WILEY | Hindawi

Research Article

Brown Rice Germ Integrity Identification Based on Deep Learning Network

Shuofeng Li, Bing Li ⬡, Jin Li, and Bin Liu

College of Automation, Harbin Engineering University, Harbin, China

Correspondence should be addressed to Bing Li; libing265@hrbeu.edu.cn

For the quality inspection of brown rice, the segmentation of connected brown rice and the identification of germ integrity are very important. However, there is no better traditional algorithm to achieve better segmentation and recognition results. This paper improves the brown rice (BR) segmentation algorithm based on background skeleton. The candidate matching points are obtained by the background skeleton method, and the optimal matching points are found by the ant colony algorithm. Experimental results show that the proposed segmentation algorithm achieves 96% accuracy, indicating that it can effectively suppress the interference from the endosperm surface. After segmentation is complete, identification of embryo integrity is performed. Firstly, a convolutional neural network (CNN) is built to identify the germ direction; then, the germ direction is normalized; finally, an improved Inception-v3 network is built to identify the germ integrity. On the basis of the Inception-v3 network, additional branches are added to improve the detection accuracy of small objects. In addition, mutual-channel loss and mlpconv are added to enable the model to better approximate the abstraction of the latent space. The experimental results show that the comprehensive recognition accuracy of the proposed algorithm is as high as 94.83%, which is significantly higher than the current mainstream recognition algorithms.

1. Introduction

Brown rice (BR), which is high in nutrients, is currently very popular; as a result, ensuring the quality of BR production is critical [1]. During production and processing, BR germ integrity is primarily manually and subjectively determined, which is a low-efficiency process, and the BR germ integrity cannot be quantitatively analyzed [2]. The BR quality is determined by how well the germ portion can be preserved. The amount of germ preserved in each BR grain has a direct impact on BR processing quality. The less the germ is preserved, the greater the loss of nutrients in BR is [3]. Therefore, it is necessary to not only preserve the germ of BR but also maximally increase the preserved extent of the germ of each BR grain during processing. Thus, it is necessary to examine the BR germ integrity during processing.

Machine vision is being increasingly applied in agriculture. An increasing number of studies have focused on the use of machine vision to examine rice quality. Machine vision can improve rice quality testing efficiency while also overcoming the effects of subjective factors, reducing test errors, and increasing test speed and accuracy. In recent years, Courtois et al. [4] proposed the use of image analysis techniques to assess the breakage of and cracks in parboiled rice grains. Lin et al. [5] proposed a method for splitting touching rice grains based on their contour curvature. Using a support vector machine, Sun et al. [6] evaluated and analyzed the chalkiness of connected rice grains. Payman et al. [7] developed a rice quality index test system based on expert vision.

In this study, BR germ integrity is comprehensively examined by image segmentation and identification. Wang et al. [8] improved the watershed algorithm to address the problem of oversegmentation of targets in image segmentation involving small connected targets. Jia et al. [9] inhibited the effects of noise on segmentation by bitmap cutting and region merging. Some researchers used various edge detection operators in the segmentation and extraction

of images with targets and backgrounds that have significantly different features [10–13]. In a fully convolutional network (FCN), fully connected layers (in a conventional network) are treated as convolutional layers, and deconvolutional layers are connected. An FCN can generate a pixel segmentation map by upsampling [14, 15].

Another method of detection is to use a target detection algorithm to directly identify and classify the targets in the original image, rather than using an image segmentation algorithm to extract single targets. Progress has been made since Krizhevsky et al. [16] proposed AlexNet [17]. At present, in computer vision tasks, convolutional neural networks are used to segment and classify targets [18–20]. In 2012, a major improvement was made to ImageNet [21–23]. Various deep learning methods, represented by convolutional neural networks (CNNs), have been used extensively well in terms of numerous vision tasks, including object detection. The single shot multibox detector (SSD) [23–25] performs exceedingly well in terms of detection accuracy and speed. In computer vision tasks, the use of multi-resolution processing data will improve the overall level of image segmentation and classification [26, 27]. The SSD was proposed in 2016 to address the problem of you-only-look-once (YOLO) algorithms in target positioning having insufficient detection accuracy [28–30]. The main idea of the SSD is the dense and uniform sampling of an image at various locations. The SSD was established on the concept of anchor points in faster region-CNNs (R-CNNs) [31, 32]. In the SSD, various prior frames are used to predict the proportion and length-to-width ratio (LWR) of the boundary frame of an object during sampling. The features are then extracted, classified, and regressed using a CNN. There is only one step in the entire procedure. In order to detect targets, the SSD employs a pyramidical feature layer group. The shallow feature map has a smaller receptive field, so it can predict a small target, while the deep feature map has a larger receptive field, so it can predict a large target. In the SSD, however, feature layers of various scales are unrelated to one another, and feature complementarity between feature layers is lacking. As a result, the SSD is relatively inefficient in detecting general and small targets. In August 2020, Li et al. [33] investigated a deep learning based algorithm for marine target detection and segmentation with a 32.7% improvement in detection accuracy. In July 2021, Krishnamoorthy et al. [34] used the Inception ResNetV2 model to identify infectious organisms in rice plant leaves with a detection accuracy of 95.6%. In August 2021, Yang et al. [35] proposed a temporal feature-based segmentation (TFBS) model for accurate crop mapping using time-series SAR images. In October 2021, Lan et al. [36] proposed two improved recognition models MobileNetV2-UNet and FFB-BiSeNetV2 for the identification of rice weeds with an accuracy of 93%. In the same month, Wang et al. [37] proposed a small target detection algorithm based on Mask R-CNN model with a detection accuracy of 66.5%. In December 2021, Guo et al. [38] proposed an image segmentation and deep learning based method to automatically identify rice grains and calculate rice fructification rate with good results. In the same month, Wang and Qiu [39] used an unsupervised classification method for fast localization of small targets to improve detection efficiency.

In the process of processing, the quality testing of brown rice is of great significance for retaining nutrients and reducing food waste. Among them, the segmentation of connected BR and the detection of germ integrity are extremely important. However, the above algorithms are unable to effectively identify all the rice grains in an image and have difficulties with the connected rice grains in particular.

The main contributions of this study are summarized as follows:

(1) The BR germ integrity is classified by image segmentation combined with image identification. Single rice grains are extracted from connected rice grains using the proposed segmentation algorithm. Then, the BR germ integrity is classified based on an improved Inception-v3 model.

(2) The background skeleton-based segmentation algorithm is improved [40]. The region growing algorithm and ant colony algorithm [41–44] are integrated into reference-point matching to improve the matching accuracy for skeleton endpoints.

(3) Two neural network (NN) models are used to identify BR germ integrity. First, the visual geometry group (VGG) model is used to identify and classify the directions of the BR germs. Then, the Inception-v3 model is structurally improved to increase the identification accuracy.

2. Proposed Work

2.1. Preprocessing. There are two steps to preprocessing a BR image. (1) After binarization, small, scattered white areas are often generated in an image due to the miscellaneous items on the platforms where the husks, bran, and rice are placed during production and processing as well as noise from some image processing procedures. These small areas may subsequently affect the identification of the effective connected areas and thus need to be eliminated. (2) The training set is image preprocessed to enhance the features of the germ portion in the images in order to improve subsequent training performance for the germ integrity of single rice grains.

2.1.1. Image Preprocessing. In the first step of image preprocessing, conventional spatial and frequency domain filters cannot simply and effectively eliminate the small areas, whereas a morphological filter can easily eliminate them. To eliminate the small areas, a structural element with a suitable dimension can be used to erode the original image. After erosion, these small areas can be eliminated. An erosion operation, on the other hand, enlarges the voids and separates the previously connected regions. A rice area can also be reduced by an erosion operation. To eliminate the negative effects of an erosion operation, a dilation operation can be subsequently performed. The combination consisting of

first an erosion operation and a subsequent dilation operation is referred to as an opening operation. The " ∘ " in the following equation is the operator of an opening operation and signifies that a structural element is first used to erode the image and then dilate the results of the erosion operation. We have

$$X \circ B = (X \ominus B) \oplus B. \tag{1}$$

A morphological filter is used to remove the small white area in the following left image. A 5×5 square is used as a structural element. The comparison image shows that the small white area is removed. The results are shown in Figure 1.

2.1.2. Image Segmentation Preprocessing.
In the second step of image segmentation preprocessing, the training set is subjected to top-hat transformation to enhance the germ features in the image. This step involves reducing the grayscale levels of the dark regions and increasing the grayscale levels of the bright regions using the following equation to increase the contrast:

$$h = f - (f \circ b). \tag{2}$$

Figure 2(a) shows the original image. Figure 2(b) shows the result of politeness transformation. As a result, after the contrast stretching process, there is a sharper contrast between the germs and other parts of rice. This is beneficial to the subsequent germ integrity identification training.

2.2. Image Segmentation.
To segment small targets, this paper uses an improved image segmentation algorithm based on background skeleton features. In addition, the endpoint matching process in this algorithm is improved. To facilitate the analysis of BR grain features, it is necessary to separate the BR and the background based on image preprocessing. The background in the image is black, and its grayscale level differs relatively significantly from that of the BR in the image. To segment the BR grains, this study first uses a threshold segmentation algorithm. The image is segmented using the threshold segmentation algorithm based on the grayscale level. This method is applied to distinguish the BR foreground image from background image and then converts the grayscale image into a binary image. The segmentation equation is shown as follows:

$$g(x, y) = \begin{cases} 1, & m(x, y) > N, \\ 0, & m(x, y) \leq N, \end{cases} \tag{3}$$

where $m(x, y)$ is the grayscale image, $g(x, y)$ is the binary image, and N is the segmentation threshold (N is calculated using a classic and highly robust algorithm—the maximum between-class variance (Otsu) algorithm). A threshold is calculated using the Otsu algorithm primarily on the basis of maximum between-class difference between grayscale levels on the BR and the background. Based on the grayscale probability distribution, the Otsu algorithm for digital images is represented by the following equations:

$$w_0 = \sum_{i=0}^{z} P_i,$$

$$w_1 = \sum_{i=z+1}^{L-1} P_i, \tag{4}$$

$$\mu_0 = \sum_{i=0}^{z} \frac{iP_i}{w_0},$$

$$\mu_1 = \sum_{i=z+1}^{L-1} \frac{iP_i}{w_1}, \tag{5}$$

$$\mu = w_0\mu_0 + w_1\mu_1, \tag{6}$$

$$\sigma^2 = \left[w_0 \left(\mu_0 - \mu \right)^2 + w_1 \left(\mu_1 - \mu \right)^2 \right], \tag{7}$$

where z is the set grayscale segmentation threshold (grayscale levels are classified into classes A and B), P_i is the probability of each gray level, w_0 and w_1 are the probability sums of grayscale levels for classes A and B, respectively, the probability measures of classes A and B grayscale levels are μ_0 and μ_1, respectively, and σ^2 is the between-class variance. The optimum segmentation threshold T that leads to a maximum σ^2 is calculated by traversing $z = [0, 255]$.

The segmentation threshold for the grayscale BR image is calculated using the Otsu algorithm. Grayscale levels greater and less than the segmentation threshold in the image are set to 1 and 0, respectively. Thus, a binary BR image is obtained. Figure 3 is a binary BR image obtained by processing a grayscale BR image.

Background segmentation is followed by region segmentation. Based on the similarities between the spatial regions, in the region segmentation process, the regions are merged according to the similarity criteria. Below, the region growing algorithm is primarily used to segment the regions in the binary BR image to complete the marking and characterization of the regions.

In region marking, the same connected components in an image are assigned the same mark, and different connected regions are differentiated with different marks. A binary image is used to mark the connected regions. The background is labeled with the same label (generally 0). Finally, the image is transformed into a mark matrix.

The main algorithm used in region marking is region growing, which was originally used in grayscale image region segmentation. The idea of this algorithm is as follows. A seed pixel is set or found in the region that needs to be segmented. The region is expanded with this seed pixel as its initial point. Subsequently, the pixels within the neighborhood of the seed pixel are examined to determine their similarities to the seed pixel. The pixels that meet the similarity criteria are merged into the seed region. After that, until no more pixels meet the similarity criteria, the newly formed region is used as a seed for the similarity judgment and region merging steps. Finally, the merged region is one connected region that is independent.
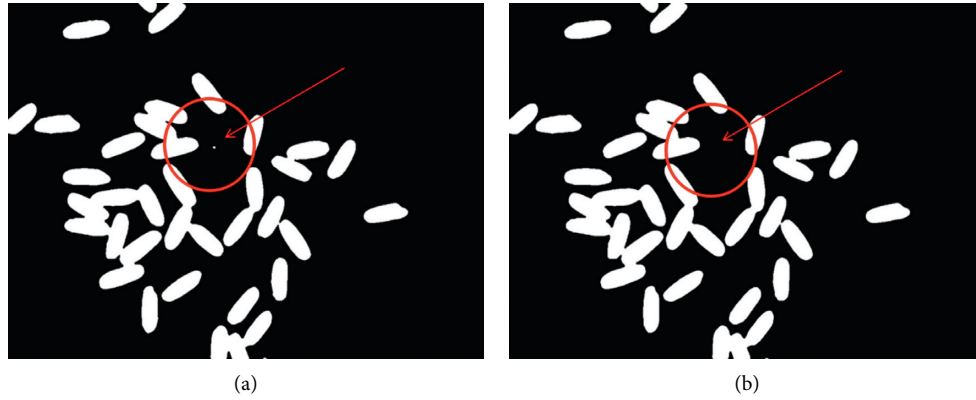
FIGURE 1: Comparison of images before and after morphological filtering. (a) The original image. (b) The image processed by morphological filter.
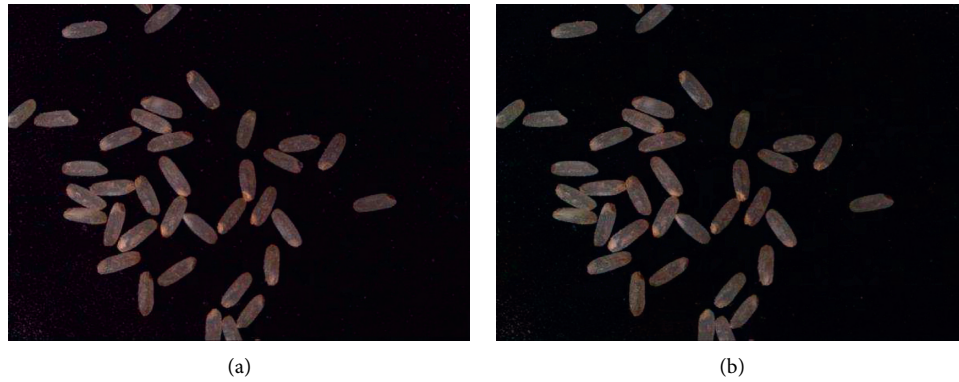

FIGURE 2: Image comparison after segmented transformation. (a) The original image. (b) The image processed by top-hat transformation.

FIGURE 3: Binary image of brown rice.

After subjecting the binary BR image to the background segmentation and region marking procedures, it is necessary to further segment different regions by marking their geometric properties. A method for describing the region features, which is involved in the extraction of the shape features from the binary BR image, is described by the following. The BR features include the perimeter, area, direction of the long axis, and center of gravity. The perimeter of a BR grain is the length of the boundary of the region. On a two-dimensional plane, the perimeter of a region can be approximately calculated by counting the number of pixels

of its boundary. Sometimes, to obtain a more accurate representation, pixels are treated as points, and the bend and linear boundaries of a right angle are distinguished by chain-code representations. When the chain-code value is an odd number (i.e., bend boundary), the length is denoted by $\sqrt{2}$. When the chain-code value is an even number (i.e., linear boundary), the length is denoted by 1. Thus, the perimeter p is represented by the following:

$$p = N_e + \sqrt{2}N_o, \tag{8}$$

where N_e is the number of chain codes with even-number values and N_o is the number of chain codes with odd-number values.

The area of a region is its basic property. For the region R in the BR image, its area is calculated based on the number of pixels contained within it. The area of the region R is represented by the following:

$$S_R = \sum_{(x,y)\in R} 1. \tag{9}$$

The above equation calculates the cumulative number of pixels in all the regions plus 1.

The position of a target region in the image is characterized by its center of gravity. In the binary BR image, all the pixels within the region have a value of 1 and thus follow a

uniform distribution. Let (x_i, y_i) be the coordinates of the position of a pixel in the image. The equation for calculating the center of gravity is as follows:

$$\overline{x} = \frac{1}{S_R} \sum_{(x_i, y_i) \in R} x_i,$$

$$\overline{y} = \frac{1}{S_R} \sum_{(x_i, y_i) \in R} y_i, \tag{10}$$

where $\overline{x}$ and $\overline{y}$ are the coordinates of the position of the center of gravity and S_R is the area of the region.

During BR identification correction, the direction of each BR grain plays an important role and is the main feature parameter for extraction correction. Under normal circumstances, for a long and narrow region, the direction of the long side of its minimum bounding rectangle is selected as the direction of its long axis.

$$\mu_{pq} = \sum \sum_R (x - \overline{x})^p (y - \overline{y})^q f(x, y),$$

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right), \tag{11}$$

where μ_{pq} is the central moment of the region, $f(x, y)$ is the pixel value of the point (x, y) in the region, and θ is the angle of the direction of the long axis.

The connected regions in the binary BR image are marked using the region growing algorithm, as shown in Figure 4. In this paper, the connected regions of the image after binarization are numbered, and each number is at the center of each connected domain.

The connected rice grains are segmented after the BR image has been subjected to the background threshold segmentation and region segmentation procedures. To segment the connected rice grains in the image, it is necessary to first determine the positions of the connected regions. The connected regions in the BR image are marked and described during the region segmentation procedure. Two of them were selected as examples for analysis. The original image is shown in Figure 5(a). First, the image is binarized, as shown in Figure 5(b). Then, according to the calculation, the minimum convex hull (MCH) (i.e., minimum convex polygon) of each connected domain image is obtained. The MCH image is shown in Figure 5(c). Furthermore, the MCH region's area is calculated using the region marking method. Equations (12) and (13) can be used to determine the connected regions.

$$f(i) = \begin{cases} 1, & S(i) < 0.85 \text{ and } A(i) > H, \\ 0, & \text{else,} \end{cases} \tag{12}$$

$$\text{solidity} = \frac{\text{area}}{\text{convex}}, \tag{13}$$

where $i \in [1, m]$, $f(i)$ is applied to judge the region connection condition ($f(i) = 1$ means that the current region is a connected region; $f(i) = 0$ means the current region is a nonconnected region), $A(i)$ is the area of the region, convex is the area of the MCH in the region, S is the ratio of the area of the region to the area of the MCH, and H is the area threshold (H is set based on the area of a single rice grain in the binary image; in this study, H is set to 2,400, which is a little bit smaller than the area of a single rice grain in the binary image).

After determining the connected regions, it is necessary to calculate the segmentation endpoints. Thus, an operation is performed on the binary image and the MCH image using equation (14). A binary image of the background image for connected rice grains within the MCH region is thus obtained, as shown in Figure 5(d).

$$C(x, y) = A(x, y) - B(x, y), \tag{14}$$

where $A(x, y)$ is the binary image of the MCH in the connected region, $B(x, y)$ is the binary image of connected BR grains, and $C(x, y)$ is the binary image of the background of the image of the connected BR grains.

In the binary image, the connected features between BR grains images have been completely extracted, so the background image is represented by a curve. The topographic and connection features in the background image are preserved in the skeleton image, as shown in Figure 5(e). The following shows the calculation of the skeleton:

$$S(A) = \bigcup_{n=0}^{N} S_n(A), \tag{15}$$

$$S_n(A) = (A \ominus nB) - [(A \ominus nB) \circ B], \tag{16}$$

$$A \ominus nB = ((\cdots (A \ominus B) \ominus B) \ominus \cdots) \ominus B, \tag{17}$$

where $\ominus$ is defined as the morphological erosion operator, the background image is represented by A, the skeleton of the background image is represented by $S(A)$, the structural element is represented by B, N is the value in $(A \ominus nB) \neq \varnothing$ and $(A \ominus (n+1)B) = \varnothing$, $S_n(A)$ is the subset of n_{th} skeleton, and $A \ominus nB$ signifies the use of B to erode A for an n number of consecutive times.

The skeleton image endpoint look-up table function is used to generate a neighborhood pixel value look-up table Lut. Subsequently, the segmentation endpoints are searched for in the look-up table for the background skeleton feature image in Figure 5(e). The segmentation points in the image's coordinates are calculated. The segmentation endpoints in Figure 5(f) are the white and shaded endpoints. The skeleton image endpoint look-up table function is described by the following: in terms of a binary image, if there is one and only one pixel with the value of 1 within the 3×3 neighborhood of a pixel with the value of 1, the function returns the value of 1; otherwise, the
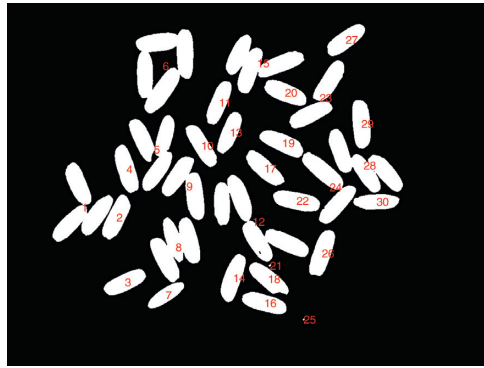
Figure 4: Rice area marker image.
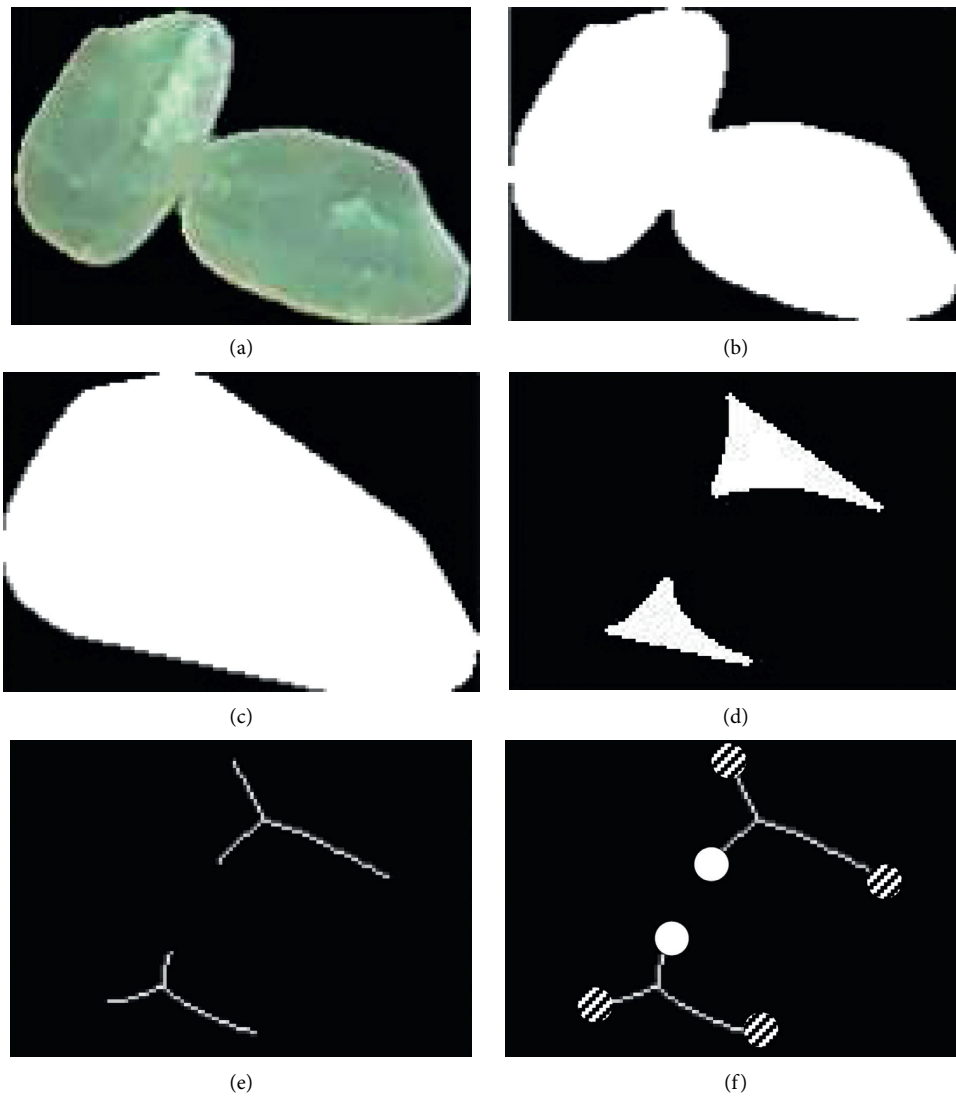

(a)

(b)

(c)

(d)

(e)

(f)

Figure 5: Schematic diagram of feature segmentation endpoints. (a) Original image. (b) Binary image. (c) MCH. (d) Background image. (e) Skeleton image. (f) Skeleton endpoints.

function returns the value of 0. Detailed calculation method is shown as follows:

$$M(x, y) = \begin{bmatrix} A_{(x-1,y-1)} & A_{(x,y-1)} & A_{(x+1,y-1)} \\ A_{(x-1,y)} & A_{(x,y)} & A_{(x+1,y)} \\ A_{(x-1,y+1)} & A_{(x,y+1)} & A_{(x+1,y+1)} \end{bmatrix}, \quad (18)$$

$$\text{Sum}M(x, y) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} A_{(i,j)}, \quad (19)$$

$$f_M(x, y) = \begin{cases} 1, & SM(x, y) = 2 \text{ and } A_{(x,y)} = 1, \\ 0, & \text{other}, \end{cases} \quad (20)$$

where $M(x, y)$ is a 3×3 neighborhood with $A_{(x,y)}$ as the center, $A_{(x,y)}$ is a pixel with the value of 1 in the background skeleton feature image, $SM(x, y)$ is totaled elements in $M(x, y)$, and the skeleton endpoint extraction function is represented by $f_M(x, y)$.

The background skeleton endpoints are fixed in binary image. Because some skeleton endpoints are formed by the MCH boundary, not all of these skeleton endpoints can be defined as segmentation endpoints. Thus, the edges of the minimum convex image are detected based on the Sobel operator equations (21)–(23) to extract its boundary curve. Then, the minimum Euclidian distance between the boundary curve of MCH and each skeleton endpoint is calculated. The results form a minimum Euclidian distance set. The nonsegmentation endpoints are then removed based on

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

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} A, \quad (21)$$

$$|G| = \sqrt{\left(G_x^2 + G_y^2\right)}, \quad (22)$$

$$\phi = \arctan\left(\frac{G_x}{G_y}\right), \quad (23)$$

$$J = \{d_1, d_2, \ldots d_N\}, \quad (24)$$

$$Td_i = \begin{cases} 1, & d_i > th, \\ 0, & \text{other}, \end{cases} \quad (25)$$

where A is the binary image of the MCH, G is the gradient of the image, G_x and G_y are, respectively, represented by the gradient of the image in x-axis and y-axis direction. ϕ is defined as the direction angle of gradient, J is defined as the set of the minimum Euclidian distances between endpoints

and MCH, N is defined as total of endpoints, th is the segmentation endpoint distance judgment threshold (th is set to 4), $Td_i = 1$ means that the point is a segmentation endpoint, and $Td_i = 0$ means that the point is not a segmentation endpoint. In Figure 5(f), the minimum distances between shaded endpoints and boundary contours are no more than th. Thus, these shaded endpoints need to be removed. In comparison, the white endpoints are determined to be segmentation endpoints.

2.3. Feature Point Matching Steps

(1) Select a segmented end point as the reference end point in the calculated segmented end point library, and confirm the segmented matching area in accordance with the position characters of the reference end point in image. First, perform morphological expansion operation in the skeleton area where the segmentation endpoints are located. Then, the boundary features of the image are extracted (dotted line T in Figure 6). In order to more distinctly show the spatial positional relationship in the image, an original contour curve of the background region of binary image is plotted (solid line T' in Figure 6). In Figure 6, X is the selected segmentation endpoint as reference. Through traversal of the pixels of the boundary curve T, the boundary point (X_2) closest to the reference segmentation endpoint is determined between the vector XX_2. Then, the boundary curve where the boundary point X_2 is located is smoothed by filtering boundary spike noise, as shown in equation (27). Based on smooth handling of contour curve, the points $(X_1$ and $X_3)$ near X_2 are selected. Thus, matching direction angle of the reference endpoint is determined on the basis of X_1, X_2, and X_3. The specific algorithm of matching angle is calculated by the following equations:

$$A_x = A \oplus B, \quad (26)$$

$$X_i = \frac{A_{(i-2)} + A_{(i-1)} + A_{(i)} + A_{(i+1)} + A_{(i+2)}}{5}, \quad (27)$$

$$X_{\text{mid}} = (x_{\text{mid}}, y_{\text{mid}}) = \left(\frac{x_1 + x_3}{2}, \frac{y_1 + y_3}{2}\right), \quad (28)$$

$$\theta = \arctan\left(\frac{y_2 - y_{\text{mid}}}{x_2 - x_{\text{mid}}}\right) + m\pi, \quad (29)$$

$$m = \begin{cases} 1, & y_2 - y_{\text{mid}} \geq 0 \,\&\, x_2 - x_{\text{mid}} < 0, \\ 0, & x_2 - x_{\text{mid}} \geq 0, \\ -1, & y_2 - y_{\text{mid}} < 0 \,\&\, x_2 - x_{\text{mid}} < 0, \end{cases} \quad (30)$$

where A_X is defined as dilated skeleton image, A is original skeleton of binary image, B is defined as a 2-radius circular structural element, $A_{(i)}$ is defined as coordinates of the contour boundary, X_i is defined as
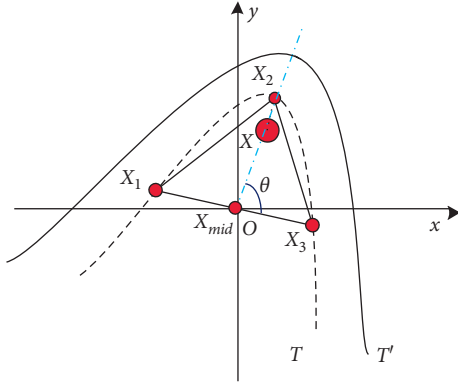
FIGURE 6: Matching angle diagram.

FIGURE 7: Baseline endpoint matching diagram.

coordinates of boundary points after smooth handling of boundary curve, X_{mid} is defined as the midpoint between X_1 and X_3, and θ is defined as the angle of the vector $\overrightarrow{X_{\text{mid}} X_2}$ and the horizontal x-axis (i.e., the direction angle of the reference point; $\theta = [-\pi, \pi)$).

(2) The selected reference point is matched with the corresponding segment end point, referring to the angle of the end point, and $\pm 45°$ is selected as the matching radiating region (the shaded region in Figure 7). In Figure 7, the shaded region shows the search area for the points to be matched.

During the matching process, a seed image is first generated using the region generation algorithm. The seed image can be obtained by eroding the binary image multiple times with structural elements differing in shape and size. The seed image, unlike the binary image obtained during the background threshold segmentation process, does not place a premium on BR contour integrity. In the generation of the binary image needed for the seed image, there is no need to reduce the contrast, because this also reduces the clarity of the boundaries and, thereby, causes the boundaries to connect with one another. To avoid insufficiently clear gaps between the BR grains during binarization due to uneven lighting, the image can be subjected to a top-hat transformation with a relatively large structural element first. The following equation shows the top-hat transformation equation (31), where f is the input image, B is the structural element, and g is the output image. Figure 8(a) shows the binary image of connected BR grains. Figure 8(b) shows the obtained seed image.

$$g = f - (f \circ B). \tag{31}$$

Afterwards, a reference point randomly selected is considered as the center of the circle and the length of rice grain is regarded as the radius with the purpose of searching all reference points in the circle. Meanwhile, the image is considered as the grid map and the pixels of reference points are treated as coordinates; then the routine is planned by the ant colony algorithm. The shortest path of all reference points in the ergodic circle is calculated (the starting point and the ending point are not connected).
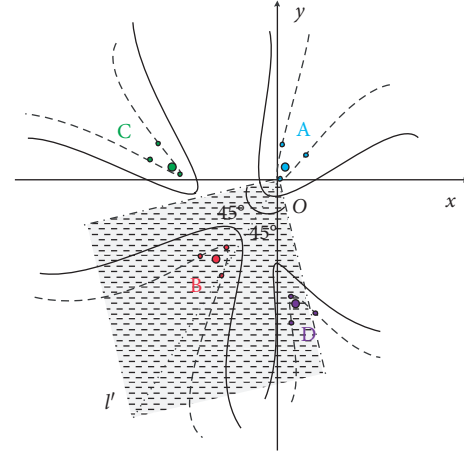
The ant colony algorithm is constructed by a kind of biological heuristic optimization algorithm with regard to cooperative behaviors of ant colony in foraging, because its significant globally researching capability is widely applied in combinational optimization problems. Based on distance matrix D, the ant colony algorithm is used to calculate a group of patrol order X for minimization of patrol distance function $V(X)$.

The number of patrol points δ and the number of ants in its colony m and distance matrix D are known; then all points passed by ant K are recorded in tabu list t_k. After the first loop, the tabu list is reset to zero; then ants are free to select again. Based on pheromone and the distribution of heuristic information, the state transition probabilities of ants are identified in its movement process, as expressed in

$$p_{ij}^k = \begin{cases} \dfrac{\tau_{ij}^\alpha(t)\eta_{ij}^\beta(t)}{\sum_{s\in\alpha_k}\tau_{is}^\alpha(t)\eta_{is}^\beta(t)}, & (j \in \alpha_k), \\ \\ 0, & (\text{other}), \end{cases} \tag{32}$$

$$\eta_{ij}(t) = \frac{1}{d_{ij}}. \tag{33}$$

Here, p_{ij}^k is state transition probability, the larger number indicates the greater probability of selecting the next point for ants. i and j indicate, respectively, the current node and the next node. $\tau_{ij}(t)$ is the amount of pheromone on the line of i and j at t moment. $\eta_{ij}(t)$ is heuristic function. a_k is searchable nodes collection not in tabu list. a is pheromone index of relative importance. β is visibility index of relative importance. d_{ij} is the data of line i and line j in distance matrix D. According to equations (1) and (2) to execute the choice of next node until all patrol points are traversed, the length of motion path for each ant is recorded as $L_k (k = 1, 2, \ldots m)$; meanwhile, the current shortest path L_{min} and patrol order X are recorded. Then the pheromone is updated pursuant to following equation and $\Delta\tau_{ij}$ is reset to 0, as expressed in

(a)   (b)

FIGURE 8: Baseline endpoint matching diagram. (a) The binary image. (b) The obtained seed image.

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij}(t), \tag{34}$$

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t), \tag{35}$$

$$\Delta\tau_{ij}^{k}(t) = \begin{cases} \dfrac{\vartheta}{L_k}, & (\text{The } k_{th} \text{ ant passes through node } (i, j)), \\ 0, & (\text{The } k_{th} \text{ ant doed not passes through node } (i, j)). \end{cases} \tag{36}$$

Here, pheromone concentration $\tau_{ij}(t)$ is pheromone concentration in previous t time. $\Delta\tau_{ij}$ is the increasing number of pheromone in the node within unit interval. $\Delta\tau_{ij}^{k}(t)$ is the increasing number of pheromone ant K. ρ is volatility coefficient of pheromone. L_k is the distance from the starting point to the end walked by ant K. θ is enhancement coefficient of pheromone. When iterations are up to its maximum Ncmax, jump out of the loop and then output the optimal patrol path X at this time as the final result. The shortest path planned by ant colony algorithm is as shown in image.

After shortest-distance path planned, the first step is to identify which two consecutive points are matching endpoints in the path and then select two points according to the order planned by the ant colony algorithm. Through the shortest distance between the reference endpoint and seed image, the nearest seed from two endpoints is judged whether it is the same one or not. If the seeds are the same one, then they are matched endpoints. Otherwise, they are unmatched endpoints. The specific judgment method is as expressed in Figure 9.

The result as shown in Figure 10 is concluded from the path planning and distance judgment of seeds. In the image, the endpoints of black line are matched endpoints. Then the rest of endpoints are reused to execute endpoint matching through ant colony algorithm until the endpoints are finally matched.

2.4. Image Identification. This study presents a CNN model for identifying BR germ positions as well as a CNN model for classifying BR germ integrity. In addition, these two models are used in combination to calculate the BR germ integrity.

2.4.1. Identification of the Germ Direction. After grain image segmentation, the images of the single BR grains show that the germ is generally located in the upper left, upper right, lower left, or lower right corner of the grain. The direction of the germ is relatively notable. Furthermore, the images of single BR grains have already been corrected horizontally along the long axis. As a result, the germ can only go in four different directions. Thus, there are only four types of VGG network outputs. Therefore, no complex feature parameters are needed, as shown in Figure 11.

2.4.2. Germ Integrity Identification. By position identification, the images of the BR germs are rotated to ensure that they are in the same direction. Then, the germ integrity is identified using an optimized Inception-v3 model. The simplest method for improving a network is to increase its depth and breadth and increase the numbers of hidden layers and neurons in each layer. However, there are some problems with this method. This method leads to a larger parameter space, renders fitting easier, and requires more computational resources. The deeper the network is, the easier it is for the gradient to vanish, and the more difficult it is to optimize the network. The core of the Inception-v3 model involves the use of 1×1, 3×3, and 5×5 small convolutional layers to replace the large convolutional layers of the network. This can significantly reduce the number of weighting parameters. Figure 12 shows the results obtained using the Inception-v3 model.

During the training process, the activation input value of the deep NN deepens prior to nonlinear transformation, resulting in the movement of the NN distribution. The convergence is slow during training because the overall distribution gradually becomes a nonlinear function. This
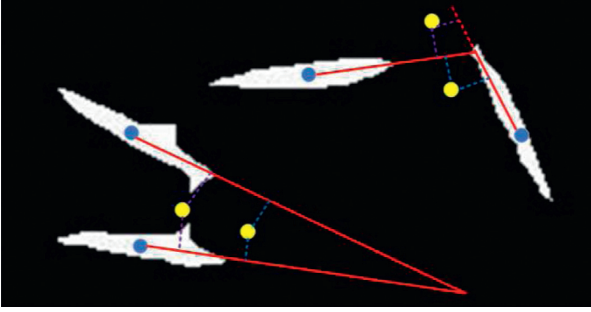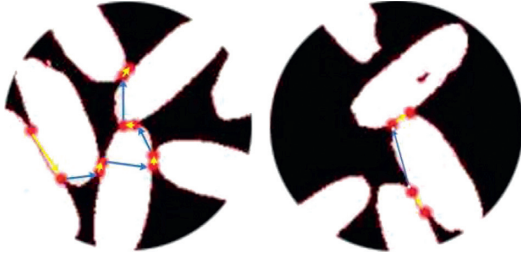
FIGURE 9: Match verification diagram.

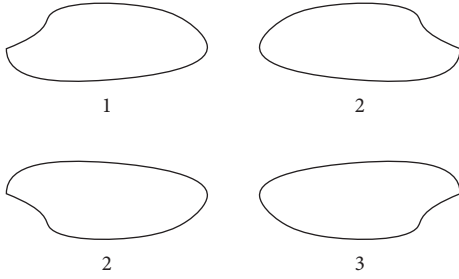FIGURE 10: Ant colony algorithm path planning image.

FIGURE 11: Image of the germ position.

results in the vanishing of the gradient of the low-level NN during backpropagation.

The batch normalization method normalizes each data output (mean: 0; variance: 1) as well as the mean and standard deviation of the data after the convolution pool, as shown in the following equation:

$$\hat{x}^{(k)} = \frac{x^k - E\left[x^k\right]}{\sqrt{\mathrm{Var}\left[x^k\right]}}. \tag{37}$$

The Inception-v3 model is used in combination with a proposed loss function, which is referred to as the mutual-channel loss (MC-loss) function, to distinguish the features of different channels. In addition, the MC-loss function is used to limit the distribution of the channels. The MC-loss function contains two channel-related components, namely, a discriminality component and a diversity component [28]. Figure 13 shows the overall model framework.

The input image $X \in R^{W \times H}$ is used as the input of the basic network. The depth feature $F \in R^{N \times W \times H}$ corresponding to the image is extracted, where N is the number of channels of the depth feature and W and H are the width and

height of each feature map, respectively. In addition, it is required that N equal $c \times \xi$, where c is the total number of classes in the training dataset and ξ is the number of feature maps allocated to each class. Therefore, the nth-dimensional feature map of the depth feature F can be represented by $F_n \in R^{WH}$, $n = 1, 2, \ldots, N$. Moreover, the depth feature of the class i can be represented by $F_i \in R^{\xi \times WH}$, $i = 0, 1, 2, \ldots, c\text{-}1$, as expressed in the following equation:

$$F_i = \left\{F_{i \times \xi + 1}, F_{i \times \xi + 2}, \cdots, F_{i \times \xi + \xi,}\right\}. \tag{38}$$

Inputting two optimization objectives with different purposes into two different branches facilitates their calculation and optimization. In Figure 13, the cross-entropy loss branch treats the feature as input. Through a fully connected layer, the conventional classification loss L_{CE} is calculated. The cross-entropy loss can drive the network to extract the meaningful discriminality information and focus on the global discriminality region. In addition, the MC-loss branch supervises the network, allowing it to learn features and focus on different local discriminality regions. Subsequently, an ultimate loss function is obtained by adding the MC-loss weighted by μ and L_{CE}. Thus, the loss function of the whole network is as follows:

$$\mathrm{Loss}(F) = L_{\mathrm{CE}}(F) + \mu L_{\mathrm{MC}}(F). \tag{39}$$

In addition, the MC-loss is obtained by a weighted addition of a discriminality component loss L_{dis} and a diversity component loss L_{div}. Thus, the MC-loss is defined as follows:

$$L_{\mathrm{MC}}(F) = L_{\mathrm{dis}}(F) + \lambda L_{\mathrm{div}}(F). \tag{40}$$

The left boxes in Figure 14 show the procedure for the distinguishability component. The diversity component is used to measure the similarities between the feature channels. As shown by the right side of Figure 14, the diversity component can drive the feature channels of the same group to become different, as shown in Figure 14.

The orthogonal classification layer adds an improved dropout, which reduces unnecessary link items of various categories, as shown in the following equation:

$$r = \mathrm{soft\,max}\left((M \bullet W)v\right). \tag{41}$$

M is a mask matrix of W, constant diagonal matrix, as expressed in the following equation:

$$M = \begin{bmatrix} M_{1,1} & 0_{1,2} & \cdots & 0_{1,k} \\ 0_{2,1} & M_{2,2} & \cdots & 0_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{k,1} & 0_{k,2} & \cdots & M_{k,k} \end{bmatrix}. \tag{42}$$

Based on Inception-v3 model, two branches are added. The first branch is the original germ rice image branch. The second branch is adding mlpconv to the Inception-v3 model which can help the model to better approximate the abstract representation of the potential space. The simple convolutional neural network cannot extract the abstract features of the germ well, and the inception structure is too
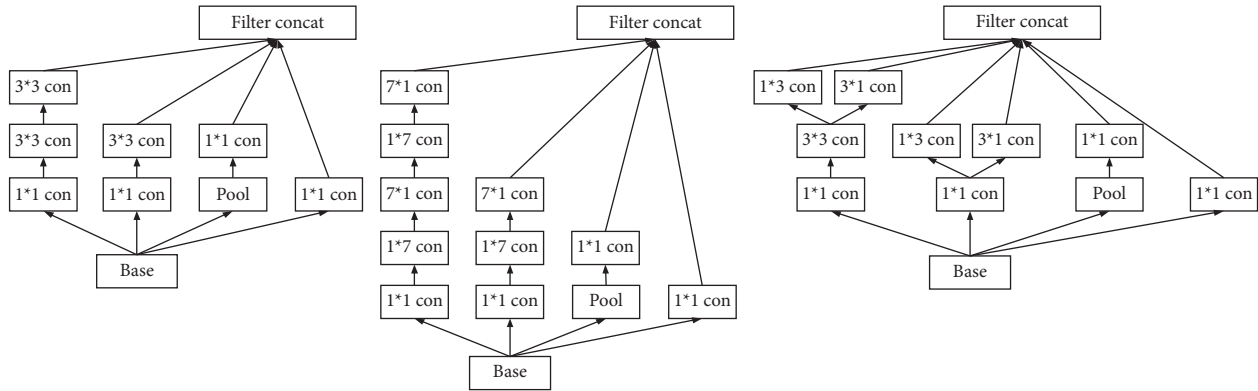
FIGURE 12: The Inception-v3 model (Block1, Block2, and Block3 from left to right).
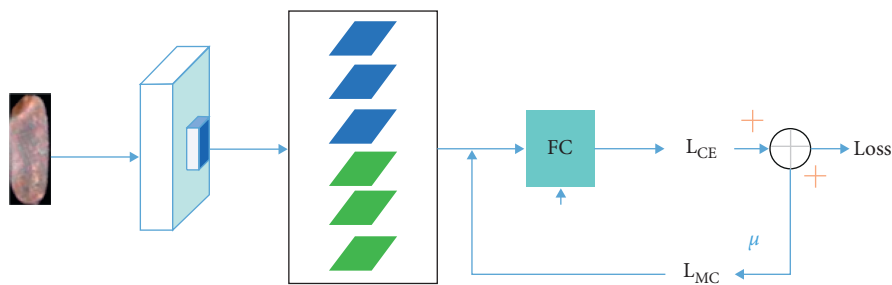

FIGURE 13: Model frame diagram.

complicated, due to the small size of the brown rice germ. Therefore, mlpconv is used to effectively extract the germ characteristics through the subsequent $1 * 1$ convolution. Figure 15 shows the structure.

At the same time, ResNet is added into the second branch [45–47]. These residual connections allow us to have a deeper network structure without increasing parameters. The second branch is multiple MBR connection including mlpconv, batch normalization, and ReLU. The integral structure is as shown in Figures 16 and 17.

The input image is divided into three parts. Through the improved Inception-v3 loss function, the first part becomes the original part (the input image is $299 * 299$). In the second part (the input image is $35 * 35$; after upsampling it becomes $225 * 225$), the BR germs are extracted based on the position direction. As the germ portion is too small in size, the number of pixels is increased by upsampling. Subsequently, the germ image is transferred to 7-layer MBR network. Through the MBR network, the deeper abstract features of the germ can be extracted. It is beneficial to improve the detection accuracy of embryo integrity. The upper part of the network focuses on the extraction of the overall features of brown rice, and the lower part of the network focuses on the feature extraction of the germ part of brown rice. Combining the two parts not only preserves the overall information of the rice, but also accurately detects the integrity of the germ. Finally, the first two parts are combined with the original germ image by global average pooling and input into a complete connection layer. Figure 18 shows the network structure.
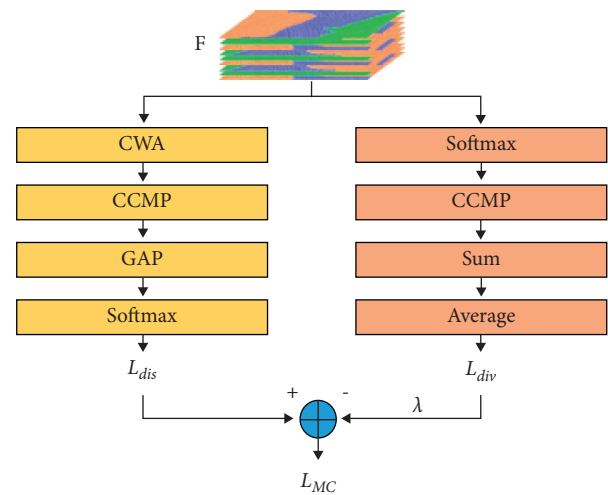

FIGURE 14: Overview of the MC-loss.

During the production and processing of BR, all the processing procedures are completed by experienced workers through visual assessment, which ensures processing quality and facilitates the adjustment of the equipment. Human visual assessment is based on a general feeling and cannot accurately determine the germ integrity. To more accurately describe the germ integrity of BR and improve its quality, BR grains are classified by observation into eight types, ranging from germ-free BR to whole-germ BR. The classification of each type of BR is based on the germ integrity. Figure 19 shows the results.
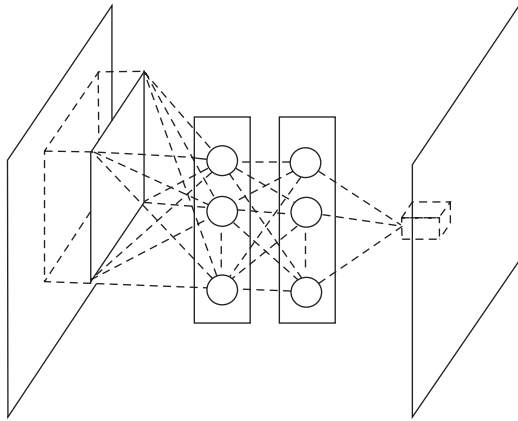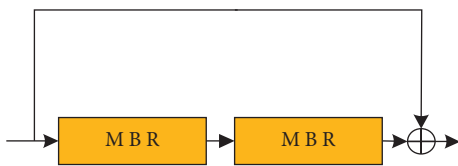
FIGURE 15: mlpconv layer.

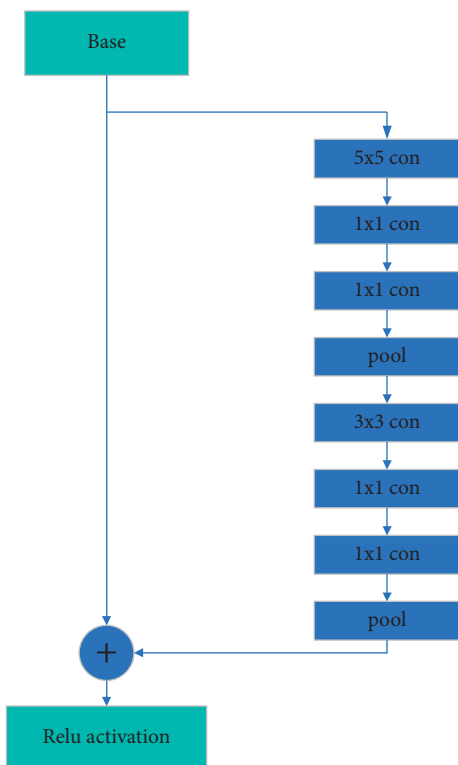FIGURE 16: The scheme of a possible residual connection in a CNN.

FIGURE 17: The overall structure of MBR.

3. Experimental Method

3.1. Segmentation Experiment. Connected BR grains were segmented using the improved background skeleton-based segmentation algorithm proposed in this study. Figure 20 shows the segmentation results.

In view of the limitations of the background skeleton-based segmentation algorithm for single rice grains, the original endpoint matching method was improved in this study. A segmentation experiment on multiple connected BR grains was performed by integrating the region growing algorithm. In addition, the image of the most complex BR grain connection condition was used in a comparison experiment. Figures 21 and 22 show the experimental results for the segmentation of randomly distributed BR grains.

Images are of various types of cereals in which rice grains are connected on a small scale. These images were processed using the classic range-transformation watershed algorithm, the edge detection algorithm with different operators, the region growing algorithm, a fully convolutional NN (FCNN), and the proposed algorithm. The ultimate segmentation results obtained using these four algorithms are shown in Figures 23–27.

During the extraction of single BR grains, operations, such as image rotation, were performed. As a result, the BR grains were horizontal throughout the process. Figure 28 shows the results for single rice grains.

To examine the segmentation performance for various BR grains, 2000 grain samples were collected for three types of BR grains differing in LWR. The connected BR grains were segmented (25 images were collected for each 2000 grains). For the range-transformation watershed algorithm, the edge detection algorithm with different operators, the region growing algorithm, the FCNN, and the proposed algorithm, the number of correctly segmented BR grains and the segmentation time were determined. In addition, the accuracy of each algorithm was calculated. Correct segmentation should ensure that single BR grains are completely segmented and that there are no oversegmentation lines in the BR grains. Tables 1–6 summarize the results.

Tables 1–6 summarize the results. The average segmentation accuracies of the classic range-transformation watershed algorithm, the edge detection algorithm, the region growing algorithm, the FCNN, and the proposed algorithm for the three types of BR grains differing in LWR (long, short, and intermediate) were 72.6%, 83.67%, 90.9%, 91.3%, and 94.96%, respectively. The segmentation accuracy of the proposed algorithm was higher than those of the other segmentation algorithms. With regard to the computational time for the segmentation operation, of the three algorithms, the proposed algorithm required the longest computational time. This is mainly because the proposed algorithm involves a relatively large number of steps in the segmentation operation and is more complex, thus requiring more computational time in processing. Because of this, the proposed algorithm effectively inhibited erroneous segmentation, had a higher segmentation accuracy, and provided more valid images of BR grains for the subsequent extraction of single BR grains.

3.2. Identification Experiment

3.2.1. Germ Direction Identification Experiment. The 6,000 BR grains obtained after image segmentation were used as a training set. A total of 500 training iterations using the VGG
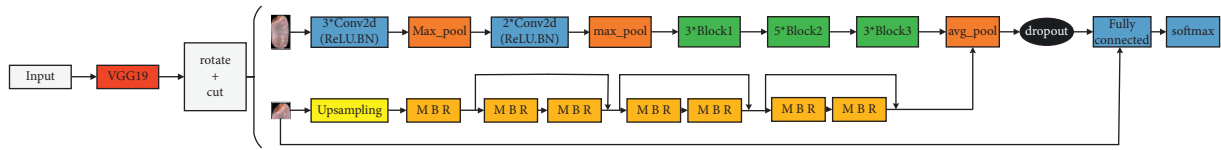
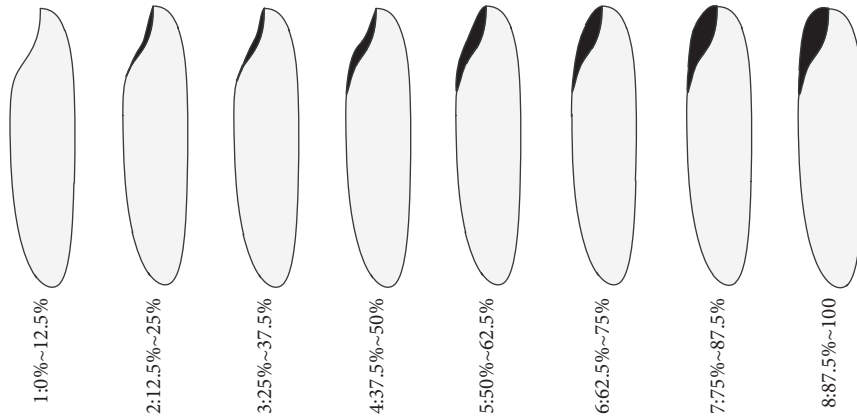FIGURE 18: Improved Inception-v3 network structure diagram.

1:0%~12.5% 2:12.5%~25% 3:25%~37.5% 4:37.5%~50% 5:50%~62.5% 6:62.5%~75% 7:75%~87.5% 8:87.5%~100

FIGURE 19: Rice germ integrity classification standard.
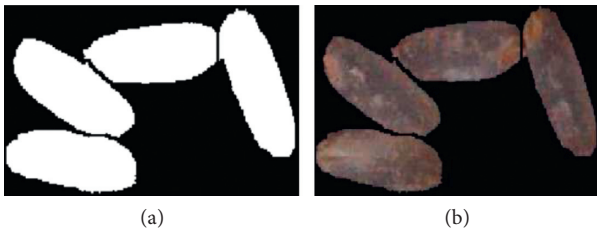

(a) (b)

FIGURE 20: (a) Binary image after segmentation. (b) Germ rice image after segmentation.
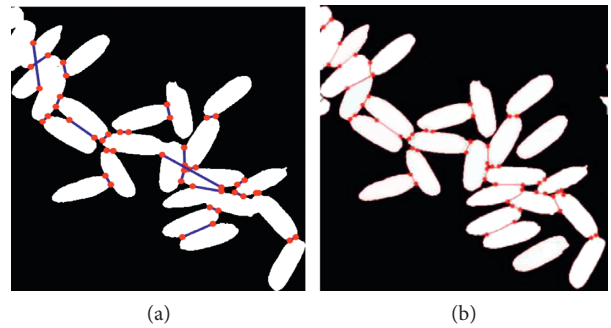

(a) (b)

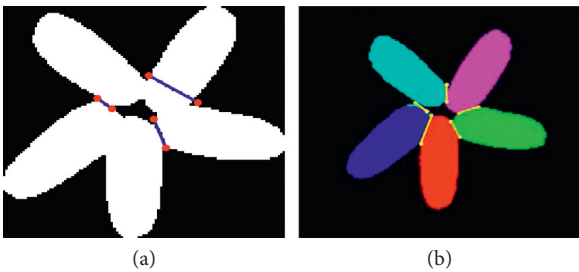FIGURE 22: (a) The unimproved segmentation effect picture. (b) The improved segmentation effect picture.

(a) (b)

FIGURE 21: (a) The unimproved endpoint matching effect picture. (b) The improved endpoint matching effect picture.

FIGURE 23: The classic range-transformation watershed algorithm.

model were performed. In addition, the accuracy of the VGG model was examined by k-fold cross-validation. The accuracy of the VGG model was found to be 98.6%, meeting the accuracy requirements for germ position and direction identification.

3.2.2. Germ Integrity Identification Experiment. First, the position of the germ of each single BR grain was determined using the germ position identification model. Then, the BR grains were rotated so that their germs were in the same direction.

A total of 500 training iterations were performed using the improved Inception-v3 model. Afterwards, training was performed in intervals of 10. The accuracy of the improved Inception-v3 model was examined based on the test set.
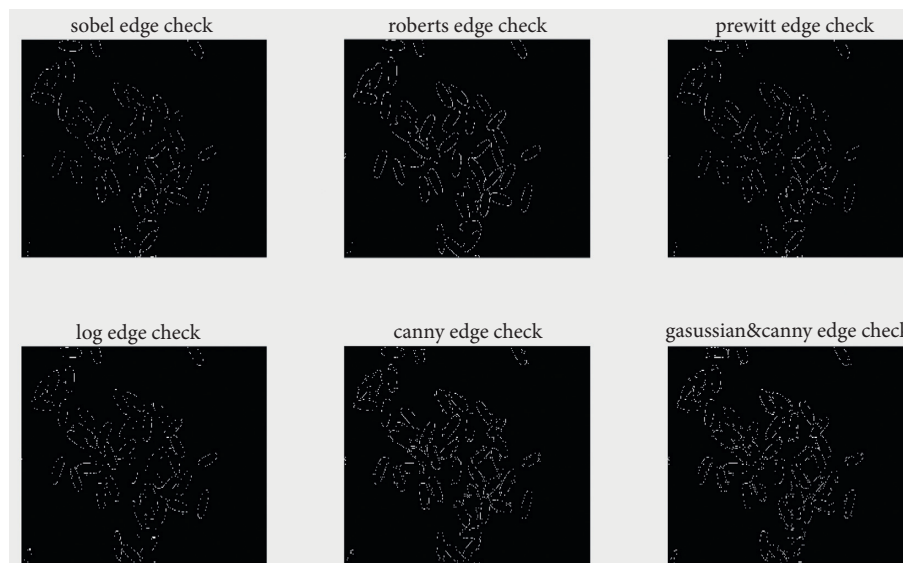
FIGURE 24: The region growing algorithm.

FIGURE 25: The edge detection algorithm with different operators.

FIGURE 26: Fully convolutional neural network segmentation.

FIGURE 27: The proposed algorithm.

Ultimately, the error rate on the test set converged to 5.9%. After 80 training iterations, there was a significant decrease in the rate of convergence. After 500 training iterations, k-fold cross-validation showed that the identification accuracy of the optimized Inception-v3 model was 95.3%.

First, an experiment was conducted to examine the germ integrity of single BR grains. In the experiment, Daohuaxiang-2 BR grains with various levels of germ integrity were used as samples. The whole batch of BR grains was first classified by visual observation according to the eight-type BR classification standard. For each type, 700 grains were selected for testing. A total of 5600 BR grain samples were used. The number of BR grains that could be collected from each image was limited. Therefore, after manual identification, the samples were divided into eight groups. Each group consisted of 25 BR grains. These samples were then subjected to image collection analysis. Image collection was
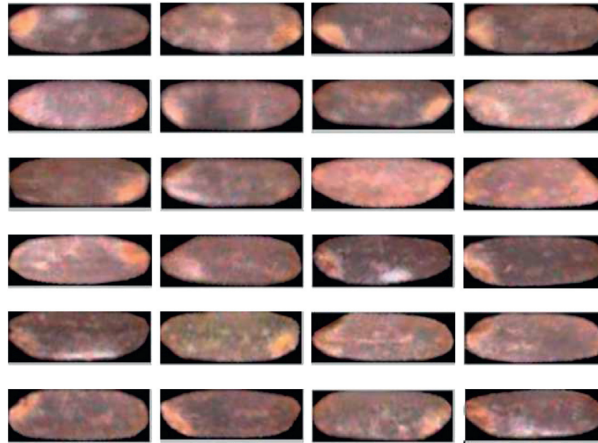
FIGURE 28: Single grain rice extraction image after segmentation.

TABLE 1: The classic range-transformation watershed algorithm segmentation result.

Types of rice	Number of rice grains/grain	Number of correctly segmented grains/grain	Accuracy (%)	Segmentation total time/s
Long	2000	1396	69.8	44.28
Short	2000	1532	76.6	42.21
Intermediate	2000	1428	71.4	45.22

TABLE 2: The region growing algorithm segmentation result.

Types of rice	Number of rice grains/grain	Number of correctly segmented grains/grain	Accuracy (%)	Segmentation total time/s
Long	2000	1578	78.9	68.10
Short	2000	1706	85.3	72.52
Intermediate	2000	1736	86.8	76.63

TABLE 3: The edge detection algorithm segmentation result.

Types of rice	Number of rice grains/grain	Number of correctly segmented grains/grain	Accuracy (%)	Segmentation total time/s
Long	2000	1056	52.8	90.76
Short	2000	1244	62.2	99.21
Intermediate	2000	1276	63.8	92.61

TABLE 4: Fully convolutional neural network segmentation result.

Types of rice	Number of rice grains/grain	Number of correctly segmented grains/grain	Accuracy (%)	Segmentation total time/s
Long	2000	1752	87.6	42.36
Short	2000	1828	91.4	45.69
Intermediate	2000	1874	93.7	47.54

TABLE 5: The proposed algorithm segmentation result.

Types of rice	Number of rice grains/grain	Number of correctly segmented grains/grain	Accuracy (%)	Segmentation total time/s
Long	2000	1842	92.1	76.31
Short	2000	1906	95.3	72.95
Intermediate	2000	1950	97.5	71.44

performed a total of five times. Finally, the manual classification and identification results were compared with the BR test results. Table 6 summarizes the experimental comparison results.

In this section, the optimized Inception-v3 model was used to identify the germ integrity. To demonstrate its advantages, the optimized Inception-v3 model was compared with several currently popular deep network identification models (VGGnet16, VGGnet19, Inception-v3, and ResNet50). The performance of these models was evaluated. In addition, a comparison with the original dataset was performed. Table 7 summarizes the experimental comparison results.

TABLE 6: Test results.

Germ integrity class	Result/grains	The results of the algorithm/grains	Germ integrity accuracy/%	Average accuracy (%)
0% ~ 12.5%	700	667	95.2	
12.5% ~ 25%	700	671	95.8	
25% ~ 37.5%	700	661	94.3	
37.5% ~ 50%	700	666	95.1	
50% ~ 62.5%	700	675	96.3	95.53
62.5% ~ 75%	700	668	95.3	
75% ~ 87.5%	700	677	96.6	
87.5% ~ 100%	700	670	95.6	

TABLE 7: Contrast test results.

Types of rice	Experiment method	Recall	Precision	Accuracy	Training samples
	VGGnet16	83.69	82.34	82.91	2000
	VGGnet19	90.54	90.32	89.98	2000
Long	Inception-v3	92.21	92.55	91.88	2000
	ResNet50	92.11	92.34	91.86	2000
	Optimized Inception-v3	93.73	94.16	94.83	2000
	VGGnet16	83.56	83.24	83.18	2000
	VGGnet19	91.21	91.48	91.33	2000
Short	Inception-v3	92.79	93.12	91.93	2000
	ResNet50	89.15	90.44	89.87	2000
	Optimized Inception-v3	95.56	94.97	95.53	2000
	VGGnet16	83.86	83.16	83.54	2000
	VGGnet19	86.85	87.66	86.82	2000
Intermediate	Inception-v3	92.77	91.68	91.72	2000
	ResNet50	91.36	92.63	92.49	2000
	Optimized Inception-v3	94.87	93.69	94.56	2000

TABLE 8: Contrast test results.

Types of rice	Network	Recall	Precision	Accuracy	Training samples
	Faster R-CNN	77.59	91.23	92.16	2000
Long	YOLO v3	88.18	91.79	91.56	2000
	SSD (VGG-16)	90.16	92.34	92.09	2000
	BSF network	92.10	94.16	94.69	2000
	Faster R-CNN	79.38	91.52	91.36	2000
Short	YOLO v3	89.49	91.92	91.55	2000
	SSD (VGG-16)	90.26	92.71	92.69	2000
	BSF network	95.30	94.16	94.25	2000
	Faster R-CNN	79.81	92.22	92.39	2000
Intermediate	YOLO v3	89.57	92.66	92.91	2000
	SSD (VGG-16)	90.86	92.89	93.21	2000
	BSF network	97.50	94.87	95.53	2000

A comparison of the models in Table 7 shows that the identification accuracy (94.97%) of the optimized Inception-v3 model was higher than those of the other models. As the initialization parameters for training the BR identification model, this model learned domain knowledge from large amounts of easily accessible same-domain data. This addressed the problem of insufficient training data for the BR identification task and further improved the identification accuracy.

Finally, an overall and comprehensive image segmentation and identification experiment was performed to examine the germ integrity. The segmentation algorithm was used to determine the position of each BR grain (first segment according to the background skeleton, and then recognize through the improved Inception-v3 network (BSF)). The BSF network was compared with faster R-CNN, SSD, and YOLO-v3 in terms of accuracy. Table 8 summarizes the experimental comparison results.

A comparison of the methods in the above table shows that the identification accuracy (94.83%) of the BSF network was higher than those of the other algorithms. As a result, the BSF network performed relatively well in image segmentation. The recall rate of the BSF network was far higher than those of the other detection algorithms.

4. Conclusions

In this study, the background skeleton-based algorithm for image segmentation is improved. Besides, segmented endpoints are obtained by detecting skeleton features of the background images of connected BR grains. Meanwhile, adopting segmentation endpoint matching algorithm is beneficial for achieving segmentation successfully. This paper addresses image problems of connected BR grains. And the algorithm accuracy applied to image segmentation of connected BR grains in various shapes is above 96% as presented in this paper. In addition, this algorithm is suitable for segmenting images with large regions of connected BR grains and is highly robust. Moreover, compared to the classic segmentation algorithm, this algorithm is suitable for relatively wide BR grains.

In addition, two networks are proposed to identify germ integrity. A position network model is used to unify the directions of BR germs. Furthermore, an algorithm based on an improved Inception-v3 model is proposed for determining BR germ integrity. The network outputs are divided into eight categories based on their germ integrity. The accuracy of the proposed algorithm is higher than those of other NN algorithms. The comprehensive accuracy of the proposed algorithm is 94.83%. This algorithm can effectively suppress interference from the endosperm surface of BR.

Finally, an overall detection function is achieved based on background skeleton features and the deep learning network algorithm. First, identifying a large number of connected BR grains can ensure high accuracy performance. Second, this can also help rapidly and accurately determine the positions of BR grains. In the experiment, this method was found effective in identifying the germ integrity of a large number of BR grains.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J. Ma and Y. Geng, "Nutritional characteristics and nutritional support of germ rice," *Food and Nutrition in China*, vol. 24, no. 1, pp. 55–58, 2018.

[2] Y. Zhang, D. Han, and P. Liu, "A review of the deductive process and related research results of germ rice in Japan," *Cereal Arid Food Iodustry*, vol. 3, pp. 10–13, 2019.

[3] T. Xie, K. Cheng, and G. Shi, "The new national standard of rice is released to highlight moderate processing and industrial upgrading," *China Food*, vol. 21, pp. 66-67, 2018.

[4] F. Courtois, M. Faessel, and C. Bonazzi, "Assessing breakage and cracks of parboiled rice kernels by image analysis techniques," *Food Control*, vol. 21, no. 4, pp. 567–572, 2010.

[5] P. Lin, Y. M. Chen, Y. He, and G. W. Hu, "A novel matching algorithm for splitting touching rice kernels based on contour curvature analysis," *Computers and Electronics in Agriculture*, vol. 109, pp. 124–133, 2014.

[6] C. Sun, T. Liu, C. Ji et al., "Evaluation and analysis the chalkiness of connected rice kernels based on image processing technology and support vector machine," *Journal of Cereal Science*, vol. 60, no. 2, pp. 426–432, 2014.

[7] S. H. Payman, A. Bakhshipour, and H. Zareiforoush, "Development of an expert vision-based system for inspecting rice quality indices," *Quality Assurance and Safety of Crops & Foods*, vol. 10, no. 1, pp. 103–114, 2018.

[8] Y. Wang, Y. Dai, and J. Xue, "Research of segmentation method on color image of Lingwu long jujubes based on the maximum entropy," *Eurasip Journal on Image and Video Processing*, vol. 2017, no. 1, 2017.

[9] J. Y. Jia, E. Zhang, and C. Wang, "Color image segmentation via wavelet frames," in *Proceedings of the 2019 IEEE 4th International Conference on Signal and Image Processing*, pp. 975–979, Wuxi, China, July 2019.

[10] J. Ren and X. Zeng, "An image segmentation method of improved sobel operator," *Tech Bull*, vol. 55, no. 1, pp. 193–199, 2017.

[11] G. N. Chaple, R. D. Daruwala, and M. S. Gofane, "Comparisions of Robert, Prewitt, Sobel operator based edge detection methods for real time uses on FPGA," in *Proceedings of the 2015 International Conference on Technologies for Sustainable Development (ICTSD)*, pp. 1–4, Mumbai, India, February 2015.

[12] G. Zhang, L. Xu, and M. Zhang, "An improved canny edge detection algorithm," *Chinese Journal of Medical Physics*, vol. 34, no. 8, pp. 811–815, 2017.

[13] Y. Zhang and B. Jiao, "Two-dimensional maximum entropy threshold segmentation algorithm based on Canny operator," *Shanghai Dianji University*, vol. 21, pp. 1–5, 2018.

[14] T. Zhiguo, O. Jianping, Z. Jun, and H. Jie, "A laminar denoising algorithm for depth image," *Acta Optica Sinica*, vol. 37, no. 5, pp. 0510002–0511434, 2017.

[15] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440, Boston, MA, USA, 2015.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing System*, pp. 1097–1105, Red Hook, NY, USA, 2012.

[17] E. A. Smirnov, D. M. Timoshenko, and S. N. Andrianov, "Comparison of regularization methods for imagenet classification with deep convolutional neural networks," *AASRI Procedia*, vol. 6, pp. 89–94, 2013.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the Advances in Neural Information Processing*

Systems, pp. 1097–1105, Lake Tahoe, NV, USA, December 2012.

[20] K. Simonyan and A. V. Zisserman, "Ery deep convolutional networks for large-scale image recognition," in *Proceedings of the ICLR 2015*, pp. 1–13, San Diego, CA, USA, May 2015.

[21] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Computing and Applications*, vol. 32, no. 10, pp. 6111–6124, 2019.

[22] W. Liu, D. Anguelov, D. Erhan et al., "SSD: single shot multibox detector," *Computer Vision-ECCV 2016*, Springer, Cham, Switzerland, pp. 21–37, 2016.

[23] G. Cao, X. Xie, W. Yang, Q. Liao, G. Shi, and J. Wu, "Feature-fused SSD: fast detection for small objects," in *Proceedings of the 9th International Conference Graphical Image Process*, p. 106151E, Qingdao, China, 2018.

[24] J. Jeong, H. Park, and N. Kwak, "Enhancement of SSD by concatenating feature maps for object detection," 2017, http://arxiv.org/abs/1705.09587.

[25] D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scaleconvolutional architecture," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2650–2658, Washington, DC, USA, December 2015.

[26] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery," *Remote Sensing*, vol. 7, no. 11, pp. 14680–14707, 2015.

[27] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the 2016 IEEE Conference Computer Vision Pattern Recognition (CVPR)*, pp. 779–788, Las Vegas, NV, USA, 2016.

[28] J. Redmon and A. Farhadi, "YOLOv3: an incremental improvement," 2018, http://arxiv.org/abs/1804.02767.

[29] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 91–99, Montreal, Qubec, Canada, 2015.

[30] R. Girshick, "Fast R-CNN," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1440–1448, Washington, DC, USA, 2015.

[31] S. Yang, B. Fang, W. Tang, X. Wu, J. Qian, and W. Yang, "Faster R-CNN based microscopic cell detection," in *Proceedings of the 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, pp. 345–350, Shenzhen, China, 2017.

[32] B. Li, S. Li, and J. Li, "Image segmentation algorithm based on the background skeleton feature," in *Proceedings of the 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS)*, pp. 138–144, Chongqing, China, 2019.

[33] G. Li, Z. Wang, and J. Ren, "Relation and application method of deep learning sea target detection and segmentation algorithm," *Mathematical Problems in Engineering*, vol. 2020, Article ID 1847517, 12 pages, 2020.

[34] N. Krishnamoorthy, L. V. Narasimha Prasad, C. S. Pavan Kumar, B. Subedi, H. B. Abraha, and V. E. Sathish Kumar, "Rice leaf diseases prediction using deep neural networks with transfer learning," *Environmental Research*, vol. 198, Article ID 111275, 2021.

[35] L. Yang, R. Huang, J. Huang et al., "Semantic segmentation based on temporal features: learning of temporal-spatial information from time-series SAR images for paddy rice mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.

[36] Y. Lan, K. Huang, C. Yang et al., "Real-time identification of rice weeds by UAV low-altitude remote sensing based on improved semantic segmentation model," *Remote Sensing*, vol. 13, no. 21, p. 4370, 2021.

[37] P. Wang, H. Wang, X. Li, L. Zhang, R. Di, and Z. Lv, "Small target detection algorithm based on transfer learning and deep separable network," *Journal of Sensors*, vol. 2021, Article ID 9006288, 10 pages, 2021.

[38] Y. Guo, S. Li, Z. Zhang et al., "Automatic and accurate calculation of rice seed setting rate based on image segmentation and deep learning," *Frontiers in Plant Science*, vol. 12, Article ID 770916, 2021.

[39] Y. Wang and M. Qiu, "Multimedia detection and processing of remote-sensing image of small target combined with variable neighborhood search algorithm," *Advances in Multimedia*, vol. 2021, Article ID 9672749, 8 pages, 2021.

[40] D. Chang, Y. Ding, J. Xie et al., "The devil is in the channels: mutual-channel loss for fine-grained image classification," *IEEE Transactions on Image Processing*, vol. 29, pp. 4683–4695, 2020.

[41] Y. Wang and Q. Ye, "Improved strategies of ant colony algorithm for solving shortest path problem," *Computer Engineering and Applications*, vol. 48, no. 13, pp. 35–38, 2012.

[42] G. Qin and J. Yang, "An improved ant colony algorithm based on adaptively adjusting pheromone," *Information & Control*, vol. 31, no. 3, pp. 198-199, 2002.

[43] S. Zhu, F. Xu, and Z. Teng, "Application of improvement ants algorithm in solving shortest path," *Computer Technology and Development*, vol. 21, no. 7, pp. 202–205, 2011.

[44] M. Zhou, K. Cheng, and Z. Wang, "Improved ant colony algorithm with planning of dynamic path," *Computer Science*, vol. 40, no. 1, pp. 314–316, 2012.

[45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR' 2016)*, pp. 770–778, Las Vegas, NV, USA, 2016.

[46] M. Boroumand, M. Chen, and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1181–1193, 2018.

[47] A. V. eit, M. Wilber, and S. Belongie, "Residual networks behave like ensembles of relatively shallow networks," in *Proceedings of the 2016 Advances in Neural Information Processing Systems, (NIPS'2016)*, pp. 550–558, Barcelona, Spain, 2016.

[48] M. F. Haque, H.-Y. Lim, and D.-S. Kang, "Object detection based on VGG with resnet network," in *Proceedings of the 18th International Conference on Electronics, Information, and Communication*, pp. 1–3, Auckland, New Zealand, 2019.