research is funded by the Special Project of Key Technological Innovation (No. CSTC2015zdcy-ztzx7003) and research program of basic research and frontier Technology (No. CSTC2016jcyjA0444) received from Chongqing Science and Technology Commission.

References
[1] Shen, Z.M. (2016) Current situation of Citrus production during the 12th Five-Year Plan and Development Trend and Countermeasure of the 13th Five-Year Plan in China[J]. Scientific breeding, 12: 5 – 8.
[2] Cai, J.R., Zhou, X.J., Li, Y.L., Fan, J. (2008) Recognition of mature oranges in natural scene based on machine vision[J]. Transactions of the CSAE, 24: 175 – 178.

[3] Chinchuluun, R., Lee, W.S. (2006) Citrus Yield Mapping System in Natural Outdoor Scenes using the Watershed Transform. In: 2006 ASABE Annual International Meeting. Portland. pp. 1 – 12.

[4] Rizon, M., Yusri, N.A.N., Kadir, M.F.A., Mamat, A.R.B., Aziz, A.Z.A., Nanaa, K. (2015) Determination of mango fruit from binary image using randomized Hough transform. In: Eighth International Conference on Machine Vision. Barcelona. pp. 6427 – 6436.

[5] Xie, Z.H., Ji, C.Y., Guo, X.Q., Zhu, S.X. (2011) Detection and Location Algorithm for Overlapped Fruits Based on Concave Spots Searching[J]. Transactions of the Chinese Society for Agricultural Machinery, 42: 191–196.

[6] 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[J]. Transactions of the Chinese Society for Agricultural Machinery, 47: 21–26.

[7] Peng, H., Wu, P.F., Zhai, R.F., Liu, S.M., Wu, L.L., Jing, X. (2012) Image Segmentation Algorithm for Overlapping Fruits Based on Disparity Map[J]. Transactions of the Chinese Society for Agricultural Machinery, 43: 167-173.

[8] Xiang, R., Ying Y.B., Jiang H.Y., Rao X.Q., Peng Y.S. (2012) Recognition of Overlapping Tomatoes Based on Edge Curvature Analysis, 43: 157-162.

[9] Feng, J., Wang, S., Liu, G., Zeng, L.H. (2012) A Separating Method of Adjacent Apples Based on Machine Vision and Chain Code Information. In: International Conference on Computer and Computing Technologies in Agriculture. Beijing. pp. 258-267.

[10] Song, H.B., Zhang, C.D., Pan, J.P., Yin, X., Zhuang, Y.B. (2013) Segmentation and reconstruction of overlapped apple images based on convex hull[J]. Transactions of the Chinese Society of Agricultural Engineering, 29: 163–168.

[11] Lu, J., Sang, N. (2015) Detecting citrus fruits and occlusion recovery under natural illumination conditions[J]. Computers & Electronics in Agriculture, 110: 121-130.

[12] Xu, Y., Li, Y.H., Song, H.B., He, D.J. (2015) Segmentation method of overlapped double apples based on Snake model and corner detectors[J]. Transactions of the Chinese Society of Agricultural Engineering, 31: 196 – 203.

[13] Wang, D.D., Xu, Y., Song, H.B, He, D.J., Zhang, H.H. (2015) Fusion of K-means and Ncut algorithm to realize segmentation and reconstruction of two overlapped apples without blocking by branches and leaves[J]. Transactions of the Chinese Society of Agricultural Engineering, 31: 227 – 234.

[14] Li, Y., Yang, C.H., Hu, Y.C., Zhang, H., Yang, Y. (2018) Overlapping citrus target recognition and localization method based on convex shell and distance transformation[J]. Modern Manufacturing Engineering, 9: 82-87.

[15] Sa, I., Ge, Z., Feras, D., Ben, U., Tristan, P., Chris, M. C. (2016) Deep Fruits: A fruit detection system using deep neural networks[J]. Sensors, 16: 1222-1228.

[16] Bargoti, S., Underwood, J.P. (2017) Image segmentation for fruit detection and yield estimation in apple orchards[J]. Journal of Field Robotics, 34: 1039 – 1060.

[17] Zhou, Y.C., Xu, T.Y., Zhen, W., Deng H.B. (2017) Classification and recognition approaches of tomato main organs based on DCNN[J]. Transactions of the Chinese Society of Agricultural Engineering, 33: 219 – 226.

[18] Fu, L.S., Feng, Y.L, Elkamil, T., Liu, Z.H., Li, R., Cui, Y.J. (2018) Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks[J]. Transactions of the Chinese Society of Agricultural Engineering, 34: 205–211.

[19] Xiong, J.T., Liu, Z., Tang, L.Y., Lin, R., Bu, R.B., Peng H.X. (2018) Visual detection technology of green citrus under natural environment[J]. Transactions of the Chinese Society for
Agricultural Machinery, 49: 45—52.

[20] Peng, H.X., Huang, B., Shao, Y.Y., Li, Z.S, Zhang, C.W., Chen, Y., Xiong, J.T. (2018) General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the Chinese Society of Agricultural Engineering, 34:155-162.

[21] Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T. (2008) LabelMe: A database and web-based tool for image annotation [J]. International Journal of Computer Vision, 77: 157 -173.

[22] Shi, J., Tomasi, C. (2002) Good Features to Track In: IEEE Conference on Computer Vision & Pattern Recognition. Seattle pp. 1 – 10.

[23] Zhou, Q.H. (2007) On an Isomorphic Direction of Improving and Optimizing an Algorithm for Determining the Convex Hull of 2D Point Set or Line Segment Set[J]. Computer Science, 34: 216-218.