

Article

Overlapping Pellet Size Detection Method Based on Marker Watershed and GMM Image Segmentation

Weining Ma ^{1,2,3} , Lijing Wang ^{1,2,3}, Tianyu Jiang ^{1,2,3}, Aimin Yang ^{1,2,3} and Yuzhu Zhang ^{3,*}

¹ Hebei Engineering Research Center for the Intelligentization of Iron Ore Optimization and Ironmaking Raw Materials Preparation Processes, North China University of Science and Technology, Tangshan 063210, China

² Tangshan Intelligent Industry and Image Processing Technology Innovation Center, North China University of Science and Technology, Tangshan 063210, China

³ College of Metallurgy and Energy, North China University of Science and Technology, Tangshan 063210, China

* Correspondence: zyz@ncst.edu.cn

Abstract: The particle size of pellets is an important parameter in steel big data, and the high density and high overlap rate of pellets bring a great challenge to particle size detection. To address this problem, a particle size intelligent detection algorithm with an improved watershed and a Gaussian mixture model (GMM) is proposed. First, the initial segmentation of the pellets and background is achieved by using adaptive binary segmentation, and then the secondary fine segmentation of the pellets and background is achieved by combining morphological operations such as skeleton extraction and marked watershed segmentation; then, the contour of the connected domain of pellets is calculated, and the non-overlapping pellets in the foreground and the overlapping pellets are filtered according to the roundness of their contours. Finally, the number of overlapping pellets is predicted by Gaussian reconstruction of the grayscale image of the overlapping pellets, and the number and granularity of the overlapping pellets are predicted by the Gaussian reconstruction of the overlapping pellets. The experimental results showed that the algorithm achieved a 91.98% segmentation accuracy in the experimental images. Compared with other algorithms, the algorithm can also effectively suppress the over-segmentation and under-segmentation problems, and it can effectively realize the pellet size detection of dense, overlapping pellets such as those on a pelletizing disk, which provides an effective technical means for the metallurgical performance analysis of pellet ore and intelligent pellet-making driven by big data.

Keywords: steel big data; particle size; intelligent pelletizing; watershed; GMM



Citation: Ma, W.; Wang, L.; Jiang, T.; Yang, A.; Zhang, Y. Overlapping Pellet Size Detection Method Based on Marker Watershed and GMM Image Segmentation. *Metals* **2023**, *13*, 327. <https://doi.org/10.3390/met13020327>

Academic Editor: João Manuel R. S. Tavares

Received: 28 December 2022

Revised: 2 February 2023

Accepted: 2 February 2023

Published: 6 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Pelletized ore is a kind of spherical artificial iron ore which has the advantages of low pollution levels, high-grade quality, uniform particle sizes, superior cold strength, high void ratios, and good reduction compared with other artificial ores, and it is one of the primary raw materials for modern blast furnace iron-making. In recent years, the pellet ore feed ratio has been rising year by year, and so it is important to vigorously develop high-quality pellet ore production process technologies to accelerate the industrial upgrading of the steel industry and realize an intelligent manufacturing system [1]. The particle size and particle size distribution in the pellet-making process is an important indicator for measuring the quality of raw balls, as the raw balls that are too small cannot meet the mechanical strength requirements and are easy to break during transportation, and broken raw balls in the roasting process will produce a lot of dust and shorten the service life of rotary kiln equipment. In a blast furnace, a pellet with a uniform particle size has a large porosity, small airflow resistance, and fast reduction rate, which provides favorable conditions for high yield and low consumption in the blast furnace. Manual sampling and sieving is a traditional method for conducting pellet size detection, but this method is susceptible

to a number of samples and the randomness of sampling, and thus, it is not suitable for detecting large quantities of pellets. The process of sampling relies heavily on manual labor and the site environment is harsh, which can easily cause safety hazards. Manual sieving and detection is time-consuming and the pellets cannot be sampled frequently enough to achieve the effect of real-time detection; in addition, physical sieving may cause the pellets to experience wear and rupture, thus affecting the accuracy of detection.

The development of machine visioning has laid a strong theoretical foundation for the promotion and application of image detection technology. The use of image detection technology to identify and detect the particle size of pellets can effectively overcome the shortcomings of the sieving method that relies on manual labor, and image detection can increase site safety and improve real-time detection. At present, some large- and medium-sized pellet plants are attempting to use machine vision technology as the basis of intelligent detection means through high-definition industrial cameras, which are used to collect images of the pellet-making site. The pellet image acquisition process as shown in Figure 1. The image segmentation algorithm is used to separate the pellets from the background, and the size of each pellet in the image and the overall size distribution are counted. However, since the pellets are not separated from each other in actual production, a large number of pellets are distributed in the form of adhesion and overlap, which makes the image segmentation of the pellets very difficult [2–9].

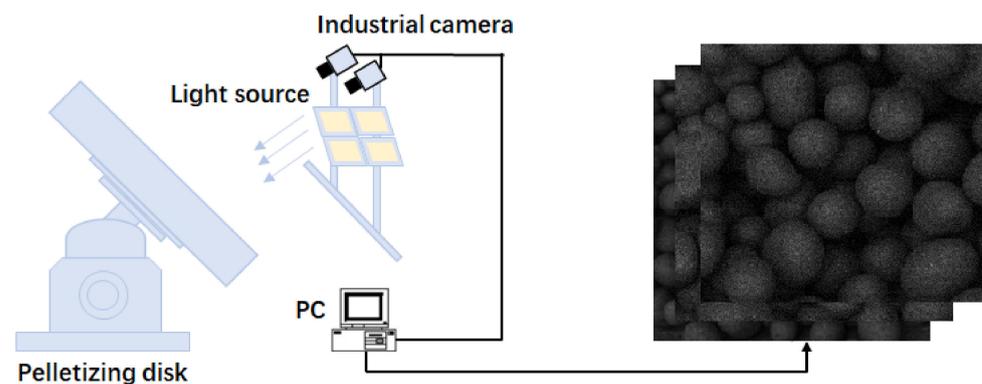


Figure 1. Schematic diagram of pellet image acquisition for pellet-making discs.

In order to solve the above problems, many scholars have carried out in-depth analyses on the characteristics of the images of agglomerated and overlapping objects, and they have proposed a series of feasible solutions. Heydari et al. proposed three methods to predict the number of overlapping pellets: Hough transform circle detection, marked watershed, and support vector machine, and they created a majority voting decision mechanism for predicting the results obtained by the three methods. This method is only applicable in the case of sparsely distributed pellets, and the algorithm ignores the adhesion of more than three pellets. Wu et al. used a disk scanning method to select the effective edge points of the overlapping contours, and then they calculated the particle size of the overlapping pellets. This method can use effective edge information, to some extent, to overcome the shortcomings of the inaccurate edge segmentation of the marked watershed method. However, the accuracy of the morphological reconstruction of the algorithm depends on the selection of the size of the super-parameter structural elements. The disk scanning method is not suitable for estimating the particle sizes of pellets with less roundness, and it needs further improvement for irregular pellets and densely distributed pellets. Zafari et al. proposed an overlapping object segmentation algorithm based on contour concave point extraction, which realized the segmentation of overlapping objects. However, this method depends on the accuracy of the contour extraction. When the background area is not clear from the target area or the edge of the target area is not clear, this algorithm cannot extract the contour of the target area well [10–12].

In order to solve the problem of overlapping pellet segmentation and particle size estimation, a machine vision image particle size detection algorithm is proposed, which indirectly estimates pellet size information through pellet image detection. The main contributions of this study are as follows:

1. The adaptive binary segmentation method is used to segment the foreground area and the background area, which overcomes the traditional global threshold binary segmentation that is susceptible to light brightness and light unevenness factors, and it ensures the anti-interference under different lighting environments.
2. Further subdivision of the foreground and background areas using morphological operations and marker watershed segmentation can effectively compensate for the disadvantages created by the imprecise foreground and background edges obtained using binary segmentation, and it can improve the accuracy of segmentation.
3. A model of binary image segmentation based on a Gaussian mixture model is established where the binary image of the overlapping pellets is used as a template and its corresponding grayscale image is extracted, a Gaussian reconstruction is performed on the extracted grayscale image, and the number of clusters and the initial value of the center are calculated by the reconstructed grayscale map. Then, the coordinates of each pixel of the binary image are used as samples, and the initial value and the samples are brought into the hybrid Gaussian model, which can realize the pixel-level segmentation of the binary image of the pellets.
4. A method is proposed for the size estimation of the edges of incomplete pellets, and the least squares method is used to fit the circle to estimate the number of pellets in the edge area, which can effectively overcome the disadvantage of the small sizes of the pellets estimated by the circumscribed circles.

The remainder of this paper is organized as follows: Section 2 presents the principle of the algorithm, Section 3 analyzes the experimental results, and Section 4 summarizes the conclusions.

2. Algorithm Principle

The flowchart of the proposed algorithm for pellet size detection is shown in Figure 2. Firstly, the input pellet image is pre-processed by median filtering to filter out the noise interference in the image; then, the foreground target is initially extracted by adaptive binary segmentation and the foreground and background areas are further segmented by watershed segmentation. The contours of the connected domain in the foreground are judged according to the roundness formula, and the non-overlapping pellets are separated from the overlapping pellets for processing. The pellet fitting method is selected according to the edge evaluation mechanism, and, finally, the selected circle fitting method is used to estimate the particle size of the pellet and obtain the particle size information of the pellet.

2.1. Foreground Segmentation

A sample with a large number of overlapping pellets is characterized by indistinguishable background and foreground areas where the grayscale value of the image of the surface pellet area is high and the grayscale value of the obscured pellet area and the background area is relatively low, and so the pellets with high grayscale values and better visualization can be treated as foreground while both the pellets with low grayscale values and backgrounds are treated as background. According to the binary segmentation algorithm, the foreground area and the background area are segmented, and the obtained foreground area includes separate pellets and overlapping pellets, which lays the foundation for further segmentation processing.

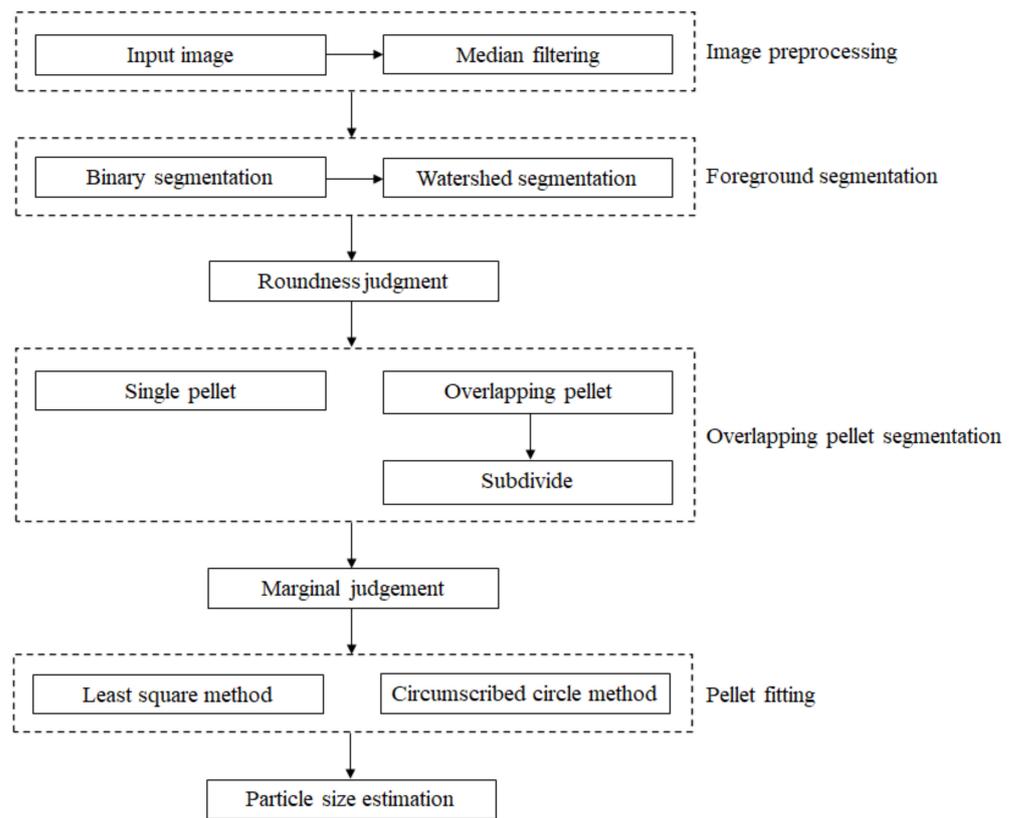


Figure 2. Flow chart of the pellet size detection algorithm.

2.1.1. Median Filtering

The pellets form unevenly during the process of pellet-forming, and their surfaces are unevenly distributed, and so the brightness of the pellets is not uniformly distributed in the imaging process. Thus, performing binary segmentation directly on the image will generate a lot of local extreme value noise, which brings great difficulties to the process of identifying of the pellets. Compared with other filtering algorithms, median filtering can effectively overcome the edge blurring phenomenon while suppressing the salt-and-pepper noise. Median filtering can be expressed as follows:

$$G(x, y) = \text{med}\{f(x - k, y - l), (k, l \in W)\} \quad (1)$$

where W represents the window size of the median filter, $f(x, y)$ represents the original image, and $G(x, y)$ represents the image after median filtering.

In the median filtered image process, the pixel value at each point is the median of all the pixel values in the window centered at that point, and the median filtered image is obtained by sliding the window through each pixel. The filtering ends when, after several rounds of filtering, the filtered image no longer changes. Compared with Gaussian filtering, the median filtered image process removes noise and retains clear edge information, preventing the loss of some detail information and small foreground areas.

2.1.2. Binary Segmentation

Before beginning the pellet image segmentation, firstly, the foreground (pellet area) should be separated from the background, and the binary segmentation is selected to separate the foreground from the background. Considering that the uneven light in the pellet production plant and the mutual occlusion between pellets can lead to an uneven grayscale distribution in the image, the binary segmentation algorithm must be highly resistant to brightness variations. The adaptive thresholding algorithm [13] is a local thresholding algorithm that is robust to images with uneven brightness.

The threshold value of adaptive binary segmentation is not a fixed value—it depends on the neighborhood s_{xy} centered on the pixel point (x, y) in the image, where m_{xy} and σ_{xy} are the mean and standard deviation of the neighborhood, T_{xy} is the threshold value, $f(x, y)$ is the original image entered, and $g(x, y)$ is the calculated binary image, and the formula is calculated for all pixel positions in the image and a different threshold is calculated at each position using the pixels in the neighborhood s_{xy} .

Adaptive thresholding segmentation is defined as follows:

$$g(x, y) = \begin{cases} 255, & f(x, y) > T_{xy} \\ 0, & f(x, y) \leq T_{xy} \end{cases} \quad (2)$$

$$T_{xy} = \sigma_{xy} - C \quad (3)$$

where C is a constant that is used to adjust the threshold T_{xy} .

Figure 3a is a pellet image with uneven grayscale, and the red dashed box area has a lower grayscale value. Figure 3b is the result obtained using the Nobuyuki Otsu method (Ostu) where, when the image brightness is not uniform, the red boxed area is treated with less grayness than the background, which causes mis-segmentation. Figure 3c shows the results obtained using the adaptive threshold segmentation method, which can effectively overcome the problem of foreground loss caused by uneven gray scale and realize the segmentation of the foreground and background.

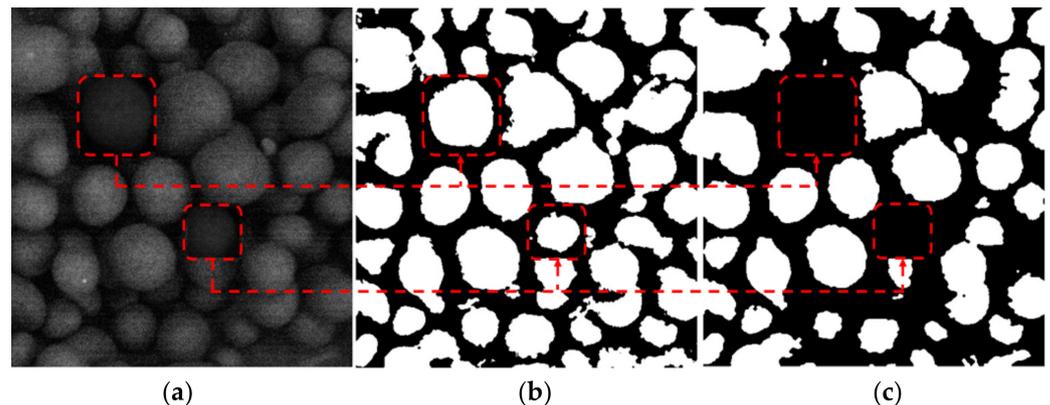


Figure 3. Binary segmentation. (a) Inhomogeneous grayscale images. (b) Adaptive threshold segmentation results. (c) Ostu segmentation results.

2.2. Foreground Refinement

The adaptive binary segmentation will initially divide the foreground and background, but because the edge areas of the foreground and background are blurred, the blurred edges will lead to errors in the binary segmentation. As the large or small segmentation of the foreground area directly affects the subsequent estimation of the pellet size, the secondary segmentation of foreground and background can further refine the edge information.

The watershed segmentation method can effectively segment the edges of the adhering objects, viewing the foreground and background as two objects, and the watershed segmentation method can further segment the edge area. First, the binary image, A , is obtained by adaptive segmentation and subjected to the morphological open operation, and B is the structure element selected by the open operation. The morphological open operation can be expressed as:

$$A \circ B = (A \ominus B) \oplus B \quad (4)$$

The morphological opening operation can be used to remove noise and fine target interference in the foreground and obtain a defined foreground area; then, the background area of the adaptive segmentation binary image is extracted and the background area is

skeletonized to obtain the determined background area. Finally, the definite and indefinite areas are well marked and processed by the watershed segmentation algorithm to further obtain the edge demarcation line between the foreground and background areas. Morphological skeleton extraction [14] can be expressed as:

$$S(A) = \bigcup_{k=0}^K S_k(A) \quad (5)$$

$$S_k = (A \ominus kB) - (A \ominus kB) \circ B \quad (6)$$

where A is the binary image of the background area, B is a structural element, and $(A \ominus kB)$ denotes the k consecutive erosions of A .

Figure 4a shows the determined foreground area obtained after the morphological opening operation, Figure 4b shows the determined background area obtained after the skeleton extraction of the background area, and Figure 4c shows the result obtained by determining the difference between the foreground area obtained after the watershed segmentation and Figure 4a, wherein the white area is the foreground area missed during the binary segmentation. The final foreground area obtained is shown in Figure 4d. Re-segmentation refines the boundary between the background and foreground areas and improves the accuracy of the foreground segmentation.

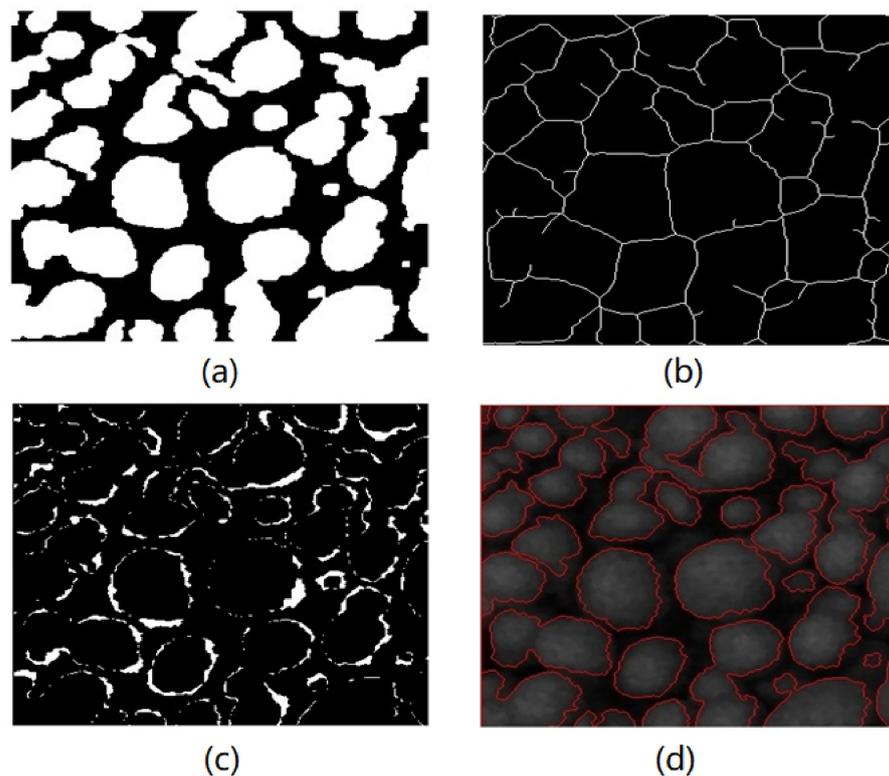


Figure 4. Watershed segmentation of the foreground area. (a) Defined foreground area. (b) The skeleton of the background area. (c) Missing foreground areas. (d) Final segmentation result.

2.3. Overlapping Pellet Segmentation

Although the pellets in the foreground area are segmented, most of them are overlapping pellets, with a few non-overlapping pellets, due to the dense distribution of the pellets. Distinguishing the foreground area between the non-overlapping pellets and the overlapping pellets and segmenting the overlapping pellets are the key issues in pellet size recognition.

2.3.1. Overlapping Pellets Discrimination

In the binary image, the foreground areas obtained are the different connected domains. The connected domains of the non-overlapping pellets and the overlapping pellets have different characteristics, and the contours of the connected domains of the non-overlapping pellets are more rounded. In contrast, the roundness of the non-overlapping pellets is often smaller because multiple pellets have adhered together. It is possible to discriminate between the overlapping pellets and non-overlapping pellets based on this characteristic of roundness. The roundness of the foreground area can be judged by its contours, and the process of calculating the roundness [15] can be expressed as follows:

$$Distance = \frac{1}{F} \sum (\|p - p_i\|) \quad (7)$$

$$Sigma^2 = \frac{1}{F} \sum (\|p - p_i\| - Distance)^2, \text{ and} \quad (8)$$

$$Roundness = 1 - \frac{Sigma}{Distance} \quad (9)$$

where p is the center of the area, p_i represents the coordinates of all pixel points on the contour, F is the area enclosed by the contour, $Distance$ is the average distance from the pixel points on the contour to the center, $Sigma^2$ is the variance in the distance from the contour pixel points to the center and the average distance, and $Roundness$ represents the relationship between the mean and the standard deviation, where the value is between 0 and 1 and the larger indicates that the contour is closer to the standard circle.

When different roundness thresholds are selected, the proportions of screened non-overlapping pellets and overlapping pellets are different, and the larger the selected threshold, the fewer non-overlapping pellets are obtained, while the smaller the selected threshold, the more non-overlapping pellets are obtained. Figure 5 shows the screening results corresponding to the roundness thresholds of 0.4, 0.5, 0.6, and 0.7, respectively, where the green contours represent pellets with roundnesses that are greater than the threshold and the red contours represent pellets with thresholds that are less than the contour of the pellet.

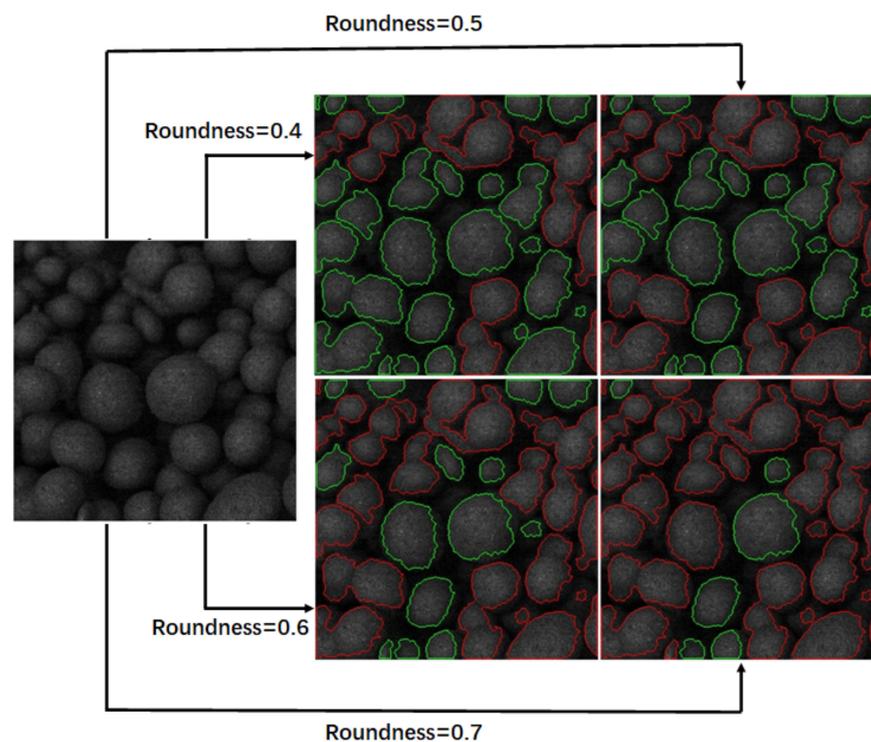


Figure 5. Screening results of different roundness thresholds.

2.3.2. Re-Segmentation of Overlapping Pellets

After roundness screening, the pellets with larger roundnesses are treated as non-overlapping pellets which can be directly fitted with a circle to estimate the particle size. The connected domain in the remaining foreground area represents the overlapping pellets, which are composed of multiple pellets overlapped, together with local occlusion between the pellets, and so the granularity information of each pellet cannot be judged directly and further segmentation is required. The pixel coordinates of the binary images of the overlapping pellets are selected as the sample variables, and a Gaussian mixture model can be used to achieve a cluster segmentation of the overlapping pellets, and then the overlapping pellets can be separated.

When the Gaussian mixture model performs the clustering operation, it is necessary to pre-specify the clustering value K , i.e., how many clusters are contained in the corresponding overlapping pellets. It is more difficult to estimate the number of pellets directly from the binary image because the overall information of the overlapping pellets in the binary image does not reflect the local feature information of the pellets; therefore, it is better to combine the binary image of the pellets with the grayscale image and use the binary image as a template to select the area in the corresponding grayscale image to estimate the number of pellets. The calculation process can be expressed as follows:

$$m(x, y) = g(x, y) * b(x, y) \quad (10)$$

where $g(x, y)$ represents the grayscale image, $b(x, y)$ represents the binary image, and $m(x, y)$ represents $b(x, y)$, which is the grayscale image with the template.

The grayscale image of the pellets is characterized by a high grayscale value in the central area and a gradual decrease in grayscale from the center outward, but the grayscale of the pellets is not completely Gaussian-distributed due to the surface inhomogeneity of the pellets and the influence of noise. The binary image to be segmented is taken as a template and its corresponding grayscale image area is selected, and Gaussian filtering is applied to the area. By reconstructing the grayscale image, local noise can be filtered out and the image becomes smooth, and the overall grayscale value tends toward Gaussian distribution. A suitable threshold is selected to perform binary segmentation on the reconstructed grayscale map, and the number of connected domains of the binary image is obtained as an estimate of the number of pellets. In the paper, the average grayscale value of the overlapping pellet area after filtering is selected as the threshold value, and the K connected domains generated after binarization are the number of pellets in the overlapping area.

The circular area in Figure 6a is selected as the study object. Figure 6b is the grayscale 3D distribution of the circular area, which contains more local extremes, and the binarization yields the 3D distribution of Figure 6d, which generates multiple connected domains, resulting in the inaccurate estimation of the number of spherical clusters. Figure 6c is the grayscale 3D distribution of the circular area after Gaussian filtering, and the binarization yields the 3D distribution of Figure 6e, which contains only one connected domain, and it can estimate the number of spherical clusters more accurately.

In order to prevent the local extremes that continue to exist in the filtered grayscale map, it is also necessary to filter the connected domains, the smaller areas in the connected domains, and the final number of connected domains K , and then the corresponding K clustering centers can be obtained. The estimation of K initial clustering centers can effectively accelerate the convergence of the Gaussian mixture model and improve the accuracy of the clustering partitioning.

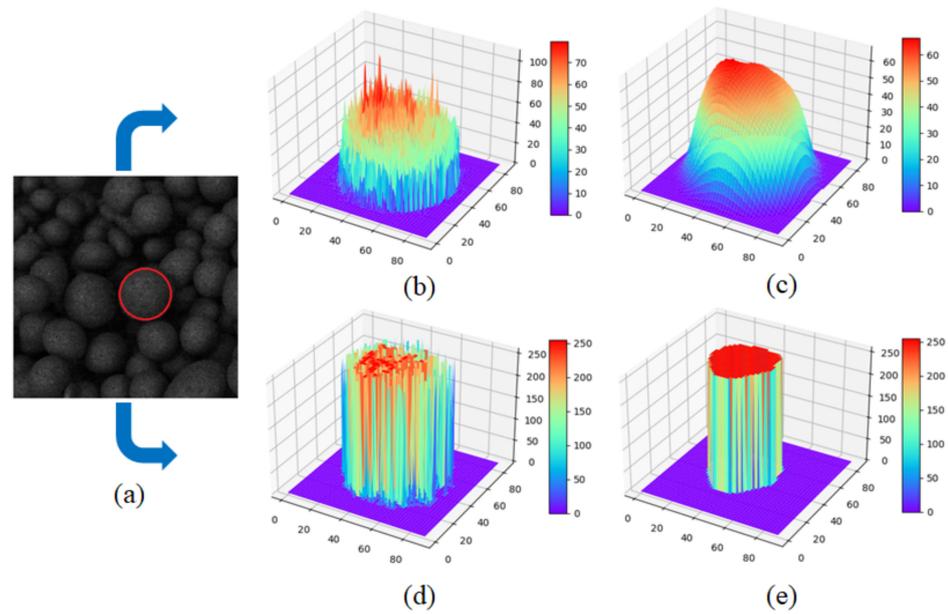


Figure 6. Gaussian reconstruction process. (a) The original grayscale map. (b) Grayscale Gaussian distribution of the original map. (c) Reconstructed map of grayscale Gaussian distribution. (d) Original image binarization distribution. (e) Reconstructed image of binarization distribution.

Figure 7 shows the segmentation results of the overlapping pellets. The three selected connected domains with smaller roundnesses, corresponding to the areas in the red dashed box, are used to estimate the K values and initial clustering centers according to the above method, and then the Gaussian mixture model (GMM) clustering algorithm is used to realize the segmentation of the overlapping pellets. The areas with different gray values in the segmentation results in the figure represent each pellet after the clustering segmentation.

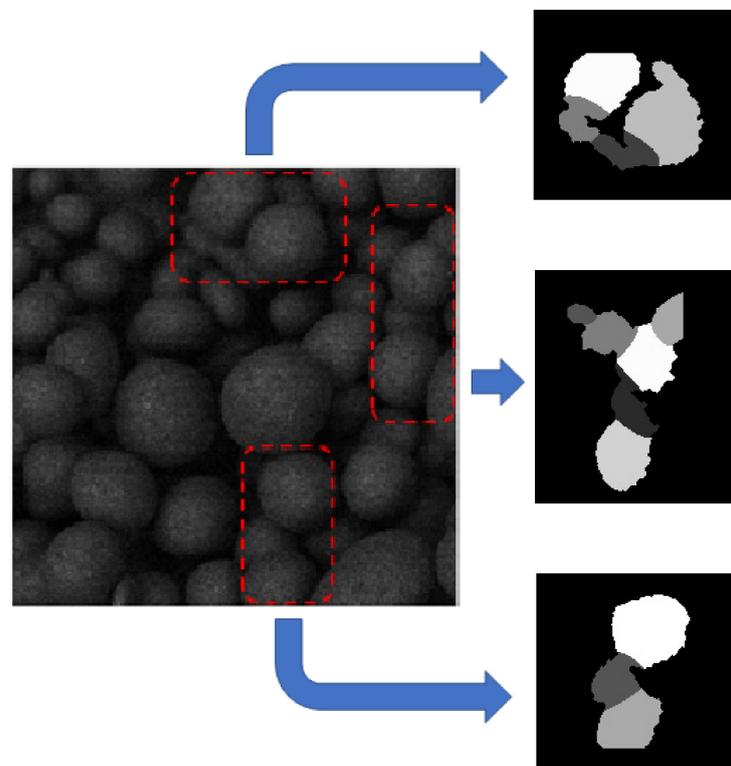


Figure 7. Overlapping pellet segmentation.

2.4. Pellet Size Estimation

The particle sizes of the non-overlapping pellets can be estimated by solving the least square circle [16] or the peripheral circle [17] of the contour. However, the overlapping pellets obtained after being divided are often incomplete pellets of poor roundness. If the least square circle is used to fit the contour of the divided pellets, the particle size obtained will be smaller. Therefore, the external circle method should be used to fit the contour of the re-segmented pellets.

For the contours with better roundnesses, the contour is the edge point of the pellet, and the particle size obtained by fitting the contour of the pellet with the least squares method is more accurate. If we let the coordinates of a single spherical contour be represented by $(x_i, y_i), i = 1, 2, \dots, n$, solving for the least squares circle of this contour, then the equation of the circular curve as follows:

$$x^2 + y^2 + ax + by + c = 0 \quad (11)$$

where x and y are the coordinates on the circle obtained from the contour point fit; a, b and c are the parameters of the equation of the curve of the circle; the equation of this curve is centered on $(-\frac{a}{2}, -\frac{b}{2})$; and the circle's radius is $\frac{\sqrt{a^2+b^2-c}}{2}$. The curve parameters can be found according to the least squares method for a, b and c . The matrix can be written as follows:

$$A = \begin{bmatrix} \sum_i x_i & \sum_i y_i & n \\ \sum_i x_i y_i & \sum_i y_i^2 & \sum_i y_i \\ \sum_i x_i^2 & \sum_i x_i y_i & \sum_i x_i \end{bmatrix}, B = \begin{bmatrix} -\sum_i (x_i^2 + y_i^2) \\ -\sum_i (x_i^2 y_i + y_i^3) \\ -\sum_i (x_i^3 + x_i y_i^2) \end{bmatrix} \quad (12)$$

where x_i and y_i represent the i -th pixel coordinate points on the contour of the sphere and n is the total number of contour pixels. Then, the parameters of the circular curve equation can be found as follows:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = A^{-1}B \quad (13)$$

By bringing a, b and c to the center, the radius expression is obtained as the particle size information of the pellet contour.

For poorly rounded contours, the contour points are not complete pellet edge points. Since the overlapping pellets obscure each other, the edges of the pellets obtained by re-segmentation contain the information of the obscured edges, and the granularity information obtained is inaccurate if the least squares method is used to fit the contours directly. When the maximum distance between the pellet contour points is equal to the pellet diameter, the derived circumscribed circle is the same as the real pellet, and when the obscured area is large and the maximum distance between the pellet contours is chosen as the pellet diameter, the derived circumscribed circle is the closest to the real pellet. Therefore, the outer circle of the contour is used to estimate the particle size of the pellet, which can, to the greatest extent possible, approximate the real particle size of the pellet.

If we let the coordinates of a single pellet contour be represented by $(x_i, y_i), i = 1, 2, \dots, n$, then the distance between any two contour points is:

$$d_i = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, i, j = 1, 2, \dots, n \quad (14)$$

We can then calculate the maximum distance between any two contour points as follows:

$$d_{\max} = \max(d_i) \quad (15)$$

Using the center of the two points at the maximum distance as the center of the circle and the maximum distance as the diameter, we can find the circumscribed circle of the “two-point formula” as follows:

$$A = \frac{x_i + x_j}{2}, B = \frac{y_i + y_j}{2}, R = \frac{d_{\max}}{2} \quad (16)$$

where (A, B) and R represent the center and radius of the circumscribed circle, respectively.

2.5. Evaluation Mechanism of Edge Pellets

Some incompleteness in the edge area is caused by the interception of the image boundary, which is different from the incompleteness of the globules caused by occlusion, and so it is inaccurate to judge such globules by using only roundness. A new and different evaluation mechanism is used for the spherical clusters that are at the edges. The red dashed line in Figure 8a shows the contour of the sphere at the edge, and the new evaluation mechanism stipulates that the accuracy of the fit is higher when the effective area of the fitted circle intersects with the contour area of the sphere in a larger ratio. This can be expressed as follows:

$$IOU = \frac{(A \cap C) \cap B}{(A \cap C) \cup B} \quad (17)$$

where A denotes the area of the fitted circle, B denotes the area of the pellet contour, and C denotes the area of the image.

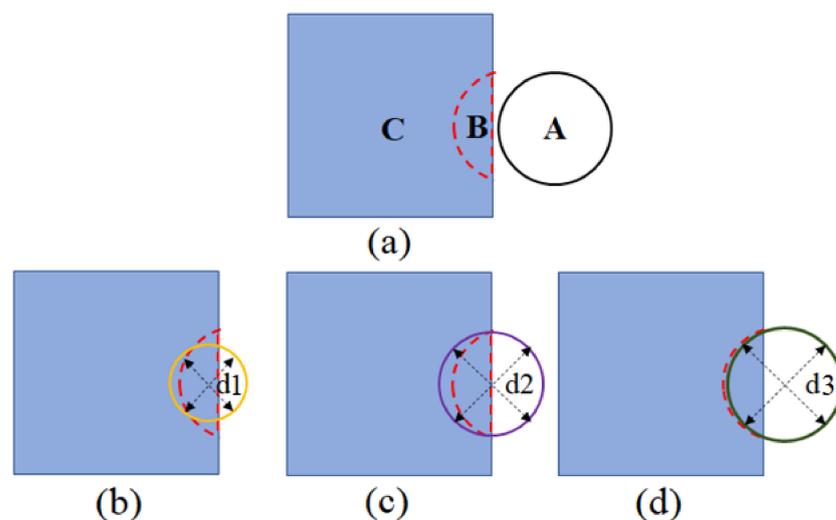


Figure 8. Edge pellet circle fitting. (a) Schematic diagram of the intersection and merging ratio of the edge pellets. (b) Least squares fitting results (with edge contours). (c) Circumscribed circle fitting results. (d) Least squares fitting results (without edge contours).

Figure 8b shows the circle obtained using least squares fitting, where the contour points contain the boundary contour. This fitting method causes the results to be small, and so the boundary contour is removed. Figure 8c shows the circle obtained by removing the sphere boundary and then fitting it with the least squares method, and Figure 8d shows the result obtained by fitting it with a circumscribed circle. The circles obtained in Figure 8c,d are taken as the larger results of the IOU, according to the new evaluation mechanism.

3. Experimental Results and Analysis

A high-speed industrial camera was used to acquire pellet images every 3 s, with images sizes of 1920×1200 , to sample the pellets in the disc pelletizer. The pellets in the images are closely distributed, with a large number of overlapping pellets blocking each other. Figure 9a intercepts four sets of 320×320 -sized images from the original image. Figure 9b is the segmentation result of the manual marked method. Figure 9b–f shows the

pellet segmentation results corresponding to the proposed segmentation algorithm, the K-means algorithm [18], the meanshift algorithm [19], and the watershed algorithm [20].

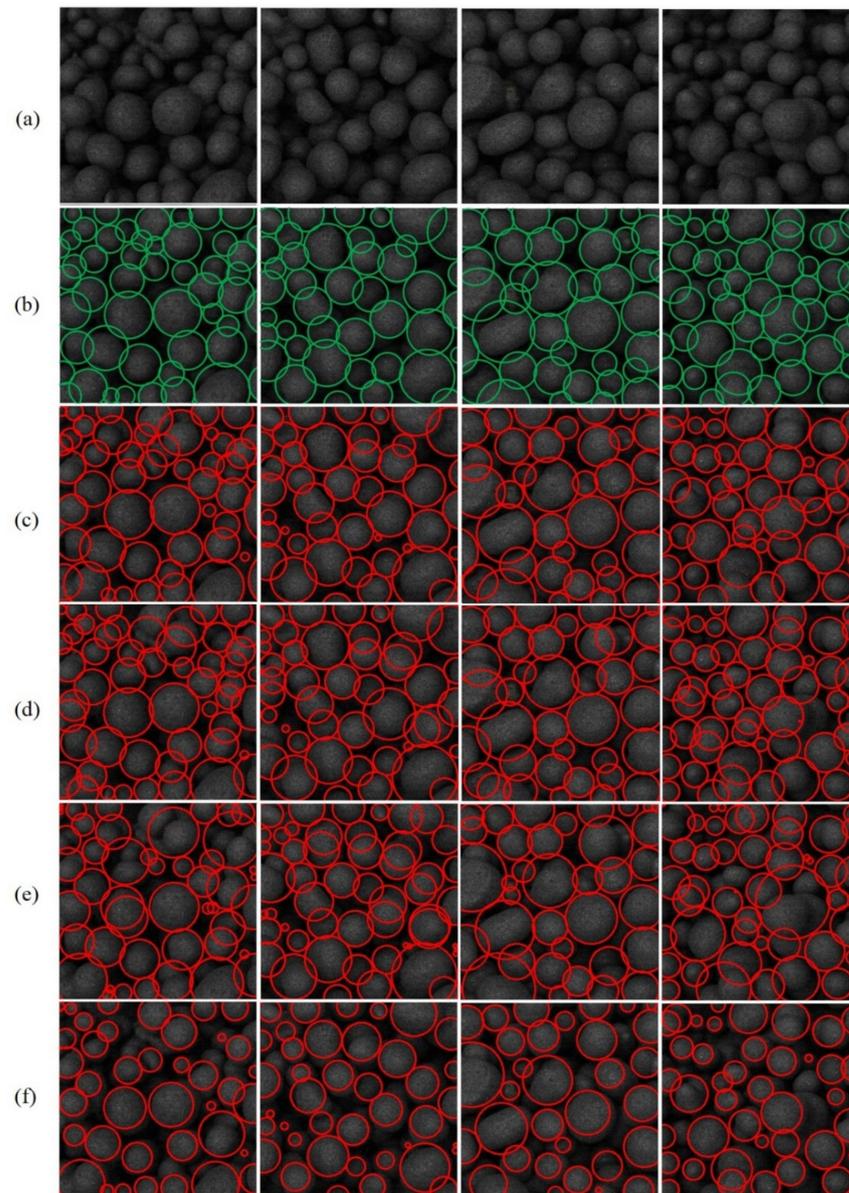


Figure 9. Comparison of different segmentation algorithms. (a) Original image. (b) Manual marked method results. (c) Proposed algorithm segmentation results. (d) K-means algorithm segmentation results. (e) Meanshift algorithm segmentation results. (f) Watershed algorithm.

The segmentation results of the different algorithms are compared with the results of the manual marked method. In the marked watershed segmentation algorithm, the effect of overlapping pellet segmentation is poor, it is easy to lose the pellets with low gray levels, and the segmentation is sensitive to image noise. Because the choice of the marked area in watershed segmentation depends on the choice of the threshold value, when threshold selection is large, the number of marked areas is small, which leads to incomplete separation of overlapping pellets. When the threshold is small, the number of marked areas is large, which easily causes over-segmentation problems. When using the Meanshift segmentation algorithm, partially overlapping pellets are not separated, leading to large pellet size predictions. At the same time, some overlaps are over-segmented, resulting in small pellet size predictions. Because there is no need to specify the number of clusters in the pellets when using the Meanshift segmentation algorithm, it automatically estimates the number of

clusters according to the shape of the connected domain of the particles, but the clustering result depends on the bandwidth setting. If the bandwidth is set too small, the number of clusters will be too large. If the bandwidth is too large, the number of clusters will be too large. When using the K-means segmentation algorithm, the number of overlapping pellet segmentations is relatively accurate, but the pellet size predictions are not accurate enough, which is due to the fact that the K-means clustering algorithm is applicable to data with circular distributions while the overlapping pellet data are unevenly distributed and have an approximate elliptical distribution, and so the inaccuracy of clustering leads to low accuracy in the pellet segmentation. In proposed algorithm, the K value and the initial center of the overlapping pellets are predicted by Gaussian reconstruction, which effectively avoids the problems of inaccurate segmentation and difficult convergence that are caused by the random initialization of the clustering center during the clustering process. In addition, when predicting pellet size, the proposed algorithm firstly determines whether it is an edge pellet or not, and it adopts different prediction methods for the sizes of the edge pellets and the non-edge pellets. Compared with the other algorithms, the proposed algorithm overcomes the problems of over-segmentation and under-segmentation, to a certain extent, and it has obvious advantages in the fitting of the edge pellets.

In order to quantitatively analyze the prediction effects of the different algorithms, five images were randomly selected from 20 320×320 images, and the prediction results were obtained using different algorithms. Figure 10 was obtained by converting the pixel size to the real size based on the ratio of the image to the real object. In Figure 10a, the pellet size distributions predicted by the different algorithms are compared, and the different pellet size intervals are shown as horizontal coordinates and the percentages of pellet numbers are shown as vertical coordinates. The proposed algorithm has a relatively large deviation in the interval with a pellet size of less than 6 mm, which is caused by a small amount of over-segmentation and image noise. In the interval of pellet sizes larger than 6 mm, it is very close to the manual marked ones. Among the four algorithms, the pellet size distribution predicted by the proposed algorithm is the closest to that of the manual marked method, which can better predict pellet size distribution. Figure 10b calculates the average pellet sizes predicted by the different algorithms. In the figure, the different algorithms are shown as horizontal coordinates and the average pellet sizes are shown as vertical coordinates. The average pellet size obtained by the proposed algorithm was 9.74 mm, and the average size obtained using the manual marked method was 10.12 mm, with an error rate of 3.7%, indicating a high prediction accuracy. In Figure 10c, the pellet sizes predicted by the different algorithms are arranged from small to large. The line charts of the pellet sizes predicted by the different algorithms are obtained by taking the serial numbers of the pellets as the horizontal coordinates and the pellet sizes as the vertical coordinates. The line charts of the pellet sizes predicted by the different algorithms are obtained. By comparing the four line charts, it can be seen that the proposed algorithm is the closest to the line chart for the manual marked method, and the predicted pellet sizes have high overall similarity with the manual marked method.

In order to measure the pellet prediction hit rate, an evaluation index of segmentation accuracy was introduced. By comparing the prediction results of the algorithm with manual marked method's results, the segmentation accuracy index AC was defined as:

$$AC = \frac{N_{right}}{N_{right} + N_{miss} + N_{over} + N_{under}} \times 100\% \quad (18)$$

where N_{right} denotes the same number of pellets in the predicted result as the manually marked results, N_{miss} denotes the number of pellets not predicted in the manually marked results, N_{over} denotes the predicted number of overlapping pellets that exceeded the manual marked ones, and N_{under} denotes the predicted number of overlapping pellets that was less than the manual marked ones.

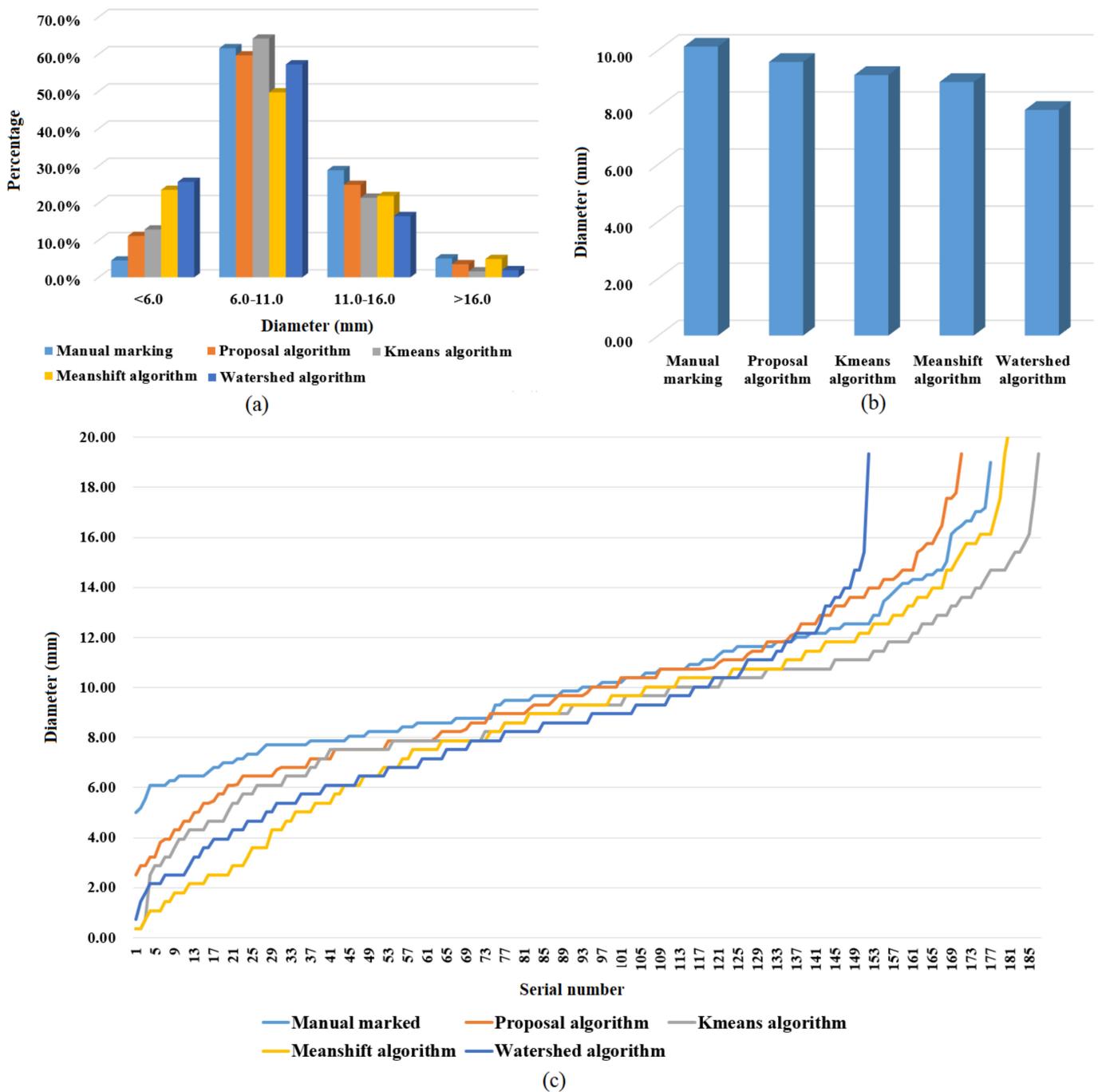


Figure 10. Comparison of the prediction results of the different algorithms and the results of the manual marked method. (a) Histogram of pellet size distribution. (b) Histogram of average pellet size. (c) Line chart of pellet size.

In this paper, 20 320 × 320 images of a total of 960 pellets were selected to gather statistics on the segmentation indexes of the different algorithms, and the watershed segmentation algorithm, the meanshift segmentation algorithm, the K-means segmentation algorithm, and the proposed algorithm were chosen for comparison. The statistical results obtained are shown in Table 1.

Table 1. Comparison of the different segmentation algorithms.

Algorithm	AC/%	N _{Total}	N _{Right}	N _{Over}	N _{Uder}	N _{Miss}
Watershed segmentation algorithm	70.94%	960	681	32	66	181
Meanshift segmentation algorithm	55.94%	960	537	102	198	123
K-means segmentation algorithm	85.21%	960	818	56	15	71
Proposed segmentation algorithm	91.98%	960	883	22	0	56

It can be seen in Table 1 that the marked watershed segmentation algorithm had a high number of missed detections, and some pellets had over-segmentation and under-segmentation problems. The number of missed detections, over-segmentations, and under-segmentations were all large for the meanshift segmentation algorithm. The number of correct segmentations by the K-means segmentation algorithm was large, but there were some problems with missing detections and over-segmentation. The proposed algorithm had the best performance of all the indexes, and the accuracy of pellet segmentation AC was up to 91.98%, which has obvious advantages compared with the other four algorithms.

4. Conclusions

In order to detect pellet size information in the pelletizing process, a pellet size prediction model based on improved watershed and GMM image segmentation was proposed. It could effectively realize pellet size prediction with a large number of overlapping pellets, which is seen when using a disc pelletizer. The marked watershed segmentation algorithm, meanshift segmentation algorithm, and K-means segmentation algorithm were selected for comparison. The experimental results showed that for pellet size prediction, the pellet size distribution and the overall similarity of pellet size obtained by the proposed method were the closest to those obtained by the manual marked method, and the predicted average pellet size error was only 3.7%. For overlapping pellet segmentation, the number of missed detections, over-segmentations, and under-segmentations obtained by the proposed method was the smallest, and the evaluation index of segmentation accuracy was up to 91.98%, which was the highest level of accuracy among the four algorithms. Therefore, the proposed algorithm can effectively detect the particle size information of dense and overlapping pellets, which provides an effective technical means for an intelligent pellet production process.

Author Contributions: Conceptualization, W.M. and L.W.; methodology, W.M.; software, W.M. and L.W.; validation, W.M., L.W., and T.J.; formal analysis, W.M.; investigation, W.M.; resources, W.M.; data curation, W.M.; writing—original draft preparation, W.M.; writing—review and editing, W.M.; visualization, W.M.; supervision, Y.Z. and A.Y.; project administration, W.M.; funding acquisition, A.Y. and W.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Hebei Province (E2022209110) and Basic Research Funds for Universities (JQN2022004).

Data Availability Statement: Not applicable.

Acknowledgments: The authors offer thanks to the Hebei Engineering Research Center for the Intelligentization of Iron Ore Optimization and Ironmaking Raw Materials Preparation Processes for their training and education. Thanks to Zhang Yuzhu and Yang Aimin for their careful guidance. And thanks to Wang Lijing and Jiang Tianyu for their cooperation.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Guo, Y.; Liu, K.; Chen, F. Effect of basicity on the reduction swelling behavior and mechanism of limestone fluxed iron ore pellets. *Powder Technol.* **2021**, *393*, 291–300. [[CrossRef](#)]
- Jägers, J.; Wirtz, S.; Scherer, V. An automated and continuous method for the optical measurement of wood pellet size distribution and the gravimetric determination of fines. *Powder Technol.* **2020**, *367*, 681–688. [[CrossRef](#)]

3. Novikova, A.; Markl, D.; Zeitler, J.A. A non-destructive method for quality control of the pellet distribution within a MUPS tablet by terahertz pulsed imaging. *Eur. J. Pharm. Sci.* **2018**, *111*, 549–555. [[CrossRef](#)] [[PubMed](#)]
4. Lumin, C.A.; Zhen, C.B.; Ansheng, F.C. Image analysis algorithm and verification for on-line molecular sieve size and shape inspection. *Adv. Powder Technol.* **2014**, *25*, 508–513. [[CrossRef](#)]
5. Zhou, S.Y.; Liu, X.Y.; Chen, Y.R. Evaluation of Deep Network-based Methods for Crack Detection of Iron Ore Green Pellet. *ISIJ Int.* **2022**, *62*, 1694–1704. [[CrossRef](#)]
6. Sun, Q.; Zheng, J.; Li, C. Improved watershed analysis for segmenting contacting particles of coarse granular soils in volumetric images. *Powder Technol.* **2019**, *356*, 295–303. [[CrossRef](#)]
7. Tripathi, A.; Kumar, V.; Agarwal, A. Quantitative DEM simulation of pellet and sinter particles using rolling friction estimated from image analysis. *Powder Technol.* **2021**, *380*, 288–302. [[CrossRef](#)]
8. Liu, X.Q.; Peng, W.B.; Zhang, B. A surface cracks detection method for nuclear fuel pellets using an improved fully convolutional network. *J. Nucl. Sci. Technol.* **2021**, *59*, 555–563. [[CrossRef](#)]
9. Wu, X.; Liu, X.; Duan, J. Online size distribution measurement of dense iron green pellets using an efficient and multiscale nested U-net method. *Powder Technol.* **2021**, *387*, 584–600. [[CrossRef](#)]
10. Heydari, M.; Amirfattahi, R.; Nazari, B. An industrial image processing-based approach for estimation of iron ore green pellet size distribution. *Powder Technol.* **2016**, *303*, 260–268. [[CrossRef](#)]
11. Wu, X.; Liu, X.Y.; Sun, W. An image-based method for online measurement of the size distribution of iron green pellets using dual morphological reconstruction and circle-scan. *Powder Technol.* **2019**, *347*, 186–198. [[CrossRef](#)]
12. Zafari, S.; Eerola, T.; Sampo, J. Segmentation of Overlapping Elliptical Objects in Silhouette Images. *IEEE Trans. Image Process.* **2015**, *24*, 5942–5952. [[CrossRef](#)] [[PubMed](#)]
13. Wang, Y.; Tan, Z.P.; Chen, Y.C. An adaptive gravitational search algorithm for multilevel image thresholding. *J. Supercomput.* **2021**, *77*, 10590–10607. [[CrossRef](#)]
14. Cai, Y.; Ming, C.; Qin, Y. Skeleton extraction based on the topology and Snakes model. *Results Phys.* **2017**, *7*, 373–378. [[CrossRef](#)]
15. Křupka, A.; Řiha, K. Minimal prerequisites for measuring two-dimensional contour roundness in a particle classification context. *Powder Technol.* **2015**, *284*, 486–495. [[CrossRef](#)]
16. Liu, D.; Wang, Y.; Tang, Z. A robust circle detection algorithm based on top-down least-square fitting analysis. *Comput. Electr. Eng.* **2014**, *40*, 1415–1428. [[CrossRef](#)]
17. Li, X.M.; Shi, Z.Y. The relationship between the minimum zone circle and the maximum inscribed circle and the minimum circumscribed circle. *Precis Eng.* **2009**, *33*, 284–290. [[CrossRef](#)]
18. Siddiqui, F.U.; Isa, N.A. Mat. Optimized K-means (OKM) clustering algorithm for image segmentation. *Oppo-Electron. Rev.* **2012**, *20*, 216–225.
19. Feng, j.; Zhan, K.; Chen, S.J. Image segmentation method of mine pass soil and ore based on the fusion of the confidence edge detection algorithm and mean shift algorithm. *Gospod. Surowcami Miner.* **2022**, *37*, 133–152.
20. Paiva, P.V.V.; Cogima, C.K.; Dezen-Kempter, E. Historical building point cloud segmentation combining hierarchical watershed transform and curvature analysis. *Parttern Recogn. Lett.* **2020**, *135*, 114–121. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.