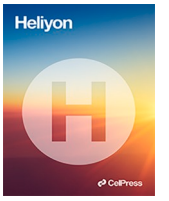
Heliyon 9 (2023) e15097

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Application of an improved watershed algorithm based on distance map reconstruction in bean image segmentation

Hongquan Liu [a,1], Weijin Zhang [b,1], Fushun Wang [b,c,**], Xiaohua Sun [d], Junhao Wang [b], Chen Wang [b], Xinxin Wang [e,f,g,*]

[a] College of Urban and Rural Construction, Hebei Agricultural University, Baoding, 071000, China
[b] College of Information Science and Technology, Hebei Agricultural University, Baoding, 071000, China
[c] Hebei Key Laboratory of Agricultural Big Data, Baoding, 071000, China
[d] Department of Digital Media, Hebei Software Institute, Baoding 071000, China
[e] Agricultural Technology Innovation Center in Mountainous Areas of Hebei Province, Baoding, 071000 China
[f] Agricultural Engineering Technology Research Center of National North Mountainous Area, Baoding, 071000, China
[g] Hebei Agricultural University Hebei Mountain Research Institute, Baoding, 071000, China

ARTICLE INFO

Keywords:
Image segmentation
Watershed algorithm
Bean image
Seed phenotype

ABSTRACT

As an important step in image processing, image segmentation can be used to determine the accuracy of object counts, and area and contour data. In addition, image segmentation is indispensable in seed testing research. Due to the uneven grey level of the original image, traditional watershed algorithms generate many incorrect edges, resulting in oversegmentation and undersegmentation, which affects the accuracy of obtaining seed phenotype information. The DMR-watershed algorithm, an improved watershed algorithm based on distance map reconstruction, is proposed in this paper. According to the grey distribution characteristics of the image, the grey reduction amplitude h was selected to generate the mask image with the same grey distribution trend as that of the original image. The original greyscale map was reconstructed with corresponding thresholds selected according to the false minima of different regions that are to be segmented, which generates an accurate distance map that eliminates the wrong edges. An adzuki bean (*Vigna angularis* L.) image was selected as the experimental material and the residual rate of the segmentation counting results of each algorithm was investigated in two cases of two-particle adhesion and multiparticle adhesion. The results of the proposed algorithm were compared with those of the traditional watershed algorithm, edge detection algorithm and concave point analysis algorithm which are commonly used for seed segmentation. In the case of two-particle adhesion, the residual rates of the watershed algorithm and edge detection algorithm were 0.233 and 0.275, respectively, while the residual rate of the concave point analysis algorithm was 0 which proved to be suitable for two-particle adhesion. In the case of multiparticle adhesion, the concave point analysis algorithm was not applicable because it would destroy the seed image. The residual rates of the watershed algorithm and edge detection algorithm were 0.063 and 0.188, respectively, while the residual rate of the proposed algorithm in the two-particle adhesion cases was 0 and the counting accuracy reached 100%, which proved the effectiveness of the proposed algorithm. The

* Corresponding author. Agricultural Technology Innovation Center in Mountainous Areas of Hebei Province, Baoding, 071000 China.
** Corresponding author. College of Information Science and Technology, Hebei Agricultural University, Baoding, 071000, China.
E-mail addresses: lhq@hebau.edu.cn (H. Liu), zhweijin1114@qq.com (W. Zhang), xxwfsh@hebau.edu.cn (F. Wang), xinxinwang.wur@qq.com (X. Wang).
[1] Co-first author.

https://doi.org/10.1016/j.heliyon.2023.e15097
Received 19 October 2022; Received in revised form 22 March 2023; Accepted 27 March 2023
Available online 15 April 2023

algorithm in this paper significantly improves the accuracy of image segmentation of adherent seeds, and provides a new reference for image segmentation processing in seed testing.

1. Introduction

The seed quality of crops directly affects their growth and final yield. Thus, it is necessary to test the seeds. At present, the adzuki bean were tested by manual measurement to evaluate its quality, subjective judgment and discrimination based on the coarse-grained national standards combined with personal experience. This method is subjective, arbitrary, inefficient and repetitive. Due to the lack of scientific and objective evaluation criteria, different people will have deviation in the judgment conclusions on the quality of the same sample, resulting in rating uncertainty [1].

Image processing technologies can rapidly analyze seed images and accurately obtain phenotypic data such as the shape, color and texture, and will not cause any damage to the seeds [2,3]. Therefore, the application of image processing in seed testing can improve the efficiency of seed testing to a certain extent [4]. Obtaining seed phenotypes based on image processing technologies can help farmers quickly determine seed quality to increase growing efficiency, which has become a popular research topic [5,6]. As an important step in image processing, image segmentation is indispensable in seed testing research. Most of the seed images collected contain some adhesive areas between seeds and if these adhesions are not processed properly, it will lead to errors in area and perimeter calculation, affecting the seed feature extraction [7,8]. To better extract the characteristics of a single seed, it is very important to divide the adhesive seeds into single seeds [9–11].

Among many segmentation algorithms, the watershed algorithm, edge detection algorithm and concave point analysis algorithm are the most used in seed image processing [9,12]. The watershed algorithm has a good response to the weak edge, which can ensure that the seed edge is closed and continuous and can theoretically achieve a good segmentation effect. However, the noise in the image and the tiny grey-level changes of the seed surface will produce the phenomenon of oversegmentation [13]. Its processing object is the grayscale distance map, forming the dividing line between two different regions with similar greyscale distances, while the adhesive region of the seed is not segmented because it is in the same region with similar greyscale gradients, resulting in undersegmentation phenomenon. The edge detection algorithm can strengthen the correct edge, but it has a poor ability to process complex adhesions [14]. Introducing this algorithm into the watershed algorithm can suppress its response to weak edges and preserve the correct edges. The concave point analysis algorithm can identify the concave point position of the convex hull graph and has good binary segmentation ability, but the effect is not good in the case of complex adhesiveness, and even destroys the seed image. However, by introducing the concave point analysis algorithm into the watershed algorithm, the adhesion position can be identified, and the undersegmentation phenomenon can be effectively alleviated.

Image processing has been applied in the research of other crop seed testing. There are also many valuable researches on the defects of watershed algorithm. The kernel of rape (*Brassica campestris* L.) obtained by selective limit corrosion is used as a marker to limit the number of watershed zones and impose a forced minimum on the original gradient image [15]. The kernel count of adhesion corn (*Zea mays* L.) kernel superimposed image can restrain the oversegmentation degree of watershed algorithm effectively [16]. The adaptive radius circle template is used to detect the corner of the adhesion of rice (*Oryza sativa* L.) image. By detecting whether the end point of the dividing line overlaps with the corner point, the undersegmentation area is recognized and merged effectively, and the error caused by the watershed algorithm is reduced [17]. The minimum transformation and filtering operation were carried out for the parts below the threshold determined according to the kernel radius, and then watershed transformation was carried out to overcome the difficulty of high-density adhesion, and the soybean (*Glycine max* Merr.) kernel with dense adhesion was accurately segmented. The mark-controlled watershed algorithm and the region threshold method were used to calculate the number of seeds in contact with each other, including Indica rice, adzuki bean, mung bean (*Vigna radiata* L.), soybean (*Glycine max* L.), peanut (*Arachis hypogaea* L.) and sesame (*Sesamum indicum* L.), and the error rate was about 3% [18].

In order to reduce the damage to seeds and the influence of subjective factors on seed rating, image processing technology was introduced to improve the seed testing efficiency of legumes. To solve the problem of oversegmentation and undersegmentation of the

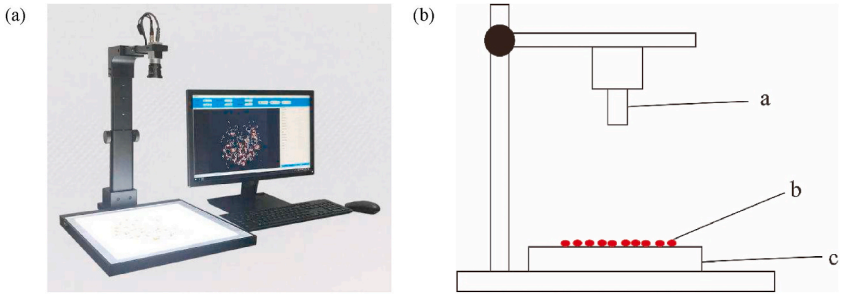


Fig. 1. (a) Diagram of the experimental setup. (b) Schematic diagram of the device mainly including a: the RGB three-channel camera used to capture images, b: the bean seeds and c: the device stage where the bean seeds were placed.

watershed algorithm, an improved watershed algorithm based on distance map reconstruction is proposed in this paper to improve the segmentation efficiency of bean seeds. The grey gradient of the local problem area of the image is adjusted, the distance map is re-generated, and the correct watershed is retained. Then edge detection and concave point analysis are combined to retain a more perfect seed image to achieve an accurate counting effect, and improve the efficiency of seed examination of bean crops such as adzuki bean.

2. Materials and methods

2.1. Experimental environment

The experimental setup including loading platform, backlight plate, camera and bracket and schematic diagram are shown in Fig. 1 (a).

Fig. 1(b) is the schematic diagram of the device. In the experimental environment, the seeds were put on the stage and the image was captured by the three-channel camera and sent back to the computer.

2.2. Experimental materials

The experimental materials were the images of adzuki beans including Baihong11, Jihong14, Pinhong2014-166, Jihong352, Yuhong2, Yuhong4, and Yuhong5, for a total of 28 images.

2.3. Image acquisition

For image acquisition, 40, 60, 80, and 100 adzuki beans of the same variety were placed on positioning paper with 5 × 5 mm anchor points of filling colors, with random adhesion among each other, and images of the seeds were collected in the experimental environment. The positioning paper was marked with an anchor point of 5 × 5 mm filling color and the image pixel was 2048 × 1536. The brightness of the light source and the focal length of the camera were adjusted. The seeds were placed into the device to collect RGB images which would be transmitted back to the computer.

2.4. Image preprocessing

To facilitate segmentation, image preprocessing was first needed. The image was greyed to reduce the image information. The influence of color information on image processing was fully considered. According to the three-channel value of $I_R(x, y)$, $I_G(x, y)$ and $I_B(x, y)$ of each pixel in the original image, the weighted greying algorithm was used to grey process the image. The grey value $I(x, y)$ of each pixel of the image was calculated through

$$I(x,y) = 0.3 \times I_R(x,y) + 0.59 \times I_G(x,y) + 0.11 \times I_B(x,y) \quad (1)$$

R, G and B in (1) represented the three-dimensional component in the RGB model where the image was located. At this time, there were still some fuzzy and some unclear details in the image, so the median filter was used to smooth the image. A square median filter was used to sort the grey value of all pixels in the region from small to large, and the pixel of the center point was the average value of all points. Then, according to the grey information of the image, the foreground and background of the image were preliminarily segmented and the segmentation threshold needed to be determined. To reduce the influence of subjective factors on the segmentation, the Otsu algorithm was used to calculate the threshold automatically and the foreground containing most seeds was retained. There was noise in the image due to the shadow caused by the light of the backlight plate. This noise could be removed by a morphological operation that could also split the slight adhesion. The open operation under the premise of protecting the image and the loss of foreground pixels could also be reduced.

Fig. 2(a) is a part of the more serious adhesion of the adzuki bean original image, and there were a lot of adhesions among the beans. Image preprocessing result of adzuki beans was shown in Fig. 2(b). There were small holes in some seeds in the binary image because the color of the umbilicus in the adzuki beans was approximately white, which can cause the umbilicus to be misclassified as background during binarization. After achieving segmentation, it was necessary to fill these holes in a timely manner to obtain more accurate seed data when continuing to calculate phenotypic information such as the seed shape and color.

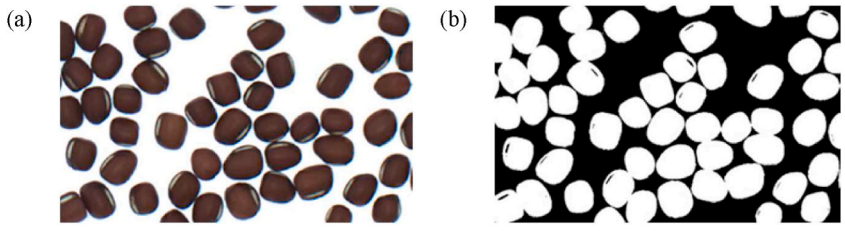


Fig. 2. (a) Adzuki bean original image. (b) Adzuki bean binary image. Image preprocessing results of adzuki beans.

2.5. Algorithm implementation

Generally, the watershed algorithm can manage the adhesion among the seeds well, but oversegmentation and undersegmentation are produced, and the segmentation accuracy is reduced. The oversegmentation is caused by the wrong edge generated by the watershed algorithm inside the seed [19,20]. The edge detection algorithm can identify the real edge and eliminate the wrong edge, which can alleviate the oversegmentation of the watershed algorithm to a certain extent. The concave point analysis algorithm can recognize the concave point of the convex hull image and divide it into two parts. It can achieve good segmentation for the two-particle adhesion. If there is still adhesion in the segmentation result of the watershed algorithm, the concave point analysis algorithm can recognize adhesion and alleviate the undersegmentation. In the following, the three algorithms were fused, and the improved watershed algorithm based on distance map reconstruction was proposed to compare and analyze the segmentation effect of adzuki beans.

2.5.1. Watershed algorithm

The watershed algorithm is based on the topological landform of geodesy, and the grey distribution of the image is regarded as the elevation distribution. Injecting water from the lowest point in each catchment basin, as the water level rises, dams are formed at the confluence of adjacent catchment basins to form different waters. The division of different regions is achieved for the image as shown in Fig. 3.

For the segmentation of adzuki bean seeds, the purpose is to divide each seed into a separate area, which requires that the grey level inside the seed is uniform and there is only one local minimum in the contour of a seed. In fact, because the color of the seed coat is not uniform, there will be a false minimum and the oversegmentation would occur inside each seed. Some bean umbilici were misclassified as background in binarization as shown in Fig. 4 due to the large difference in color between the umbilicus and seed coat, resulting in certain segmentation errors.

2.5.2. Edge detection algorithm

Canny operator edge detection is one of the most widely used edge detection algorithms [21], and its calculation steps are as follows.

i. Smooth the image with a Gaussian filter and eliminate excess noise.
ii. Calculate the gradient intensity and direction of each pixel in the image.
iii. Suppress the non-maximum pixels to eliminate spurious responses.
iv. Detect the real and potential edges with double threshold detection [22].
v. Suppress isolated weak edges.

The double threshold includes the upper and lower threshold, which need to be artificially set. If the grey level of the pixel is larger than the upper threshold, it will be considered an edge. If it is smaller than the lower threshold, it will be considered not to be an edge. That is, the lower threshold is used to control the edge connection and the upper threshold is used to control the initial segmentation of strong edges. The grey levels of the pixels between the double thresholds are convolved with the Sobel operator template

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{2}$$

to calculate the first difference, partial derivative, gradient value and direction. G_x and G_y in (2) are the gradient operators in the x direction and the y direction. Relatively, the edge is thinner and the error is smaller. The resulting image is a single channel binary image without further segmentation [23]. Therefore, the segmentation effect of the Canny operator is better, which also proves the ability of introducing it into watershed algorithm to reduce false edges. At the same time, the contour of the seed umbilicus is

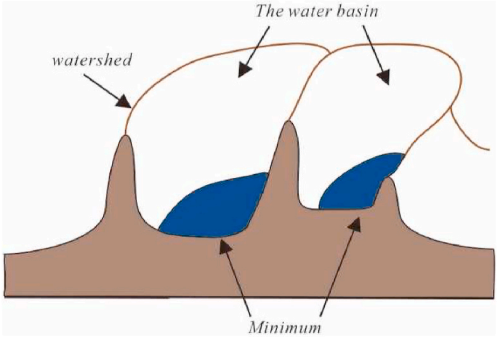


Fig. 3. Principal diagram of the watershed algorithm. Black region: water rising from the minimum point.

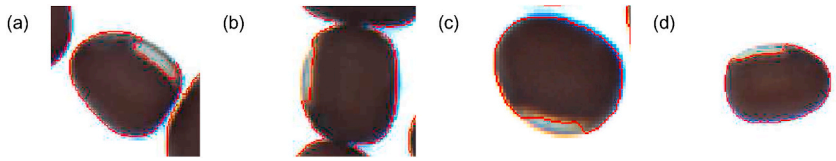


Fig. 4. Due to the large color difference between bean umbilicus and seed coat, some bean umbilici were misclassified as background during binarization. (a) The umbilicus of a non-cohesive adzuki bean is misclassified. (b) The umbilicus of a closely adhesive adzuki bean is misclassified. (c) and (d) The umbilicus of the single adzuki bean arranged in different directions is misclassified.

preserved, which preserves relatively complete original image information for the study of the seed umbilicus.

2.5.3. Concave point analysis algorithm

The concave point analysis algorithm includes two processes, concave point detection and matching. The adhesion position is identified and extracted, the real concave point is calculated by analyzing the convex hull, and the corresponding concave points are matched and connected according to certain rules to achieve the segmentation of the adhesion area [24]. The calculation steps are as follows.

i. Extract the seed binary image contour and obtain the minimum convex closure of the current contour using Graham scanning.
ii. Subtract the convex closure graph from the binary image to obtain the concave region.
iii. Extract the contour of the concave region.
iv. Traverse the two regions and search for the nearest point between the two regions to label as a concave point.
v. Draw the line to achieve segmentation based on the two concave points.

For adzuki bean images with only two seeds of adhesion, accurate segmentation can be achieved, which also proves that the

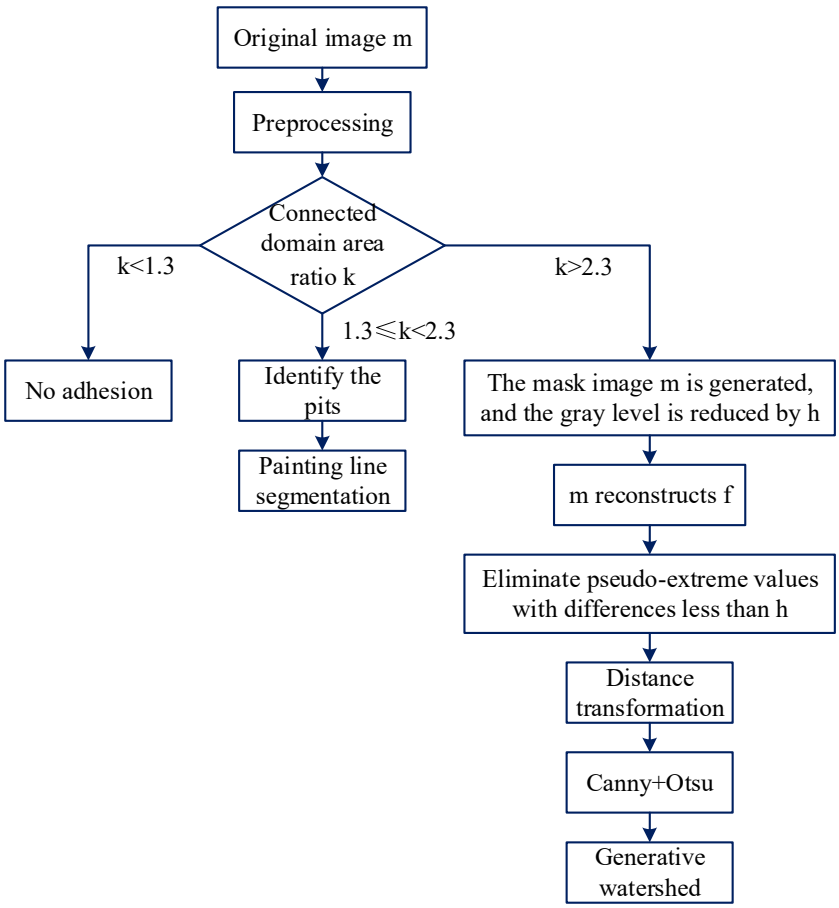


Fig. 5. Algorithm flow chart.

concave point analysis algorithm can be used to segment the under-segmented region left by the watershed algorithm.

2.5.4. Improved watershed algorithm based on distance graph reconstruction

Firstly, the adhesion condition of the seeds was determined, and a more suitable algorithm was selected. The reason of segmentation error caused by the watershed algorithm was the uneven greyscale of the image and the pseudo minima. Reconstructing the distance graph to eliminate the pseudo minimum could reduce the occurrence of oversegmentation and improve the segmentation accuracy. The algorithm process is shown in Fig. 5.

2.5.4.1. Determination of the seed number. The position of adhesion is located in the connected region with a large area, so the edge processing of the connected region is particularly important. The number of seeds contained in the current connected region was determined, the area of the connected domain was calculated, and the ratio k was solved with the average value of the five smallest areas in the prospective area set. The number of seeds contained in the connected domain was determined according to

$$f(x)=\begin{cases}1, \mathrm{k}<1.3\\ 2, 1.3\le \mathrm{k}<2.3\\ n, \mathrm{k}\ge 2.3\ (\mathrm{n}>2)\end{cases} \tag{3}$$

The calculation formula of k in (3) was:

$$\overline{S_{\min}}=\sum_{i=1}^{5} S_i \Big/ 5 \tag{4}$$

$$k=\frac{S_i}{\overline{S_{\min}}} \tag{5}$$

S_i (i = 1,2,3,4,5) in (4) and (5) is the area values of the five smallest connected domains in the current foreground area set and represents the current seed area and i indicates the serial number of the current seed. $\overline{S_{min}}$ in (4) and (5) is the average value of the five smallest areas. The value range of k is defined according to the statistical calculation of the image area of the current variety. According to the statistics of the image area of all varieties, the area of single seed is less than 1.3 times of $\overline{S_{min}}$, and the area of connected domain of adhesion between two seeds is generally less than 2.3 times of $\overline{S_{min}}$. When the ratio k exceeds 1.3, adhesion is proven, and when the ratio exceeds 2.3, there is more than two-particle adhesion. The decision relation is shown in Table 1.

If the number of seeds in the connected region is 1, no segmentation is needed, and the next connected region is traversed. If the number is 2, the concave points of the connected region are detected according to the concave point analysis algorithm and connected according to the shortest Euclidean distance to achieve segmentation. If the number is 3 or more, further operations are needed.

2.5.4.2. Distance graph reconstruction. The images m and f are set to be two greyscale images defined in the same discrete region D, whose value range is the discrete set [25] and $m \le f(\forall p \in D, m(p) \le f(p))$, using m to reconstruct f, can be expressed as

$$\forall p \in D, \rho_f(m)(p)=\max\left(k \middle| p \in p_{T_k(f)}(T_k(m)), k \in \{0, n-1\}\right) \tag{6}$$

The identification image m in (6) satisfies $m(p) = f(p) - h, p \in D$, and h is the preset constant, which refers to the decrease in the pixel grey value of Image f, as shown in Fig. 6. $T_k(f) = \{p | p \in D, f(p) \ge k\}$ and $f(p)$ are the grey values of Point p in Image f.

In the interval of Adzuki bean 2 in Fig. 6, dividing f with two minima, a and b, directly by the watershed algorithm will form an edge at c. To suppress the watershed, it can be seen from Fig. 6 that the minimum b should be eliminated.

The grey value of Image f is reduced by h to obtain the mask Image m. An appropriate threshold μ is selected for the minimum value b, namely the red dashed line in Fig. 6, whose value is equal to the grey value of Point b with a larger grey value There are two intersection points d and e between the threshold line and the grey curve of Adzuki bean 2, and the area between d and e is just a local maximum for forming the minimum b. The grey value of the pixel point between d and e is set to μ to obtain the blue curve. At this time, the minimum b has been eliminated, and the image of Adzuki bean 2 contains only one minimum point, Point a, which is the final reconstructed image.

Using m to reconstruct f, the image obtained can maintain the shape of Image f and eliminate the small difference of the amplitude less than h, which effectively extracted each maximum region from the greyscale map and provided a better distance greyscale map for the watershed algorithm. The selection of Parameter h mainly depends on the distribution characteristics of the pseudo minimum. The

Table 1
The connected domain contains the relation between the number of seeds and k.

k range	The state percentage of connected domain/%		
	Single	Two	Three and above
0 < k < 1.3	100	–	–
1.3 = k < 2.3	–	100	–
k ≥ 2.3	–	–	100

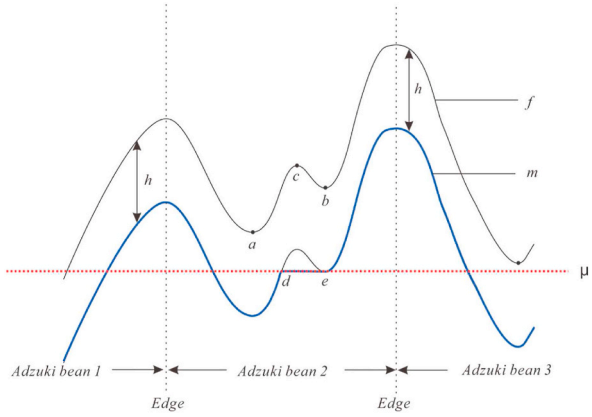


Fig. 6. Distance map reconstruction process.

empirical value was selected according to the particle characteristics of different adhesion or overlap states, and the real regional extreme values were retained while the pseudo minima were removed as much as possible.

2.5.4.3. Boundary strengthening. To protect more edge information to strengthen the boundary, the edge detection result of the Canny operator was shown in Fig. 7(a) and the Otsu segmentation result was shown in Fig. 7(b). Fig. 7(c) was the superimposition of (a) and (b) after distance map reconstruction, such as Baihong11 in Fig. 7.

It can be seen in Fig. 7 that the missing boundaries were preserved and the position marked by the red rectangular frame was the boundary reserved by superposition. It can be seen from the superposition diagram in Fig. 7(d) that enough foreground information was retained in the image, including the seed umbilicus and edges. Although some redundant information in the Canny detection results is retained, it can be eliminated by morphological operation. On the basis of protecting the edge, the possibility of under-segmentation of the watershed algorithm is reduced.

2.5.4.4. Watershed generation. In the connected region of the image to be segmented, due to the improvement of the grey gradient by distance image reconstruction, the gradient of the distance image was uniform, which established a good foundation for watershed generation. Recalculating the grey gradient inside the seed image can effectively improve the influence of oversegmentation of the watershed algorithm. As the water rose from each valley floor and the level increased, when the water level of two adjacent valley floors rose higher than the peaks between them, a confluence line was formed, which was the watershed to achieve the division. The pixel that constitutes the watershed in the mask image was marked as −1, and the pixel at the corresponding position in the binary image was marked as 0 to achieve segmentation.

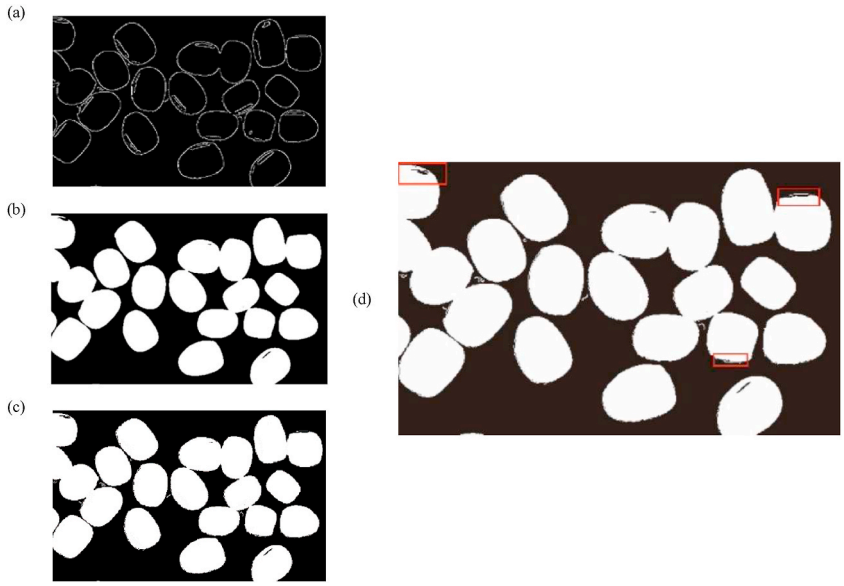


Fig. 7. (a) Canny edge detection result. (b) Otsu threshold segmentation results. (c, d) Superposition of the result.

3. Results and discussion

The proposed algorithm was used to segment 7 varieties of adzuki bean seeds and compare them with the segmentation results of other algorithms.

3.1. Result of two-particle adhesion segmentation

For the seed image containing only two adhesive seeds, such as Pinhong2014-166 and Yuhong4, 40 and 60 seeds were taken for the experiment respectively, which are numbered as 1 and 2. The residual rate of the counting results of different algorithms was compared, and the results are shown in Table 2.

The residual rate in Table 1 is listed as the sum of the ratio between the difference of the two counts and the true number of seeds, which is used to fit the deviation of the results of the calculation method, and the calculation formula is

$$z = \sum \left| \frac{x_m - x_i}{x_i} \right| \tag{7}$$

z is the percentage difference, x_i is the exact number of seeds in each experiment, and x_m is the counting result of two experiments in (7).

It can be seen from the experimental results in Table 1 that the residual rates of the watershed algorithm for the four images of the two varieties were small, both less than 0.25, which proved the existence of the phenomenon of oversegmentation and undersegmentation of the watershed algorithm. However, the residual rate of processing Yuhong4 by the edge detection algorithm was as high as 0.308, and the counting result was less than the correct value due to the undivided adhesion. These experimental data indicate that the watershed algorithm for processing seed adhesion was better than the edge detection algorithm. The residual rate of the concave point analysis algorithm was 0, which proved that it could handle the segmentation of the adhesion of two adzuki bean seeds. After determining the area of the connected region by the algorithm in this paper, the connected region of two-particle seeds was segmented by the concave point analysis algorithm. The residual rate reached 0, and the segmentation accuracy reached 100%. The segmentation result of the algorithm proposed for two-particle adhesions is shown in Fig. 8.

Fig. 8(a) and (c) were the parts including the more serious adhesion of the Pinhong2014_166 and Yuhong4 original images. It can be seen in Fig. 8(b) and (d) that the position of the adhesion of two seeds in Pinhong2014-166 and Yuhong4 were perfectly separated. The single seed was not affected by calculating the area of the connected region, and accurate segmentation was achieved.

3.2. Result of multiparticle adhesion segmentation

The concave point analysis algorithm is suitable for dividing the two-particle adhesion seeds and the contour of the seeds would be destroyed in the image of multiparticle adhesion in Fig. 9. Thus, the concave point analysis algorithm not applicable. For images including 80 seeds, such as Baihong11, Jihong14, Pinhong2014-166 and Jihong352, the data comparison results of adhesions processed by different algorithms are shown in Table 3.

Similar to the results of two-particle adhesion processing, the segmentation and the counting of the watershed algorithm were not stable, but the residual rate was stable within 0.14. Due to the oversegmentation and undersegmentation, the count was inaccurate. In comparison, the count of the edge detection algorithm was all lower than the correct value, and the residual rate was higher than 0.15. The residual rate of Jihong352 was as high as 0.350, which indicated that the edge detection algorithm has a poor segmentation ability for complex adhesions. The proposed algorithm could count all images correctly and the residual rate was 0. The comparison of the segmentation results among different algorithms and the proposed algorithm for multiparticle adhesions is shown in Fig. 10.

The image (a), (b), (c) and (d) in Fig. 10 respectively represented the original image, watershed algorithm result, edge detection result and the result of proposed algorithm of Baihong11. The image (e), (f), (g) and (h) in Fig. 10 respectively represented the original image, watershed algorithm result, edge detection result and the result of proposed algorithm of Jihong14. The image (i), (j), (k) and (l) in Fig. 10 respectively represented the original image, watershed algorithm result, edge detection result and the result of proposed algorithm of Pinhong2014-166. The image (m), (n), (o) and (p) in Fig. 10 respectively represented the original image, watershed algorithm result, edge detection result and the result of proposed algorithm of Jihong352.

The red line is the seed boundary generated in the images (b), (f), (j) and (n) in Fig. 10, and it can be seen that the traditional

Table 2
Data comparison of two-particle adhesion results processed by different algorithms.

	Image number	Watershed algorithm		Edge detection algorithm		Concave point analysis algorithm		Algorithm in this paper	
		count	Percentage difference	count	Percentage difference	count	Percentage difference	count	Percentage difference
Pinhong2014-166	1	46	0.233	35	0.275	40	0	40	0
	2	55		51		60		60	
Yuhong4	1	36	0.216	33	0.308	40	0	40	0
	2	67		52		60		60	

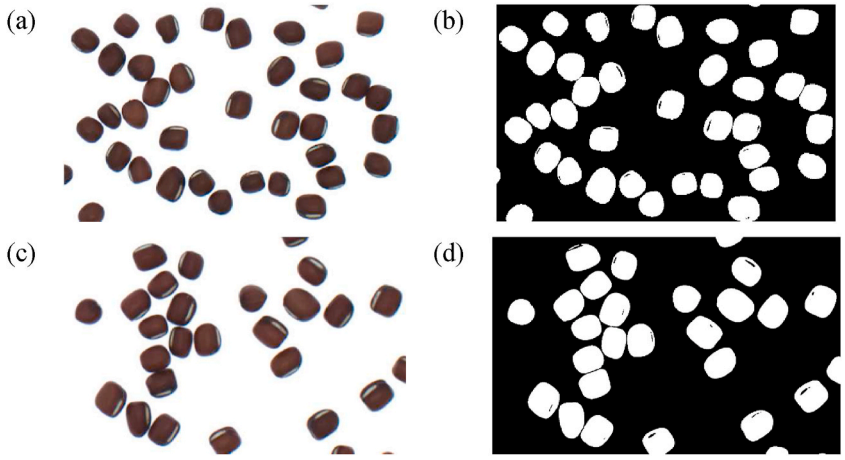


Fig. 8. (a) Pinhong2014-166 original image. (b) Pinhong2014-166 segmented result. (c) Yuhong4 original image. (b) Yuhong4 segmented result.

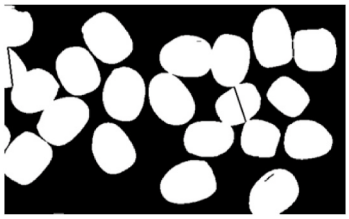


Fig. 9. The image was destroyed by the concave point analysis algorithm.

Table 3
Comparison of multiparticle adhesion data processed by different algorithms.

	Watershed algorithm		Edge detection algorithm		Algorithm in this paper	
	Count	Percentage difference	Count	Percentage difference	Count	Percentage difference
Baihong11-80	89	0.113	67	0.163	80	0
Jihong14-80	70	0.125	68	0.150	80	0
Pinhong2014-166-80	85	0.063	65	0.188	80	0
Jihong-352-80	91	0.138	58	0.350	80	0

watershed algorithm could cause the adzuki bean umbilicus to be misclassified as background or destroyed, resulting in over-segmentation and some seeds were not separated from each other. It can be seen from image (c), (g), (k) and (o) in Fig. 10 that the edge detection algorithm extracted the edge of the seed well. The umbilical was complete and the outline was clear, however, the complex adhesions were not be segmented. The segmentation result of the proposed algorithm is good. It can be seen from image (d), (h), (l) and (p) in Fig. 10 that good segmentation was achieved under the case of complex adhesions, the edge was relatively complete, and the seed umbilicus was preserved, which was convenient for further study.

3.3. Discussion

In the above experiments, due to the oversegmentation and undersegmentation, the watershed algorithm was unstable in its percentage difference, while the calculation value of the edge detection algorithm was lower than the correct value, resulting in a large counting error. Because segmentation based on concave points might destroy the whole seed, it was not applicable in the case of multiple-particle adhesions. In the case of only two-particle adhesions, the concave point analysis algorithm could be further processed after initial segmentation, which could improve the segmentation effect. The algorithm proposed combined the advantages of the above algorithms and eliminated the false boundary generated by the watershed algorithm. Under the premise of protecting the seed boundary, the precision of segmentation accuracy was improved as much as possible.

The background of the image collection was white, which was close to the color of the umbilical of the adzuki bean. During the binarization of the image, the umbilical part may be recognized as the background, resulting in certain foreground loss. In the process of segmentation, the seed umbilicus of adzuki seeds was retained, which may have a certain impact on the gradient of the distance map. Although it did not affect the segmentation count, it might affect the segmentation accuracy. In the processing results of

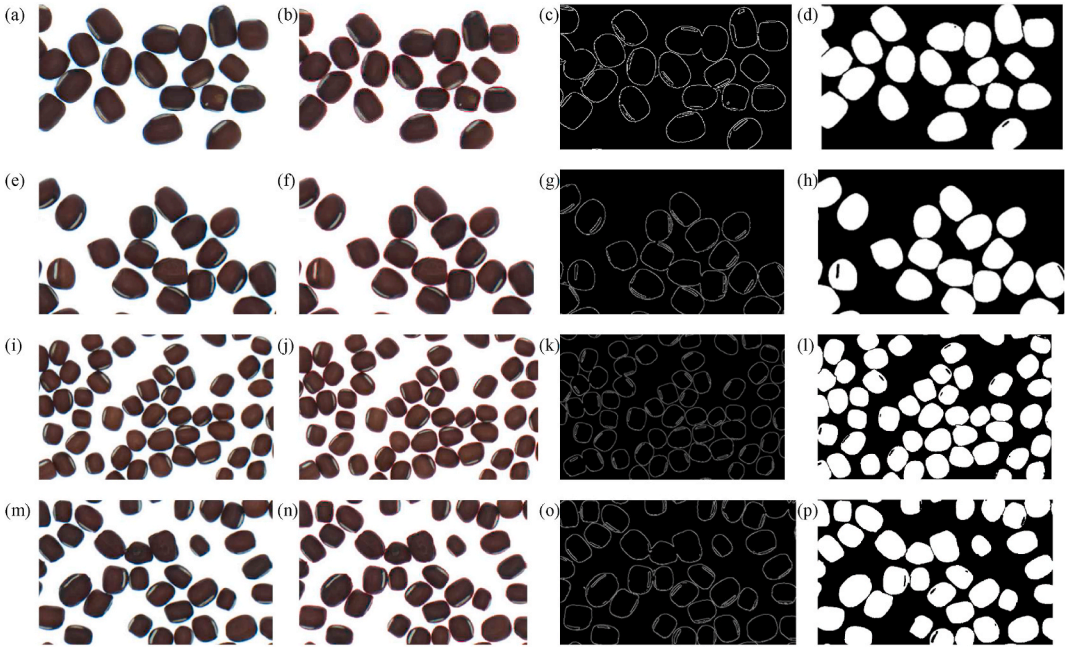


Fig. 10. The comparison of the segmentation results of Baihong11, Jihong14, Pinhong2014-166 and Jihong352 among different algorithms and the proposed algorithm for multiparticle adhesions. The images in four rows respectively represent the relate images of Baihong11, Jihong14, Pinhong2014-166 and Jihong352. The images in four columns respectively represent the original images, the processing result images of watershed algorithm edge detection algorithm and the proposed algorithm.

Jihong352, it can be seen in Fig. 11 that there is a certain error in the edge contour of the seed. It did not affect the counting accuracy; however, it may affect the calculation of the seed area and perimeter in further processing.

Previous studies on adzuki beans are few at present, mostly on seeds with uniform skin color, such as soybeans and peanuts. There may be no misclassification of seed umbilicus, and segmentation can also achieve high accuracy. The algorithm proposed in this paper was suitable for crop image segmentation with umbilicus, and can be further extended and applied to the segmentation of other legume seeds.

4. Conclusion

In this paper, a fusion algorithm for the accurate segmentation and counting of bean seed images with adhesions was proposed. Based on the traditional watershed algorithm, morphological reconstruction was used to reduce the grey difference between the seed coat and the umbilicus of adzuki bean and the possibility of misclassification. Canny edge detection was introduced to strengthen the right edge and reduce the error edge, which may cause oversegmentation. For the adhesion between two seeds, the concave points were identified by the concave point analysis algorithm and matched according to the shortest Euclidean distance to achieve accurate segmentation. For more complex adhesion, the distance map reconstruction was used to adjust some areas of the distance map to adapt to the segmentation of the watershed algorithm, and reduce the influence of oversegmentation and undersegmentation. Finally, the watershed was generated, which could better segment the adhesion area in general and accurate counting was achieved. Experimental results shown that the proposed algorithm significantly improved the accuracy of seed segmentation and achieved the automatic counting of adzuki beans. It overcame the shortcomings of manual placement of adzuki beans in previous studies on seed counting, as well as reduced the serious miscounting caused by the inaccurate segmentation of adhesion images, and provided a reference for the

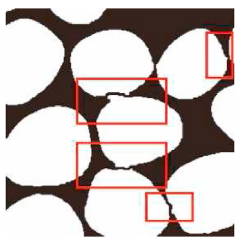


Fig. 11. The lack of precision in the segmentation results of Jihong352.

automatic detection of adzuki beans. At the same time, the proposed algorithm could be extended to the adhesion segmentation of other bean images.

In the process of seed image processing, the lack of precision of the edge existed in some segmentation results. Although the counting accuracy was not affected, it would introduce errors in the subsequent extraction of seed phenotype information and reduce the calculation accuracy. Therefore, further research is needed. Improving the accuracy of edge division on the premise of preserving bean umbilicus is the key issue to be considered in the next step of algorithm optimization and improvement.

Acknowledgements

This paper was supported by the China Agriculture Research System of MOF and MARA, China-Food Legumes (CARS-08-G-22) and the Science and Technology Project of Hebei Education Department (QN2020421).

References

[1] M.A. Momin, K. Yamamoto, M. Miyamoto, N. Kondo, T. Grift, Machine vision based soybean quality evaluation, Comput. Electron. Agric. 140 (2017), https://doi.org/10.1016/j.compag.2017.06.023.
[2] J. Wang, J. He, Y. Han, C. Ouyang, D. Li, An Adaptive Thresholding algorithm of field leaf image, Comput. Electron. Agric. 96 (2013) 23–39, https://doi.org/10.1016/j.compag.2013.04.014.
[3] Q. Wang, M. Cheng, X. Xiao, H. Yuan, J. Zhu, C. Fan, J. Zhang, An image segmentation method based on deep learning for damage assessment of the invasive weed Solanum rostratum Dunal, Comput. Electron. Agric. 188 (2021), 106320, https://doi.org/10.1016/j.compag.2021.106320.
[4] J. Ma, K. Du, L. Zhang, F. Zheng, J. Chu, Z. Sun, A segmentation method for greenhouse vegetable foliar disease spots images using color information and region growing, Comput. Electron. Agric. 142 (2017) 110–117, https://doi.org/10.1016/j.compag.2017.08.023.
[5] D. Patrícioa, R. Riederb, Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review, Comput. Electron. Agric. 153 (2018) 69–81, https://doi.org/10.1016/j.compag.2018.08.001.
[6] X. Yang, X. Shi. Computer image processing technology and its application in agricultural engineering. Digital Communication world, 2021(10): 45-46+50. doi: 10.3969/J.ISSN.1672-7274.2021.10.020.
[7] L. Zhou, W. Wu, T. Liu, C. Sun, Research status and prospect of rice and wheat grain counting methods, Xiandai Nonge Keji (12) (2020) 18–20+22.
[8] X. Zheng, F. Yang, Z. Li, Overview of crop image segmentation algorithm, Mod. Comput. (19) (2020) 72–75, https://doi.org/10.3969/j.issn.1007-1423.2020.19.015.
[9] P. Huang, Q. Zheng, C. Liang, Summary of image segmentation methods, J. Wuhan Univ. (Nat. Sci. Ed.) 66 (6) (2020) 519–531, https://doi.org/10.14188/j.1671-8836.2019.0002.
[10] S. Singh, N. Mittal, D. Thakur, H. Singh, D. Oliva, A. Demin, Nature and biologically inspired image segmentation techniques, Arch. Comput. Methods Eng. (2021) 1–28, https://doi.org/10.1007/S11831-021-09619-1 (prepublish).
[11] H. Gao, T. Zhen, Z. Li, A review of segmentation methods for adhesive particle images, J. Chin. Cereals Oils Assoc. 37 (3) (2022) 186–194.
[12] S. Divyameena, M. Mangaleswaran, A study on various image segmentation algorithms, Int. J. Sci. Res. Sci., Eng. Technol. (2018) 272–276, https://doi.org/10.32628/ijsrset21841134.
[13] Y. Xue, J. Zhao, M. Zhang, A watershed-segmentation-based improved algorithm for extracting cultivated land boundaries, Rem. Sens. 13 (5) (2021) 939, https://doi.org/10.3390/RS13050939.
[14] S. Qiao, Q. Yu, Z. Zhao, L. Song, H. Tao, T. Zhang, C. Zhao, Edge extraction method for medical images based on improved local binary pattern combined with edge-aware filtering, Biomed. Signal Process Control 74 (2022), 103490, https://doi.org/10.1016/J.BSPC.2022.103490.
[15] Y. Yao, Y. Li, W. Zou, Y. Liu, R. He, Determination method on thousand-seed weight of rapeseed based on image processing, Oil Crop Science 44 (1) (2022) 201–210, https://doi.org/10.19802/j.issn.1007-9084.2020368.
[16] Z. Shuo, Real-time localization approach for maize cores at seedling stage based on machine vision, Agronomy 10 (2020), https://doi.org/10.3390/agronomy10040470.
[17] S. Tan, X. Ma, Z. Mai, L. Qi, Y. Wang, Segmentation and counting algorithm for touching hybrid rice grains, Comput. Electron. Agric. 162 (C) (2019) 493–504, https://doi.org/10.1016/j.compag.2019.04.030.
[18] W. Wu, L. Zhou, J. Chen, Z. Qiu, Y. He, T.K.W. Gain, A measurement System of thousand kernel weight based on the android platform, Agronomy 8 (9) (2018), https://doi.org/10.3390/agronomy8090178.
[19] M. Gamarra, E. Zurek, H.J. Escalante, L. Hurtado, H. San-Juan-Vergara, Split and merge watershed: a two-step method for cell segmentation in fluorescence microscopy images, Biomed. Signal Process Control 53 (2019), 101575, https://doi.org/10.1016/j.bspc.2019.101575.
[20] Q. Sun, J. Zheng, C. Li, Improved watershed analysis for segmenting contacting particles of coarse granular soils in volumetric images, Powder Technol. 356 (C) (2019) 295–303, https://doi.org/10.1016/j.powtec.2019.08.028.
[21] L. Libor, K. Martin, Automatic buildings detection using Sobel, roberts, Canny and prewwitt detector, J. Electr. Eng. 72 (4) (2021) 278–282, https://doi.org/10.2478/JEE-2021-0039.
[22] K. Wang, P. Liu, L. Wang, Infrared image adaptive Canny edge-detection of aircraft skin based on fast fireworks algorithm, Infrared Technol. 43 (5) (2021) 443–454.
[23] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 8 (6) (1986), https://doi.org/10.1016/B978-0-08-051581-6.50024-6.
[24] B. Zhang, H. Lu, F. Wang, M. Li, X. Wang, Study on segmentation method of adhesive chicken body based on concave point analysis, J. Chin. Agric. Mech. 42 (2) (2021) 164–170+183, https://doi.org/10.13733/j.jcam.issn.2095-5553.2021.02.025.
[25] T. Wu, Sophisticated Phenotyping and Equipment Development of Individual Wheat Seed, Yang Ling: Northwest A&F University, 2020, https://doi.org/10.27409/d.cnki.gxbnu.2020.000018.