Research Article

The Research of Disease Spots Extraction Based on Evolutionary Algorithm

Kangshun Li,[1] Lu Xiong,[1] Dongbo Zhang,[2] Zhengping Liang,[3] and Yu Xue[4]

[1]*College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China*
[2]*School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, China*
[3]*College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China*
[4]*School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China*

Correspondence should be addressed to Lu Xiong; luzi0822@163.com

According to the characteristics of maize disease spot performance in the image, this paper designs two-histogram segmentation method based on evolutionary algorithm, which combined with the analysis of image of maize diseases and insect pests, with full consideration of color and texture characteristic of the lesion of pests and diseases, the chroma and gray image, composed of two tuples to build a two-dimensional histogram, solves the problem of one-dimensional histograms that cannot be clearly divided into target and background bimodal distribution and improved the traditional two-dimensional histogram application in pest damage lesion extraction. The chromosome coding suitable for the characteristics of lesion image is designed based on second segmentation of the genetic algorithm Otsu. Determining initial population with analysis results of lesion image, parallel selection, optimal preservation strategy, and adaptive mutation operator are used to improve the search efficiency. Finally, by setting the fluctuation threshold, we continue to search for the best threshold in the range of fluctuations for implementation of global search and local search.

1. Introduction

In recent years, as the country pays more attention to the problem of "agriculture, countryside, and farmers," "precision agriculture" concept is gradually on the rise. The new ideal of using information technology and computer technology to assist farmers to carry out agricultural production has been put on the agenda. In order to solve the problem of fuzziness and subjectivity in the detection and identification of crop diseases and insect pests, graphics and image processing technology are introduced into the field of agricultural pests identification and control. Image classification and recognition technology of crop pests and diseases is the application of image processing technology in the field of crop diseases and insect pests identification [1]. The image classification and recognition technology will be able to rely on computer image to obtain a large amount of information, as the basis for the diagnosis of diseases and insect pests can make up for the shortcomings of traditional

diagnostic techniques, which has the characteristics of less input parameters, fast diagnosis, high accuracy, and good real-time performance, so it has important significance for the timely provision of the necessary information and the prevention and treatment of diseases and insect pests of agricultural workers.

Research on the diagnosis and image recognition of crop diseases and insect pests on machine vision in foreign countries should be carried out earlier [2–4]. As early as 1985, related researchers identified disease spot by the shape characteristic of grain. In recent years, with the development of computing power, especially the development of pattern recognition and image processing, computer vision technology has been developed. Then, computer vision technology has been applied to the identification and detection of pests and diseases of crops. Keagy, Zayas, and Christopher et al. studied the recognition technology of computer vision and applied it to the recognition of stored grain pests. The experimental results show that the research has high recognition

efficiency. Gassoumi studied the identification of insects in cotton field and applied the computer image processing technology based on feature extraction to identify the pests. Gassoumi also proposed a fuzzy neural network recognition algorithm; the algorithm can identify 12 kinds of pests in the research content, and the recognition rate is above 90%. Murakami et al. sampled the leaves of cucumber and used a variety of methods such as gray level cooccurrence matrix to identify the pests. Where the thrips pest identified, recognition rate is as high as 90%. Shariff proposes a classification algorithm based on fuzzy logic, which is aimed at the 6 kinds of pests in rice. Experimental results show that the algorithm has good classification and recognition rate.

Relatively speaking, the research on the field of classification and identification of crop diseases and insect pests in China is more advanced than foreign countries. In spite of this, some universities and research institutes in China have accumulated a certain research foundation in this field. This is mainly due to the release of China's "food industry science and technology development plan"; the plan clearly shows that the increase in food production needs to have information technology support. At present, China's agricultural information construction is relatively backward, so there is a need to strengthen the agricultural information technology research and gradually the world's leading edge of Agricultural Technology. However, there are many research directions in the agricultural information construction, in which the intelligent identification and detection of diseases and insect pests is one of the key problems to be solved. Therefore, it has a long-term strategic significance to solve practical problems. In our country, many researchers have studied and solved this problem by using computer vision, and some practical problems have been solved [5, 6]. Among them, Qiu et al. and others used artificial neural network technology [7]. They developed a real-time detection system for 9 kinds of pests in the field of crops. With the rapid development of computer vision technology, the problem of population density calculation of micro insect pests has been solved. The laboratory of IPMIST of the China Agricultural University calculated the population density of the whitefly in the crops based on their research results. The experimental results show that the recognition accuracy can reach more than 90%.

Chen et al. used digital image processing technology and local threshold method to segment the target region and nontarget region, and the research results were applied to the identification of the cotton leaf worm hole area and the edge of the leaves of defect degree and pest level. The experimental results show that the algorithm is effective, and the error range is less than 0.05 [6]. Zhang et al. use neural network to determine the degree of breeding of tomato. The training of neural network is realized by genetic algorithm. Xu et al. are also studying tomato species, but their study is to identify the missing nutrients in the culture process. The feature extraction improved by using the percentage difference histogram after eliminating the influence of leaf shape and leaf environment. Therefore, the precision of extracting the texture characteristics of the missing nutrients is improved, and the recognition rate is as high as 80%. Wang also used the digital image processing and neural network

[8], but he combined expert knowledge and information technology, so as to play a long distance identification and detection of crop diseases and pests. Cheng et al. set the color and texture of the leaves as a starting point. It is done through the use of color theory, based on these two features to select the appropriate parameters, and the use of fuzzy K-nearest neighbor classification. The image of rice pests and diseases was collected by Wang and Zhou and others [9], who take borer as object based on neural network to complete the identification of the experimental results, with accuracy of up to 90%.

Liang et al. collected different kinds of insect images [10] to obtain the characteristics of the image. They showed the characteristics obtained based on neural network in mathematical form to join the rough set. The experimental results show that the classification accuracy is high. In 2009, Tian et al., Niu et al. and others propose a method for the identification of plant diseases and insect pests based on the SVM. The method is based on the linear kernel function and radial basis function in the SVM based on the identification method for cucumber disease and insect pests. Experimental results show that the proposed algorithm has better recognition accuracy than other support vector machines in the field of cucumber and show that the algorithm has better classification accuracy for small sample data sets. Based on the experiment results, the method proposed in this study provides a classification accuracy of up to 94.12% in training set and 81.82% in testing set with the SVM classifier [11, 11–14]. The extracted features are input into some classifiers including the SVM (support vector machine) [11, 13, 15]. Classification is a classical data mining method, which have been applied in many fields [1, 16–21]. In 2011, Liu et al. also used support vector machine to identify plant diseases and insect pests. They took the traces that the insects left on the persimmon surface as research object, and then the texture and color features were extracted and used as parameters. Support vector machine uses Sigmoid as the kernel function for classification and recognition. The experimental results show that the algorithm can better identify fruit diseases and insect pests. Qiu et al. studied the classification and recognition based on fuzzy decision classifier, and the research results will be applied to the recognition of stored grain pests. In the form of insect pests, the complexity of the recognition feature, through the classification of the classifier, and the accuracy are of up to 95.2%. The experimental results show that this algorithm is aimed at the effect of pest classification. In addition, Qiu et al. also realized the monitoring of field pests and diseases; the experimental results show that the monitoring technology has high practicality. Chen et al. study wheat pests. In addition to the study of the classification of pests, the image segmentation technology is applied to the study. They take the aphids as the target; the experimental results show that the algorithm is effective, and the recognition accuracy is up to 90.7%.

2. Evolutionary Algorithm

Evolutionary algorithm is an iterative search algorithm proposed by Holland in 1975. Evolutionary algorithm [22, 23]

is based on Darwin's theory of biological evolution, which is the nature of biological evolution and the nature of the population to be integrated into the algorithm, so that the algorithm is intelligent and parallel. At the same time, the reasonable design evolution algorithm [24] can effectively avoid the local optimization problem in the process of solving the problem, and the algorithm has strong robustness. Evolutionary algorithm essentially is a kind of probability of random search, but with adaptive, it refers to the "survival of the fittest" evolutionary ideas. The algorithm gives a fitness value of each individual in the population, which represents the survival probability of the individual in the process of biological evolution [25]. The higher the fitness of the individual is, the higher the probability of its entry into the next generation of reproduction is. It is this kind of selection strategy that makes the evolutionary algorithm have self-adaptability in the process of solving practical problems and produce better offspring. Evolutionary algorithm for the problem is not directly solving the problem of the solution space for optimal solution search, but the problem of the solution of gene coding and the establishment of the mapping between problem solving and gene coding. Each potential candidate solution has a corresponding gene encoding. In the process of the algorithm, the solution of the code is searched. After determining the coding method for the potential candidate solution, the initial population is generated. Initial population is the starting point of the search algorithm, which is the first generation of the population multiplication and scientific and reasonable setting of initial population, which helps the algorithm to get the optimal solution of the problem faster and more efficiently. In the process of population evolution, evolutionary algorithm is used as the only index to evaluate the quality of the population. The algorithm is analyzed by the model of the problem. The fitness function is designed to solve the problem, and the fitness value is used to evaluate the probability of survival to the next generation in the evolution process of the population and the survival of the fittest. As well as the process of biological evolution, evolutionary algorithms need to experience the three processes of selection, crossover, and mutation. Selection, crossover, and mutation are referred to as genetic operations of evolutionary algorithms. It is because of the genetic operation, which makes the population of individual reproduction of more diversity, we can produce different solutions, to avoid falling into the local optimal dilemma. Evolutionary algorithm, for the whole process as natural biological evolution generally, continues to produce better offspring and algorithm convergence in the last population. The optimal solution of the individual with the highest fitness value is obtained by genetic decoding.

3. Analysis of Diseases and Pest of Crop

Because of the diversity of plant diseases, many difficulties are brought to the diagnosis of machine vision. Field crop pests and diseases have a wide variety of categories and complex and diverse symptoms, which makes the field of crop pests classification and identification more difficult. The research and analysis of crop diseases and insect pests showed that because the leaf diseases and insect pests have the same occurrence, some symptoms are similar; it is difficult to distinguish. Leaf diseases and insect pests are difficult to distinguish because the part of the occurrence is the same, and some symptoms are similar. In this section, maize disease and insect pest image will be used as the research object, study on leaf plant diseases and insect pests, and analysis the lesion area of image visual features and statistical features. Using image analysis and combined machine vision technology and converting some of the information in the image that is difficult to describe to a target that the computer can recognize and deal with have become an important basis for the identification of leaf disease and disease in the classification and identification of diseases and pests.

3.1. Analysis of Image Gray. The gray images of crop diseases and insect pests include many features of the image. Among them, the gray histogram can be generalized to represent the gray statistical information of the image and can get the description information for a particular type of image. As shown in Figure 1, it takes the maize big spot disease image as the analysis target, carries on histogram the original image to 256 gray level, and uses the gray histogram to carry on the analysis. From the visual angle of analysis, leaf disease has obvious characteristics of the external position through the observation. The gray values of disease spot are significantly higher than those of no symptoms around the lesion area. From the point of view of statistical information, it can be found that the gray value of the image of the maize spot disease showed a single peak distribution, and in the 100 areas of gray value changes significantly, as well as the frequency of rapid decline.

The infection process of the leaf of crops is the pathogen invading the local side and gradually spreads outward to form the green spots which differ in size, color, and shape. The lesion and healthy leaf area vary in the color difference, and this difference appears in gray image performance as gray inhomogeneity. As shown in Figure 2, the histogram showed that the gray value of lesion location broadly distributed in the area more than 100, and the gray value of good leaf location broadly distributed below 100. For a particular type of pest disease images, the gray value of lesion and nondisease area is in different gray range. This is consistent with the observed appearance characteristics of the diseased and nondiseased leaves; it is also consistent with the rule of formation of lesion. It is obvious that gray mean value of pest lesion position in the image is higher than that of the disease-free area from Tables 1 and 2. The gray value of variance in the image spot area is large, the distribution is not uniform, and the gray value of no symptoms area convergence and its distribution is uniform. In the histogram, the lesion of the gray value distribution is more average. In the image, the distribution of gray value is scattered, and the distribution of the health area is concentrated.

3.2. Analysis of HSI Chroma. HSI color space is a model of image color which is described by using three kinds of indexes, the chroma value, the saturation value, and the

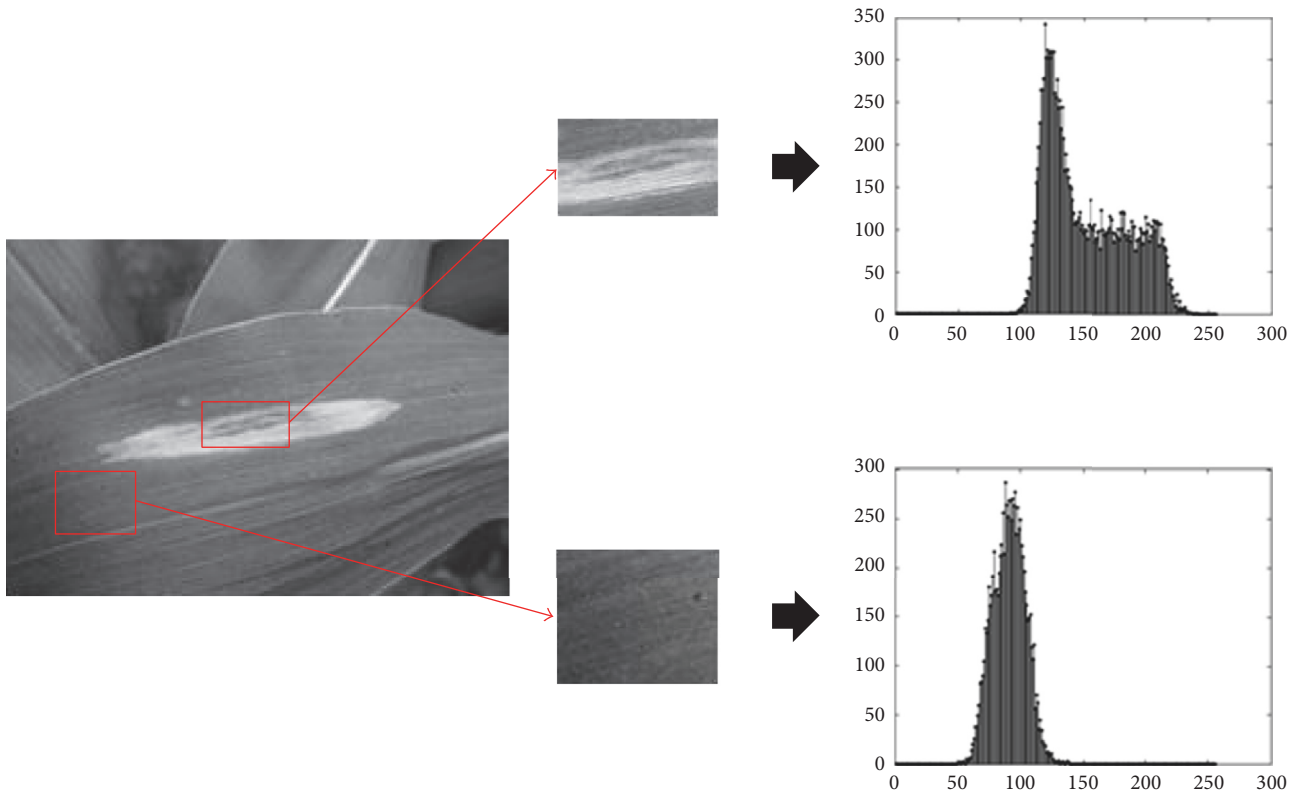FIGURE 1: Analysis of gray big spot disease image.

FIGURE 2: Gray histogram distribution characteristics of large spot disease image.

brightness value [26]. The chroma value is defined as the color wavelength, which reflects the color degree of color image. The infection process of the leaf of crops is the pathogen invading the local side and gradually spreads outward to form the green spots which differ in size, color, and shape. Therefore, the color value of HSI color space can be more clearly reflected in the disease and insect pests image of the color difference. General pest spots will appear yellow, red brown, brown, and other strange colors. Therefore, it can be more effective to reflect the visual characteristics and statistical characteristics of the diseased leaves by analyzing the color value of the image of plant diseases and insect pests in the HSI color space. As shown in Figure 3 the big spot disease image of maize is converted to HSI color space from the RGB color space for analysis. In order to facilitate the observation and

analysis and to obtain the color value H in the HSI color space, calculate the conversion image I which is divided into 256 levels. The position of disease spot in converting image I has obvious external dissimilarity compared to disease-free position, and the bimodal phenomenon can be observed in the statistical histogram.

The position of big spot disease of conversion image I and leaf position without symptoms were analyzed, as shown in Figure 4. Analysis results show that the stationarity of lesion location chroma values of H and disease-free position without obvious difference and chroma values of H are concentrated in a certain interval, and distribution is uniform. In Tables 3 and 4, it can be found that, in HSI color space chroma, value is an important index to evaluate the object surface color, while the leaf color changed significantly after

TABLE 1: Gray information of plant diseases and insect pests image.

Image type	Mean value	Mean square error	Entropy	Energy
No symptoms of image	147.67	13.87	5.82	0.019
Big spot disease	109.30	30.27	6.80	0.012
Cochliobolus heterostrophus	125.11	43.41	7.01	0.010
Gray leaf spot	127.28	32.63	6.85	0.011
Cercospora leaf spot	146.01	26.39	6.64	0.012
Anthrax	147.14	32.11	6.90	0.009

TABLE 2: Gray information of pest disease image.

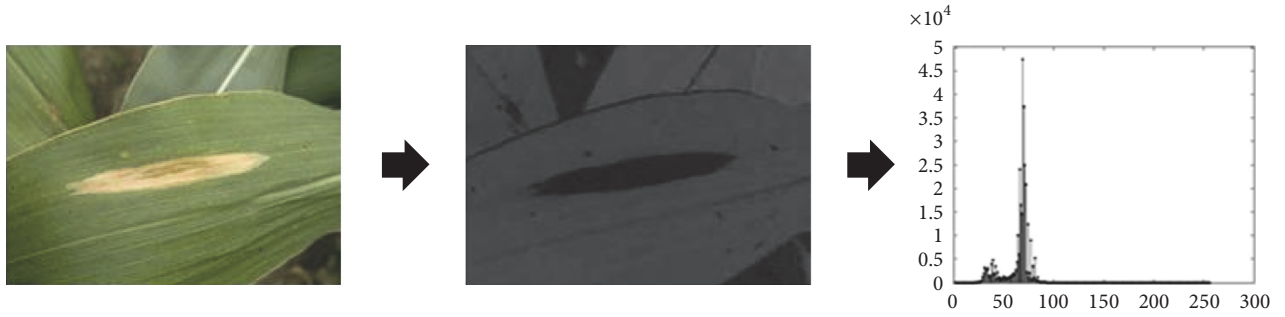| Types of spot	Mean value	Mean square error	Entropy	Energy
Cochliobolus heterostrophus	183.76	35.94	6.83	0.009
Gray leaf spot	170.04	32.03	6.45	0.013
Cercospora leaf spot	140.95	23.34	6.50	0.014
Anthrax	149.67	19.32	6.24	0.014

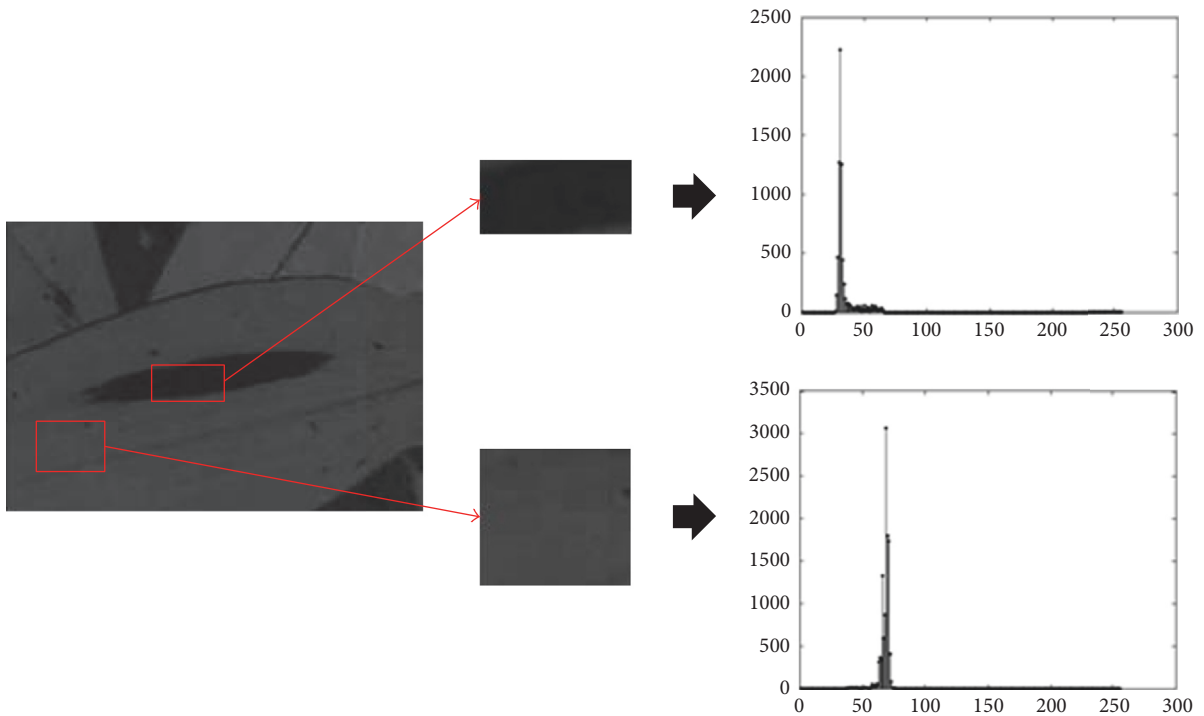FIGURE 3: Analysis of HSI chroma of big spot disease.

FIGURE 4: Distribution characteristics of HSI color in the image of the large spot disease.

TABLE 3: Chroma information of image of plant diseases and insect pests.

Types of spot	Mean value	Mean square error	Entropy	Energy
No symptoms image	67.34	2.76	3.11	0.153
Big spot disease	63.77	11.94	4.58	0.069
Cochliobolus heterostrophus	55.16	10.46	5.01	0.041
Gray leaf spot	75.98	21.85	6.13	0.017
Cercospora leaf spot	58.54	9.22	4.87	0.044
Anthrax	44.74	10.99	5.04	0.038

TABLE 4: Chroma information of image of plant diseases and insect pests.

Types of spot	Mean value	Mean square error	Entropy	Energy
Big spot disease	32.44	6.90	3.32	0.170
Cochliobolus heterostrophus	43.06	8.84	4.88	0.038
Gray leaf spot	41.39	20.15	5.45	0.034
Cercospora leaf spot	55.91	9.18	4.94	0.040
Anthrax	37.68	5.11	4.33	0.056

the occurrence of the disease. And this change can reflect the use of chroma, as an important basis for judging the location of spot disease.

4. The Segmentation Scheme of Spot Disease Based on Genetic Algorithm

4.1. The Main Steps of Segmentation of Spot Disease Extraction Based on Genetic Algorithm. Digital image processing, image matching, image classification, and recognition technology, are the bases of the work on which the image segmentation is generally based on [27]. Image segmentation can be used to extract meaningful features in the image, which can be used to obtain the required image information. The object of image segmentation is to decompose the modules with the same or similar features in the image and get some basic elements with different characteristics. The basic element which is obtained by image segmentation is called image element. The pest images can be more easily and quickly processed after extracting the image element from pest spot image. Therefore, the realization of image classification and recognition play a fundamental role, according to the characteristics of crop diseases and insect pests image, to design an effective image segmentation algorithm, extraction plant diseases, and insect pests in the lesion area. The image features of plant diseases and insect pests can be divided into visual features and statistical features [28]. The visual features of the image of pests and diseases refer to the characteristics which can be identified by human visual system, such as image color, leaf lesion brightness, texture, and shape. Statistical characteristics of the image of plant diseases and insect pests are obtained by means of a certain image transformation, such as color moment, statistical histogram, and image spectrum.

Threshold segmentation based on image gray value is a very effective method, which can reduce the computational cost and can be used in special real-time applications. Therefore, the threshold method is widely used. In this section, we present a threshold segmentation algorithm based on evolutionary algorithm whose fusion guides filtering algorithm to set the weights matrix to achieve the extraction of diseased spots. The specific process is as shown in Figure 5.

The following are details about the major steps of the processing flow:

(1) Construct two-dimensional histogram, using color and gray information to establish two-dimensional histogram after conversing original pest damage image to gray and color image. Plant diseases and insect pests lesion image design for a target class; the other parts are set as the background which makes the histogram able to reflect more pest image information. From one-dimensional to two-dimensional transformation, it can more clearly reflect the spot information distribution and information clustering.

(2) Get two valued matrix marked spot diseases based on application of genetic algorithm. Threshold search using a genetic algorithm is based on the design of the chroma and gray histogram, and threshold value is determined by the second optimization. Design chromosome coding suitable for the image threshold characteristic of diseases and insect pests combined with the results of the image analysis; the initial population is selected by the empirical value combined with the random value, and the optimization efficiency is improved. Measure function of spot disease segmentation designed to evaluate the merits of individual, according to the measurement method of the target and background in the combination of one-dimensional histogram and two-dimensional histogram. Combined with the characteristics of the
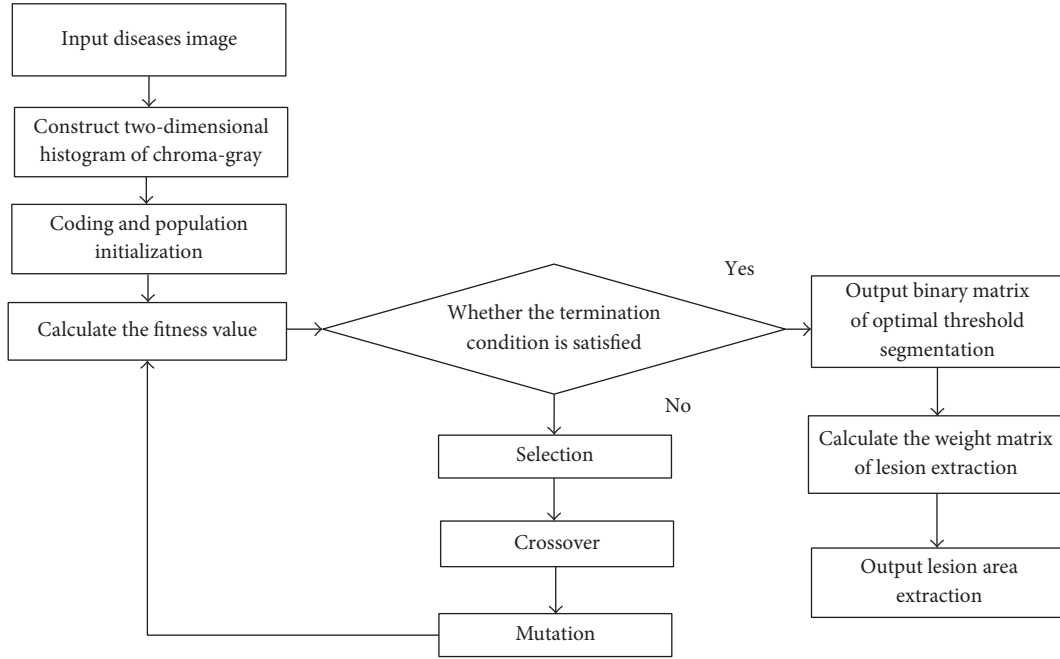
FIGURE 5: Flow chart of algorithm implementation of extraction of disease spot.

image of pests and diseases, the strategy of genetic operation is designed to improve the search efficiency and to speed up the convergence of the algorithm. Finally, use genetic algorithm to derive the optimal threshold to obtain binary matrix of spot disease segmentation, labeled pest damage in the image of the spot disease.

(3) The weight matrix of spot extraction is calculated by using directed graph. Simple binary segmentation results are used to extract the disease region; being easy to lose the original lesion edge, shape, texture, and other characteristics, it is not conducive to the classification and identification of plant diseases and insect pests. Therefore, the idea of guiding filtering algorithm is introduced. The filtering operation on the image segmentation is based on the gray orientation guide. Calculate the weight matrix of lesion region extraction and optimize results of lesion extraction. Restore the segmented lesion edge, texture fuzziness, and rough location; it is better to keep pest lesion image edge, shape, texture, and characteristics.

4.2. The Combination of Genetic Algorithm and Chroma-Gray Histogram. Threshold segmentation is an effective method of image segmentation, and the ideal situation of threshold segmentation is the statistical histogram of the image to be bimodal, and the valley in the histogram is used as the segmentation threshold to separate the object and background [29]. But in most cases, the image histogram is not the bimodal form; it is more of a single peak or multipeak shape; this will make the selection of the threshold of the histogram segmentation difficult. Two-dimensional histogram

segmentation method is developed from the traditional one-dimensional histogram segmentation, which is based on the threshold value of image segmentation. Two-dimensional histograms, i, j, respectively, represent the image of the two indicators. Take the frequency of two tuples (i, j) as the vertical coordinate to establish 3D information coordinate of image. Set the frequency of two tuples $(i, j) f_{ij}$, respectively, to be corresponding to the gray value of the pixel and the right pixel gray value; the joint probability density is defined as follows:

$$p_{ij} = \frac{f_{ij}}{M \times N}, \tag{1}$$

where $M \times N$ is the size of image. In order to form the gray level cooccurrence matrix, which represents the gray spatial variation, two-dimensional histogram is divided into four regions by using two thresholds (t, s), as shown in Figure 6, where areas 0 and 1 represent background and target, respectively, and region 3 and region 2 represent the information of boundary texture and noise.

In the previous section of pest damage image analysis, we can know lesion location in the image gray value characteristics and that no symptom location gray value feature has an obvious gap, but single gray-scale information cannot well determine the location of the lesion in the image; also it is prone to regional misjudgment of the situation. Therefore, this study proposed a segmentation method based on chroma-gray histogram. At the same time, this method uses the image gray level information and takes full account of the chroma information of image to determine the spot area. Images are converted to HSI chroma space and HSI gray space for extracting gray value and chroma value, respectively. The chroma value is divided into K and the chroma
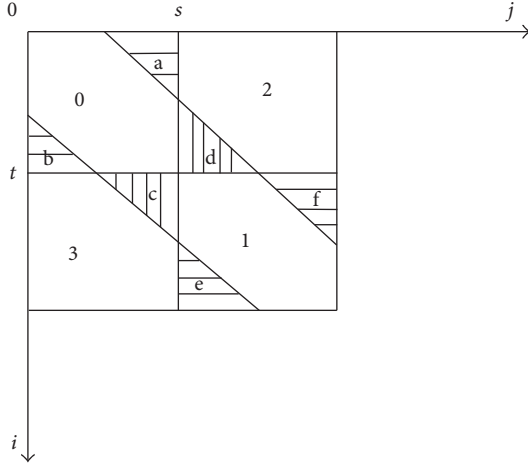
FIGURE 6: Region division of two-dimensional histogram.

of nongreen area, which is divided in $[t, K-1]$, and the gray value is divided into L. Count the frequency of two tuples (i, j) which are consisting of chroma value and gray value. The distribution probability of region 0 and region 1 is represented by P_0 and P_1 after threshold (t, s) segmentation. Formula (2) and formula (3) are used to represent the distribution probability of the two regions.

$$p_0 = \sum_{i=0}^{t} \sum_{j=0}^{s} p_{ij}, \qquad (2)$$

$$P_1 = \sum_{i=t+1}^{K-1} \sum_{j=s+1}^{L-1} P_{ij}. \qquad (3)$$

Area 0 is defined as the background region and region 1 is the target area. Region 2 and region 3 are negligible noise points and scattered points, expressed as follows:

$$P_0 + P_1 \approx 1. \qquad (4)$$

Thus, in a two-dimensional histogram, the distance measure defined by region 0 and region 1 can be expressed as follows:

$$\delta_B (t, s) = P_0 \left(\xi_0 - \xi \right)^2 + P_1 \left(\xi_1 - \xi \right)^2, \qquad (5)$$

where ξ_0, ξ_1 are mean values of element of regions 0 and 1. ξ is population mean value. Figure out the value of (t^*, s^*) as the threshold of image segmentation when δ_B is the largest value.

As shown in Figure 7, taking the gray leaf spot of maize as an example, the disease and insect pests image is transformed into the two-dimensional histogram of the combination of the image of gray value and chroma value, and the image of gray and chroma value is set to 256 levels, as in $[0, 255]$. Joint probability density of two tuples (i, j) is P_{ij}.

If one uses the two-dimensional histogram of the Otsu method as in formula (5) for segmentation, then the amount of computation is exponential growth. At any t and s, P_0, P_1, ξ_0, and ξ_1 need to execute cumulative calculation, perform 6 times' division, and multiply. The time complexity of arbitrary

image processing is $O(N^4)$. In the segmentation of the chroma-gray two-dimensional histogram, in this paper, genetic algorithm is used to find the optimal threshold; combining the goal of one-dimensional and two-dimensional histogram and the background of the measurement method, two-time optimization is performed based on histogram segmentation threshold values t, s. Using genetic algorithm in threshold optimization is faster than the exhaustive search to realize threshold optimization, and the optimal solution of the corresponding image segmentation performance is good and stable.

4.3. Coding Design and Population Initialization

4.3.1. Coding Design. Using the genetic algorithm optimization to segment two-dimensional histogram of crop diseases, the main steps are as follows: in the process of image threshold optimization, according to the characteristics of the diseases image, the genetic algorithm is used to solve the latent solution in the solution space. This key step will also affect the optimization efficiency, convergence speed, and so on. In the analysis of Section 3.2, we can use the statistical analysis of the data obtained by the different diseases image gray and chroma distribution and characteristics to coding. Using binary encoding, the gray and chroma of the image are divided into 256 levels, with 8-binary number coding, that is, from 00000000 to 11111111, corresponding to the diseases image gray and chroma value of the 256 decimal number interval of 0 to 255.

4.3.2. Population Initialization. The initial population is the starting point of the optimal threshold search. According to the gray and chroma features of the pest image, we give the initial population in the gray space and chroma space. In this paper, we adopt the same strategy for gray and chroma image; 5 empirical values with good effect on the experiment and 5 random values randomly generated from 0 to 255 were selected as the initial population.

4.4. The Design of Adaptive Value Function.

The genetic algorithm is used to evaluate the fitness of individuals, and based on this, it can produce better offspring. Fitness function is a function of calculating the fitness value of the population, and it is an important function that directly affects the search efficiency and convergence of the algorithm. Therefore, the fitness function that is designed for the two-dimensional histogram threshold segmentation is very important. Fitness function design, usually achieved by solving the problem of modeling and computing the objective function, in the transformation process formed by some specific practical problems of the fitness function, reflects the individual merits.

Set a given image chroma level to be L, segmentation threshold value of chroma to be t, gray level to be L, and segmentation threshold value of gray to be s. For the first search, suppose $s = L - 1$, that is, $P_1 = 0$. The distribution
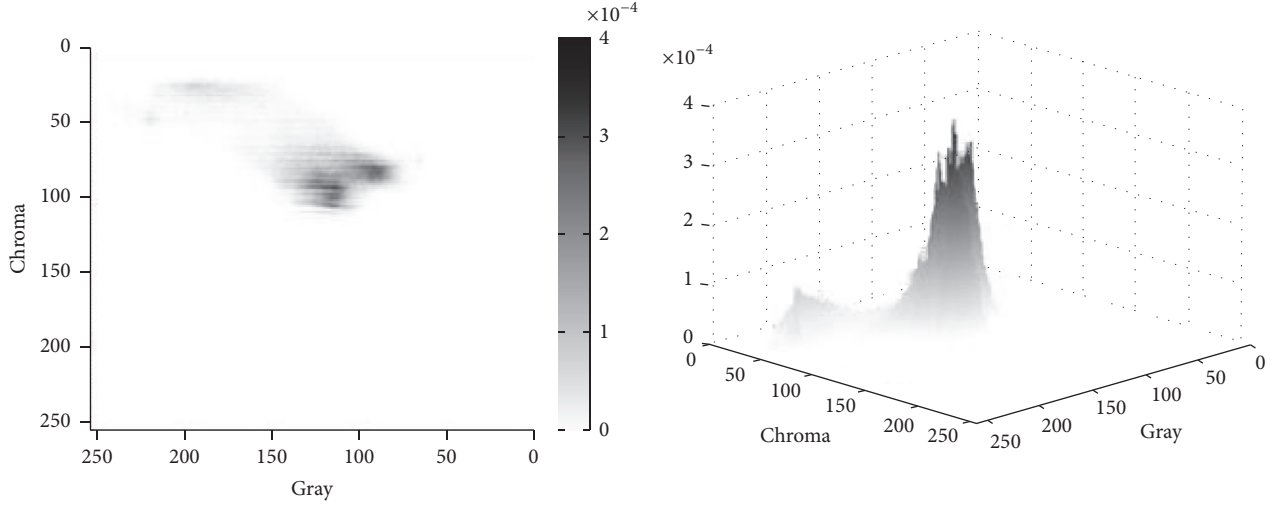
FIGURE 7: Two-dimensional histogram of chroma-gray.

probability of area 3 cannot be ignored, and the distance measure of the first optimization is expressed as follows:

$$\sigma_{B,1}(t) = P_0 (\xi_0 - \xi)^2 + P_3 (\xi_3 - \xi)^2, \tag{6}$$

where P_3 is the distribution probability of area 3 and ξ_3 is the mean of elements in area 3. Seek t^* as the threshold of the first search when $\sigma_{B,1}$ is the maximum.

On the second search for optimization, put $t = t^*$ into formula (5); then formula (7) is

$$\sigma_{B,2}(t) = \sigma_B (t^*, s) = P_0 (\xi_0 - \xi)^2 + P_1 (\xi_1 - \xi)^2. \tag{7}$$

Seeking s^* as the threshold of the second search when $\sigma_{B,2}$ is the maximum, the result of the two optimizations (t^*, s^*) is taken as the final threshold.

Therefore, formula (6) and formula (7) are used as the fitness function of the two-segmentation method that this paper adopted, respectively.

4.5. The Design of Genetic Strategy

4.5.1. Parallel Selection. Selection is the first step of genetic manipulation; the purpose is to select the population of individuals for the crossover, mutation, and other operations. Individuals with better fitness values can produce a new generation of populations by selection. In order to guarantee the multiplicity of the population in the process of reproduction and to keep pace with the convergence speed of the algorithm, this paper adopts the method of combining fitness value selection with optimal preservation strategy.

The specific design is as follows: first, the individual selection probability is given to each individual according to the proportion distribution method of population fitness. After the crossover and mutation operation, the individuals of the parent population and the offspring population are sorted according to the fitness value, respectively. The highest individual adopts a conservation strategy, substituting the optimal individuals in the parent population directly for the non-optimal individuals in the offspring population to form new

offspring populations. The ensemble selection method can guarantee the diversity of population by adapting the fitness value proportion, and, at the same time, it can save the optimal individuals accelerating the convergence speed and improve the solution precision.

In the algorithm, the evolutionary computation process is divided into three stages: initial, middle, and final stage, and different optimal preservation strategies are adopted for different periods. In the early stage of evolutionary computation, the individuals with the best fitness (the largest fitness) of the previous generation are substituted for the individuals with the fifth-highest fitness value in the generation, so as to avoid the local convergence of the search results and to enlarge the optimization interval of the division threshold. In the middle period of evolutionary computation, the best individuals of the previous generation are substituted for the individuals whose fitness is ranked last, so that the optimal individuals are preserved in the next generation, and the convergence of the algorithm is guaranteed. At the end of the evolutionary computation, the individuals with the first two ranks of the previous fitness values are substituted for the individuals with the second-order of fitness in the population, so that the algorithm converges quickly in the optimal solution neighborhood and improves the efficiency of optimization.

4.5.2. Single Point Crossover. Crossover [30] is the process of generating new individuals. Through the structural exchange of the individuals, the new individuals are different from the fathers, which ensure the diversity of the population in the process of evolution. Through the intersection operation, the searching space of the genetic algorithm is expanded, and the algorithm has powerful searching ability. In this paper, a single point crossover method is adopted, in which the paired individuals selected for crossover operation are assigned a cross bit in the binary coded bit string, and the crossover bits are the structural demarcation points for the exchange and reorganization of the pairing individuals. The latter part of the encoded bit string structure is to exchange the formation of a new population of individuals. Generate random number by

random function to determine the intersection point and set the crossover probability to be 0.7.

4.5.3. Adaptive Mutation.

Mutation operator [31] is also an important part of the genetic algorithm. It can effectively increase the local search ability of the algorithm by mutation operator to produce an individual different from the parental property; at the same time, it can increase the diversity of individuals and avoid premature phenomenon. The selection of mutation probability is particularly important; the value of pm is too large and is easy to lose outstanding individual; otherwise it is difficult to jump out of local optimum. In this paper, using adaptive mutation operator, pm will change with the individual fitness value; the larger fitness of individuals is with smaller mutation probability, while the smaller fitness of individuals is with a large probability of mutation.

The calculation formula of mutation probability is as follows:

$$
p_m = \begin{cases} \dfrac{k_1 \left(f_{\max} - f \right)}{f_{\max} - f_{\mathrm{avg}}} & f \geq f_{\mathrm{avg}} \\ k_2 & f < f_{\mathrm{avg}}, \end{cases} \tag{8}
$$

where f is the fitness value of the individual to be mutated, $f_{\max}$ is the largest fitness value in the population, f_{avg} is the average fitness value in the population, and k_1, k_2 are the coefficients, representing the change of p_m, $k_1 = 0.2$, and $k_2 = 0.2$.

4.6. Optimization of Extracting Effect Based on Gray-Oriented Graph.

In order to make the binary image segmentation have a more delicate edge texture, this paper proposed a new method for calculating the weight of lesion extraction by using the idea of oriented filtering [32, 33] that performs the filtering operation on the binarization matrix obtained by image segmentation and calculates the weight matrix by transforming to the lesion region. Guided filtering is a smoothing operation that preserves the edges of an image under the guidance of a graph. Definition p is the input image to be processed, I is the guided graph, and q is the output image after filtering operation. Guided filtering is assumed to be in a window w_k centered on pixel k, where q is a linear transformation of the guided graph I and is expressed by (9) as follows:

$$
q_i = a_k I_i + b_k, \quad \forall i \in w_k, \tag{9}
$$

where coefficients a_k and b_k are constants in window w_k.

In this paper, we define a weight matrix H in the process of lesion area extraction and let D be the original diseases image and J is the lesion extraction image; then the region extraction process is expressed as follows:

$$
J_i = H_i D_i, \quad H_i \in [0, 1]. \tag{10}
$$

Gray image has a fine edge texture; therefore, the gray image is taken as the guide graph I, introducing the idea of directed filtering; the binary matrix of the image segmentation result is defined as A, in a window w_k centered at the

pixel k. Convert H to a linear transformation of the graph I, expressed as follows:

$$
H_i = a_k I_i + b_k, \quad \forall i \in w_k. \tag{11}
$$

The guidance filtering requires coefficients a_k and b_k to minimize the difference between the output image q and the input image p. Similarly, coefficients a_k and b_k are required to minimize the difference between the weight matrix H and the binarization matrix A. In the window w_k, we consider the cost function as follows:

$$
E\left(a_k, b_k\right) = \sum_{i \in w_k} \left(\left(a_k I_i + b_k - A_i\right)^2 + \varepsilon a_k{}^2 \right), \tag{12}
$$

where ε is a regularization matrix to prevent a_k from being too large and keeping the filtering computationally stable. According to the linear regression analysis, the optimal solution expression of a_k, b_k can be expressed as follows:

$$
a_k = \frac{(1/w) \sum_{i \in w_k} I_i A_i - \xi_k \overline{A_k}}{\delta_k^2 + \varepsilon}, \tag{13}
$$

$$
b_k = \overline{A_k} - \xi_k a_k, \tag{14}
$$

where δ_k^2 and ξ are the variance and mean of gray-guided graph I in window w_k and $|w|$ is the number of elements in window w_k. At the same time, $\overline{A_k}$ is the mean of all the elements in window w_k. Since element i can be contained in multiple windows and the value of a_k, b_k changes due to the change of the window center k, it is necessary to calculate the mean value of a_k and b_k in the window centered on element i. The leading filter output of the final weight matrix H is the result of the following formula:

$$
H_i = \frac{1}{|w|} \sum \sum_{i \in w_k} \left(a_k I_i + b_k \right) = \overline{b}_i I_i + \overline{b}_i, \tag{15}
$$

where $\overline{a_i} = (1/|w|) \sum_{k \in w_i} a_k$, $\overline{b}_i = (1/|w|) \sum_{k \in w_i} a = b_k$. In order to reduce the noise point, the impact of scattered points, here set a coefficient θ, the weight matrix into the formula (16).

$$
H_i = \begin{cases} 0, & H_i < 0 \\ H_i, & H_i. \end{cases} \tag{16}
$$

It can be seen that the weight matrix H has a detail texture and edge effect similar to that of the gray-scale guided graph I while having features similar to the binarization matrix A. And guiding the filter with a fast implementation, only the time complexity, is a good performance and can quickly achieve the filtering algorithm.

5. Experimental Results and Analysis

5.1. Experimental Results.

Figure 8 shows the effect of the spot extraction algorithm. The image test is focused on the image of corn diseases at 200×200 pixels, which includes the common leaf blight, *Cochliobolus heterostrophus*, and gray
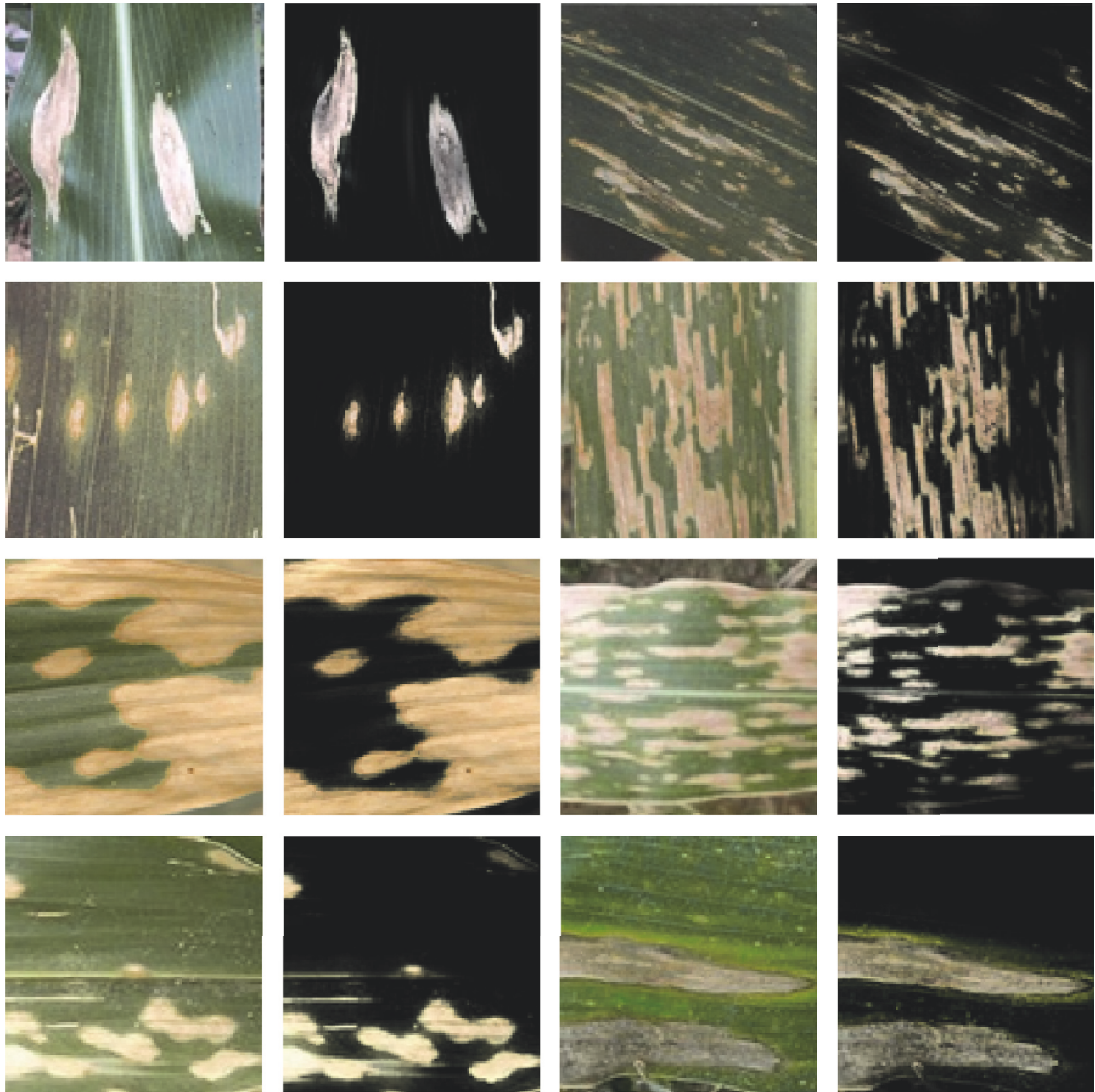
Figure 8: Results of lesion extraction.

spot disease. The experimental results show that the algorithm can effectively segment lesion location, lessen error points, and retain the original lesion images in color, edge, and texture features.

Figure 9 shows the effect of setting different values of on the extraction of lesion area in weight matrix calculation. Experiments show that, the lesion extraction area retains a fine and complete edge and texture but contains a lot of scattered points. With the increase of value, the scattered points are gradually reduced, and the lesion edge texture gradually becomes blurred and rough. Therefore, the value is set to 0.2.

5.2. *Comparison and Analysis.* Figure 10 is the comparison of the algorithm of this paper with one-dimensional Otsu threshold segmentation method and EM clustering segmentation method. The experimental results show that there are some misjudgment points when the Otsu method and the EM method are used in the segmentation, and the nonlesion location is misdiagnosed as the lesion and the lesion location is misjudged as nonlesion. The algorithm can effectively determine the location of the lesion, the less misjudgment, the region with fine texture edge extraction, and the original image of the lesion area to maintain good consistency.
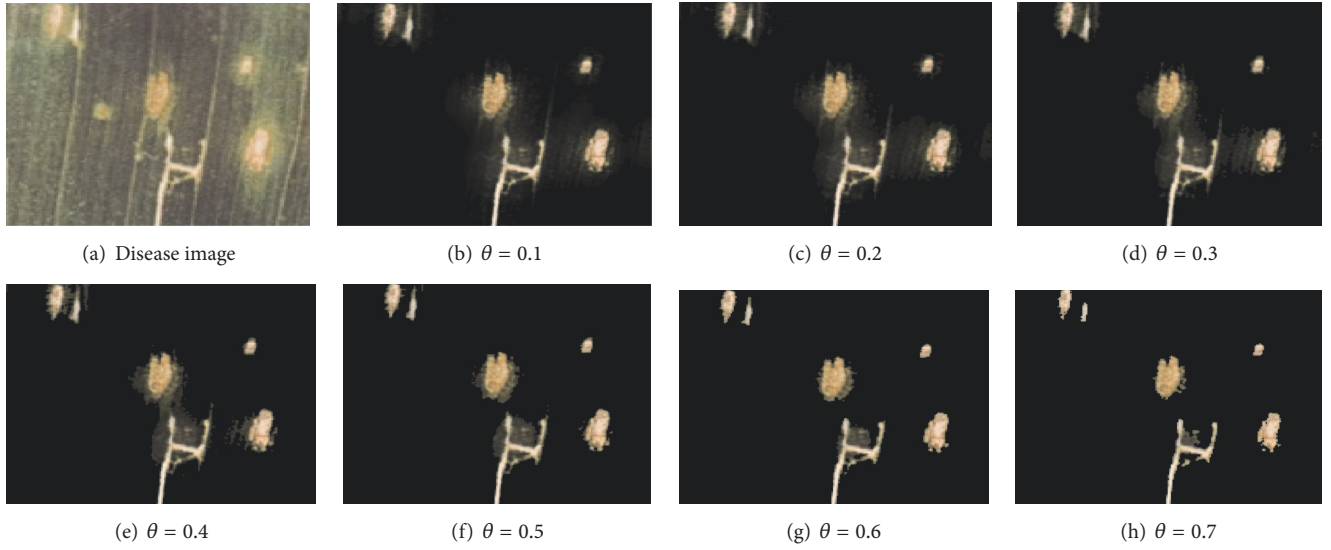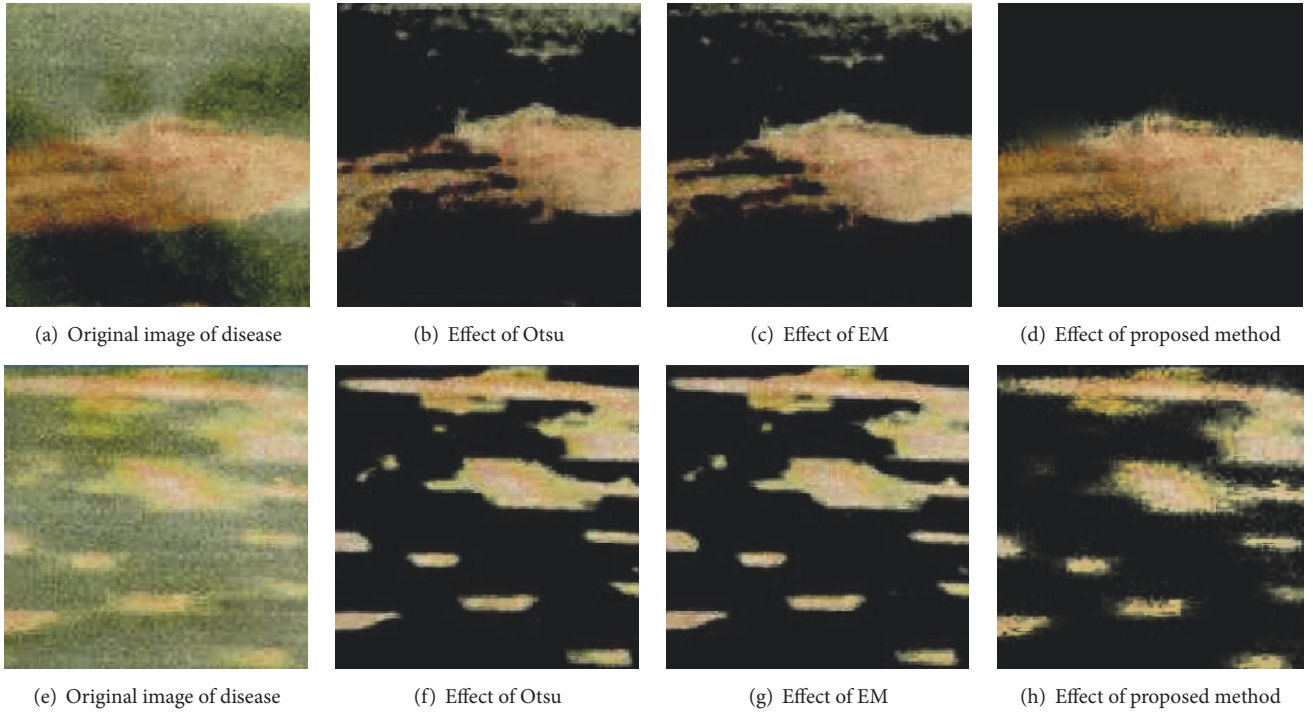
(a) Disease image (b) $\theta = 0.1$ (c) $\theta = 0.2$ (d) $\theta = 0.3$

(e) $\theta = 0.4$ (f) $\theta = 0.5$ (g) $\theta = 0.6$ (h) $\theta = 0.7$

FIGURE 9: The influence of different θ on the experimental results.

(a) Original image of disease (b) Effect of Otsu (c) Effect of EM (d) Effect of proposed method

(e) Original image of disease (f) Effect of Otsu (g) Effect of EM (h) Effect of proposed method

FIGURE 10: Comparison of the proposed algorithm with Otsu and EM algorithm.

The G-MRF [34] and TSRG algorithm [35] are the lesion segmentation algorithms for corn pests and diseases. Figure 11 shows the experimental comparison results of this algorithm with GMRF and TSRG algorithm. The experimental results show that the G-MRF algorithm can preserve the shape and texture of the lesion to be more intact but is susceptible to other scattered points, resulting in a small number of false points. While the TSRG algorithm can avoid the effect of scattered points of other nonlesion areas, its lesion area cannot retain the original shape of the lesion well. The proposed algorithm can not only preserve the original features of the lesion texture, but also avoid the influence of the scattered points and has high extraction quality.

6. Summary

First of all, this paper analyzes the data of maize diseases and pests damage image, in terms of both gray and chroma space of image in global and local statistical characteristics of the data analysis, taking differences in lesion location

(a) Original image of disease (b) Effect of G-MRF (c) Effect of TSRG (d) Effect of proposed method

(e) Original image of disease (f) Effect of G-MRF (g) Effect of TSRG (h) Effect of proposed method
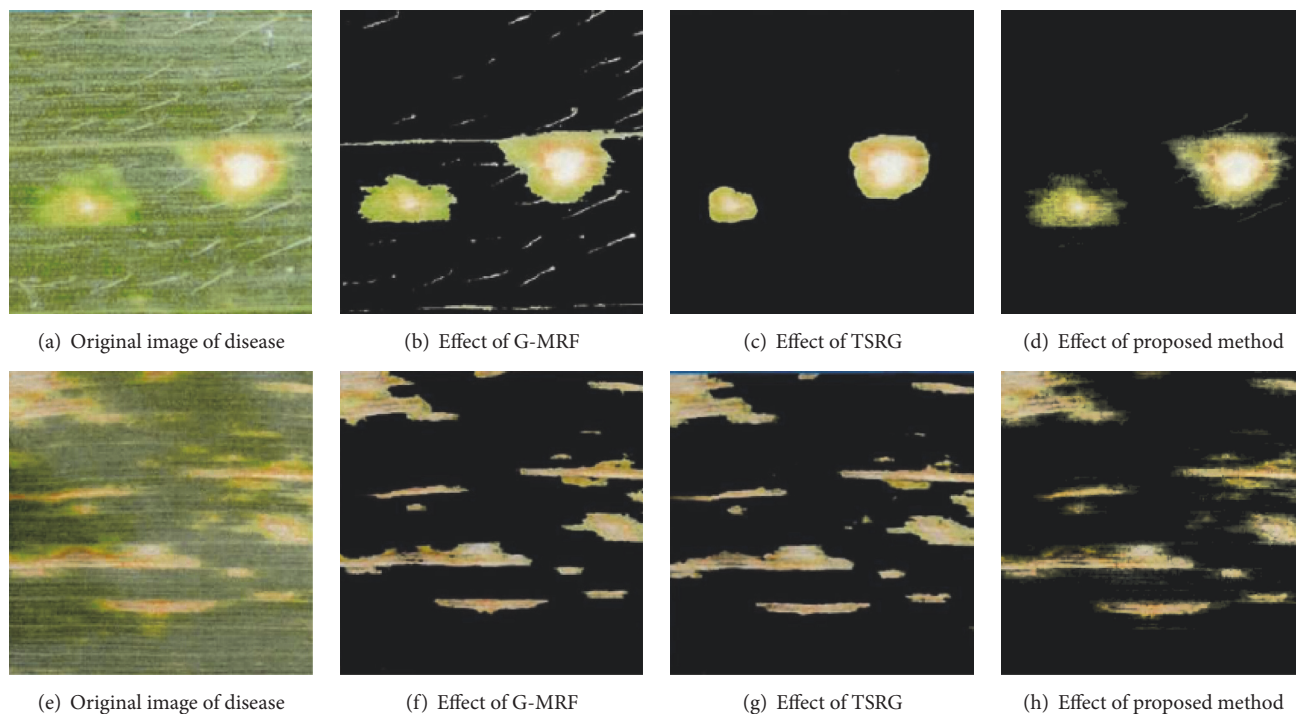
FIGURE 11: Comparison result of proposed algorithm with G-MRF and TSRG algorithm.

into account to obtain disease-free position of the color, texture, and so on, as for pest image segmentation and extraction of the lesion. According to the analysis of plant diseases and insect pests image, gray image and the HSI color space of color image are composed of two tuples to build a two-dimensional histogram, to better describe the distribution of pixel, to solve the one-dimensional histograms that cannot be clearly divided into target and background bimodal distribution situation, and to improve the use of traditional two-dimensional histogram in pest damage lesion extraction. For two-dimensional histogram threshold optimization problems, in this paper, the design of threshold optimization process is based on genetic algorithm; design for plant diseases and insect pests image data features of chromosome coding, combined with image analysis, results in the selection of initial population, using parallel selection, optimal preservation strategy, and adaptive mutation operator, to improve the searching efficiency; at the same time by setting the threshold fluctuations, the optimal threshold field fluctuation range continues to search and to achieve combined global search and local search. For binary segmentation of the lesion extraction results, this paper introduces the guidance filtering algorithm, using gray-oriented graph to compute the guiding and the filtering operation on the image segmentation, calculates the weight matrix, in the segmented lesion edge, texture fuzziness and rough position for restoration, to optimize the lesion extraction results and for better retention of plant diseases and insect pests in the edge of the lesion images and texture feature.

Finally, experimental results show that the algorithm of lesion region extraction is effective and discuss the weight matrix parameters setting effect of experimental results. Through comparison with the traditional Otsu algorithm, EM clustering algorithm indicates the applicability of the algorithm of crop disease, insect, and pest image, at the same time, through comparative tests with Gauss MRF and tsrg algorithm; the algorithm in the spot region extraction affects the superiority.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J. Wang, W. Zhang, L. Liu, and S. Huang, "Summary of crop diseases and pests image recognition technology," *Computer Engineering & Science*, vol. 36, no. 7, pp. 1363–1370, 2014.

[2] P. M. Keagy and T. F. Schatzki, "Machine recognition of weevil damage in wheat radiographs," in *Optics in Agriculture and Forestry*, vol. 1836 of *Proceedings of SPIE*, pp. 108–119, The International Society for Optical Engineering, 1993.

[3] C. Ridgway, R. Davies, and J. Chambers, "Imaging for the high-speed detection of pest insects and other contaminants in cereal grain in *Transit*," Sacramento, Calif, USA, July 29-August 1 2001.

[4] I. Y. Zayas and P. W. Flinn, "Detection of insects in bulk wheat samples with machine vision," *Transactions of the American Society of Agricultural Engineers*, vol. 41, no. 3, pp. 883–888, 1998.

[5] B. Chen, X. Guo, and X. Li, "Image diagnosis algorithm of diseased wheat," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 40, no. 12, pp. 190–195, 2009.

[6] X. Zou, "Research status of crop pest recognition over computer vision," *Computer Systems and Applications*, vol. 20, no. 6, pp. 238–242, 2011.

[7] D. Qiu, H. Zhang, X. Liu, and Y. Liu, "Design of detection system for agriculture field pests based on machine vision," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 38, no. 1, pp. 120–122, 2007.

[8] K. Wang, "Diagnosis of crop disease, insect pest and weed based on image recognition," *Chinese Academy of Agricultural Science*, 2005.

[9] J. Wang and G. Zhou, "The research of yellow stem borer identification system based on neural network," *Agriculture Network Information*, vol. 2, pp. 39–41, 2006.

[10] Z. Liang, F. Liu, Q. Zhao, and R. Du, "Application of mathematical morphology to superfamily level in insect taxonomy," *Zoological Systematics*, vol. 32, no. 1, pp. 142–152, 2007.

[11] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 7, pp. 1403–1416, 2015.

[12] B. Gu and V. S. Sheng, "A robust regularization path algorithm for ν-support vector classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 1, pp. 1–8, 2016.

[13] Z. Xia, X. Wang, X. Sun, Q. Liu, and N. Xiong, "Steganalysis of LSB matching using differences between nonadjacent pixels," *Multimedia Tools and Applications*, vol. 75, no. 4, pp. 1947–1962, 2016.

[14] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, 2016.

[15] Z. Xia, X. Wang, X. Sun, and B. Wang, "Steganalysis of least significant bit matching using multi-order differences," *Security and Communication Networks*, vol. 7, no. 8, pp. 1283–1291, 2014.

[16] H. Chang, M. Yang, and J. Yang, "Learning a structure adaptive dictionary for sparse representation based classification," *Neurocomputing*, vol. 190, pp. 124–131, 2016.

[17] X. Qi, G. Zhao, L. Shen, Q. Li, and M. Pietikäinen, "LOAD: local orientation adaptive descriptor for texture and material classification," *Neurocomputing*, vol. 184, pp. 28–35, 2016.

[18] X. Song, J. Zhang, Y. Han, and J. Jiang, "Semi-supervised feature selection via hierarchical regression for web image classification," *Multimedia Systems*, vol. 22, no. 1, pp. 41–49, 2016.

[19] Z. Liang, J. Sun, Q. Lin, Z. Du, J. Chen, and Z. Ming, "A novel multiple rule sets data classification algorithm based on ant colony algorithm," *Applied Soft Computing Journal*, vol. 38, pp. 1000–1011, 2016.

[20] Y. Lu, Z. Lai, Z. Fan, J. Cui, and Q. Zhu, "Manifold discriminant regression learning for image classification," *Neurocomputing*, vol. 166, pp. 475–486, 2015.

[21] R. Yang and Z. Wang, "Cross-oriented choquet integrals and their applications on data classification," *Journal of Intelligent and Fuzzy Systems*, vol. 28, no. 1, pp. 205–216, 2015.

[22] J. Zhang, Z.-H. Zhang, Y. Lin et al., "Evolutionary computation meets machine learning: a survey," *IEEE Computational Intelligence Magazine*, vol. 6, no. 4, pp. 68–75, 2011.

[23] M. J. Kavetha, "Coevolution evolutionary algorithm: a survey," *International Journal of Advanced Research in Computer Science*, vol. 4, no. 4, 2013.

[24] X. Jiang, J. Ma, and C. Lei, "Kernel Evolutionary Algorithm for Clustering," in *Bio-inspired Computing—Theories and Applications. BIC-TA*, M. Gong, L. Pan, T. Song, and G. Zhang, Eds., vol. 682 of *Communications in Computer and Information Science*, Springer, Singapore, Singapore, 2016.

[25] C. Huo, R. Zeng, Y. Wang, and M. Shang, "A Multi-Parent Crossover Based Genetic Algorithm for Bi-Objective Unconstrained Binary Quadratic Programming Problem," in *Bio-inspired Computing—Theories and Applications. BIC-TA*, M. Gong, L. Pan, T. Song, and G. Zhang, Eds., vol. 682 of *Communications in Computer and Information Science*, Springer, Singapore, Singapore, 2016.

[26] J. Yang and C. Liu, "Research on color space and its transformation in digital image processing," *Journal of Shangqiu Vocational and Technical College*, vol. 8, no. 2, pp. 25–27, 2009.

[27] C. Lei, J. Ma, and X. Jiang, "Unsupervised Image Segmentation Based on Watershed and Kernel Evolutionary Clustering Algorithm," in *Bio-inspired Computing—Theories and Applications. BIC-TA*, M. Gong, L. Pan, T. Song, and G. Zhang, Eds., vol. 682 of *Communications in Computer and Information Science*, Springer, Singapore, Singapore, 2016.

[28] Y. Wang, *Research on Key Technology of The Image Segemention*, Jiangnan University, Beijing, China, 2008.

[29] Y. Liang, R. Pang, and Y. Zhu, "Two dimensional Otsu line segmentation method for gray level images," *Computer Engineering and Applications*, vol. 48, no. 33, pp. 178–182, 2012.

[30] Z. Chen, Y. Jiang, and R. Wang, "Ant Colony Optimization with Different Crossover Schemes for Continuous Optimization," in *Bio-Inspired Computing—Theories and Applications*, M. Gong, L. Pan, T. Song, K. Tang, and X. Zhang, Eds., vol. 562 of *Communications in Computer and Information Science*, Springer, Berlin, Germany, 2015.

[31] K. Li and L. Xiong, "Community detection based on an improved genetic algorithm," *Communications in Computer and Information Science*, vol. 575, pp. 32–39, 2016.

[32] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 120–129, 2015.

[33] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.

[34] S. Li, *Research on Maize Diseases Intelligent Diagnosis Based on Disease Image*, Shihezi University, Beijing, China, 2010.

[35] J. Li, *Research and Implementation of Intelligent Image Processing Algorithm on Maize Leaf Ill Spots*, Beijing University of Posts and Telecommunications, Beijing, China, 2010.