

Citrus Huanglongbing Recognition Algorithm Based on CKMOPSO

Hui Wang, Shenzhen Institute of Information Technology, Shenzhen, China

Tie Cai, Shenzhen Institute of Information and Technology, Shenzhen, China

Wei Cao, Shenzhen Institute of Information Technology, Shenzhen, China

ABSTRACT

In view of the similarity of characteristics between the features of the disease images and the large dimension and the features correlation of the disease images, this will lead to the generation of feature redundancy and will introduce a serious impact on the recognition efficiency and accuracy of citrus Huanglongbing. In addition, they have the defects of high cost of detection algorithms and low detection accuracy. This will occur in the image cutting feature extraction stage, so this paper uses the citrus Huanglongbing recognition algorithm based on kriging model simplex crossover local-based search multi-objective particle swarm optimization algorithm (CKMOPSO) with strong classification capabilities from the original disease image features to select feature vectors. Experimental results show that this is an effective recognition method.

KEYWORDS

Citrus Huanglongbing Recognition Algorithm, Feature Extraction, Multi-Objective Particle Swarm Optimization Algorithm, Riging Model Simplex Crossover Local

1 INTRODUCTION

Citrus fruits are one of the most popular fruits in the world, and they are also an important part of the agricultural economy. However, at the beginning of the 20th century, citrus huanglongbing was found in many areas of southern China for the first time. Later, citrus huanglongbing was discovered in other provinces (Sichuan, Guangxi and Jiangxi) in citrus production areas. The pathogen of Huanglongbing is a bacterial pathogen that is mainly transmitted by psyllids (Grafton-Cardwell et al., 2013), and it is one of the biggest virus hazards affecting citrus cultivation in the world. The pathogen is extremely infectious and has a very high pathogenicity rate. Once a citrus tree is infected with Huanglongbing pathogen, the citrus tree will lose the ability to grow fruits in a short time, which will cause the production of citrus to drop sharply. Therefore, the citrus huanglongbing of citrus growers has evolved into a state of “talking about Huanglongbing discoloration”. This has led to the research on the identification and prediction of citrus Huanglongbing becoming an urgent problem to be solved. Scholars have conducted in-depth research on the characteristics of citrus Huanglongbing (Gottwald, 2010). From the current development analysis of citrus Huanglongbing detection and identification technology, near-infrared spectroscopy technology and detection and prevention technology based on the Internet of Things have become the mainstream development of

DOI: 10.4018/IJCI.20211001.0a10

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

the technology for identifying citrus Huanglongbing. Near-infrared spectroscopy analysis technology has the advantages of rapid identification and better reliability of results (H. Xiao, 2019). The detection and prevention technology of citrus Huanglongbing based on the Internet of Things has the advantages of low cost and rapid identification, but its recognition rate needs to be improved. The key to improving its recognition rate is to find an effective disease image recognition algorithm. So, citrus Huanglongbing has attracted the attention of domestic and foreign scholars, and they have published many research results. Sankaran et al. used a model based on deep convolutional networks to identify citrus Huanglongbing. The authors studied whether periodic or “temporary” models are effective for tasks involving sequence, vision and other aspects, and developed a cyclic convolution architecture. This is suitable for end-to-end trainable large-scale visual learning. Authors demonstrate the value of these models in benchmark video recognition tasks, image description and retrieval problems, and video narration (Sankaran et al., 2013). Mishra et al. study the potential ability of visible light and near infrared (VIS-NIR) spectroscopy to identify Huanglongbing-infected citrus trees (Mishra et al., 2012). Two different spectroradiometers with a spectral range of 350 to 2500 nm are used to collect the canopy reflectance spectrum data, and three classification techniques are used to classify the data: nearest neighbor (KNN), logistic regression (LR) and support vector machine (SVM). The authors concluded that using only one canopy reflectance observation per tree is not enough. Due to the large differences in canopy reflectance data, no classification method can successfully distinguish healthy trees from Huanglongbing infected trees. Alireza et al. extracted several types of texture features from leaf images, and used five different feature selection methods to rank the best feature set that can describe infection features. The performance of the seven classifiers is based on a stepwise classification method. An assessment was conducted (Alireza, 2013). However, using this method will incorrectly classify some undernourished samples into other categories. In view of the similarity of characteristics between the characteristics of the disease images and the large dimension, especially the characteristics of the disease images may also be related, resulting in the generation of feature redundancy, which has a serious impact on the efficiency and accuracy of citrus Huanglongbing image recognition. In addition, they have the defects of high cost of detection algorithms and low detection accuracy. This effect will occur in the image cutting feature extraction stage, so this paper uses the citrus Huanglongbing recognition algorithm based on CKMOPSO, which uses CKMOPSO to automatically select feature vectors with strong classification capabilities from the original disease image features for use in Huanglongbing image recognition, experimental results show that this is an effective recognition method.

2 CITRUS HUANGLONGBING RECOGNITION ALGORITHM BASED ON CKMOPSO

2.1 Image Acquisition

Under natural conditions, taking the common characteristics of Huanglongbing (leaf yellowing, swollen veins, uneven pages) in the growth of citrus as the research object, image acquisition is carried out through the Internet of Things network based on the routing algorithm proposed in (H. Wang et.al, 2020). To obtain better quality images of citrus Huanglongbing, the camera of the Internet of Things uses a camera with automatic gain control (AGC) function (Fowler, 2004). If the image is too dark, the average gain is increased, and if the image is too bright, the gain is reduced. For commercial cameras, AGC is usually a must-have function. Since this study requires the true reflectivity of the blade, the effect of AGC is not ideal. In order to eliminate the influence of AGC on images, similar backgrounds are used in all images. Then, the ratio of the average value of the background gray in all images to the average value in each image is used as the gain multiple of each individual image to calibrate them. The finally obtained image is of high quality, which is convenient for extracting the characteristics of citrus Huanglongbing.

2.2 Image Preprocessing of Citrus Huanglongbing

Since the quality of the image will be affected by certain factors, the collected Citrus Huanglongbing images often contain some random interference factors. Therefore, if you want to improve the accuracy of Huanglongbing disease image recognition, you must first perform image preprocessing. And you provide suitable images for feature extraction.

A bilateral filter is used to smooth the citrus Huanglongbing image. The advantage of this filter is that it considers the degree of pixels between pixel values, which can effectively protect high-frequency detail pixels, which enables edge protection for images with a high degree of blur, and effectively overcomes the problem of high blurring and some loss of detail. At the same time, the image sampling window is adjusted by adaptively adjusting the image window. This method is combined with a bilateral filter and used for image smoothing processing of Citrus Huanglongbing. When it is easy to meet the filtering requirements, the image sampling window is reduced to reduce resource consumption; when it is difficult to meet the filtering requirements, in order to obtain a better filtering effect, the image sampling window needs to be enlarged.

The image segmentation algorithm (Generic fuzzy rule based image segmentation algorithm, GFRIS) (Karmakar et al., 2002) is used based on the general fuzzy rule. Due to the difference between the leaf color of the diseased spot and the normal part, we compares the ultra-green feature with the largest class based. The threshold segmentation method of variance is combined to realize the segmentation of diseased images. This algorithm uses the spatial relationship between pixels. The GFRIS automatically approximates the key weighting factors and thresholds in the definition of fuzzy rules and neighborhood systems. Therefore, an image segmentation algorithm based on general fuzzy rules is used to separate Huanglongbing spots and background leaf parts, so as to obtain images of Huanglongbing spots, and then the characteristics of Huanglongbing spots can be extracted. However, due to the difference and closeness of the leaf colors between the diseased spots and the normal parts, it is difficult to segment them by ordinary methods. Therefore, we combines the super green feature with the GFRIS algorithm to achieve precise segmentation of disease images and facilitate the extraction of Huanglongbing spots.

2.3 Extraction of Feature of Citrus Huanglongbing Image

We extracted color and texture features from the preprocessed citrus Huanglongbing image. Using gray-scale statistical histogram features for color feature extraction, the formulas (6.2.1) ~ (6.2.7) give the gray scale of each image when k is used to represent the gray scale (from 0 to 255). Extract color statistical features from the histogram ($p(k)$). Local binary mode and local similarity mode are used for image texture feature extraction. Local binary mode is an innovative image texture descriptor, which was first introduced by Ojala et al. In this method, each matrix element in the 3×3 neighborhood is compared with the center element. If it is greater than the value of the center element, it is replaced with 1; if it is less than the center element, it is replaced with 0. Then, the threshold is assumed to be an 8-bit binary number, and its corresponding decimal number will replace the central element. Similarly, in the local similarity mode, the neighborhood element is compared with the center element. However, the relationship between each adjacent element and the central element is described by a 2-digit code. In this method, if the neighborhood elements are located within, below, or within the defined similarity range, then 00, 01, or 10 are used instead, then, multiply each element code by its respective weight (Pourreza et al. 2014). The local binary mode and the local similarity mode use the same method to prevent changes caused by code rotation. Rotation invariance is obtained by rotating the code to obtain the smallest possible decimal value. We extract texture feature extraction by formulas (6.2.8) ~ (6.2.17).

Mean

$$\mu = \sum_k k p(k) \quad (1)$$

Standard

$$\sigma = \sqrt{\sum_k (k - \mu)^2 p(k)} \quad (2)$$

Smoothness

$$1 - 1 / (1 + \sigma^2) \quad (3)$$

Evenness

$$\sum_k p(k)^2 \quad (4)$$

Entropy

$$-\sum_k p(k) \log\{p(k)\} \quad (5)$$

Maximum gray level probability

$$(k | p(k) = \max) \quad (6)$$

Range length

$$\{\max(k | p(k) \neq 0) - \min(k | p(k) \neq 0)\} \quad (7)$$

Mean

$$\mu' = \sum_{i,j} i g(i, j) \quad (8)$$

Standard

$$\sigma'^2 = \sum_{i,j} (i - \mu')^2 g(i, j) \quad (9)$$

Entropy

$$-\sum_{i,j} g(i, j) \log\{g(i, j)\} \quad (10)$$

Evenness

$$\sum_{i,j} \{g(i,j)\}^2 \quad (11)$$

Homogeneity

$$\sum_{i,j} g(i,j) / \{1 + (i - j)^2\} \quad (12)$$

Inertia

$$\sum_{i,j} (i - j)^2 g(i,j) \quad (13)$$

Cluster shadow

$$\sum_{i,j} (i + j - 2\mu')^3 g(i,j) \quad (14)$$

Cluster protrusion

$$\sum_{i,j} (i + j - 2\mu')^4 g(i,j) \quad (15)$$

Maximum gray level probability

$$\max\{g(i,j)\} \quad (16)$$

Relativity

$$\sum_{i,j} (i - \mu')(j - \mu') / \sigma^2 g(i,j) \quad (17)$$

2.4 Feature Optimization Algorithm Based on CKMOPSO

As the number of features in image recognition increases, the complexity of recognition also increases. The main reason is that it cannot quickly and accurately find the optimal feature that can represent the nature of the problem from a large number of candidate features. It is especially important to find a better feature selection method. According to the characteristics of the feature set (discrimination, reliability, independence), the methods of feature selection mainly include exhaustive method, random method and heuristic method. The exhaustive method ensures that the optimal can be found. Its disadvantage is that it has high complexity and low robustness; the random method is simple and fast, but its disadvantage is that it cannot guarantee the optimal result; the heuristic method is fast and strong robust, but its disadvantage is that it cannot guarantee the optimal result and does not consider the restrictive conditions. Therefore, we use CKMOPSO to select representative features. The specific steps are as follows:

- (1) Build the objective function

The main feature basis in the process of image feature recognition is color feature and texture feature. Usually, it is necessary to select the optimal feature from many features to recognize the citrus Huanglongbing image. Therefore, the feature selection process is a multi-objective optimization process. The goal is to select the maximum color feature value and the maximum texture feature value, as shown in (6.2.18) and (6.2.19).

For the characteristics color, the leaves with Huanglongbing spots also have a lot of green parts. Therefore, the optimal color characteristics are selected by normalizing color components b, Cb, and the correlation between color information. The formula of the objective function is as follows.

$$f_1(x) = \frac{d_y}{d_g} x - \alpha_1 * \frac{k_1}{n_1} \quad (18)$$

Where, n_1 is the total number of original features, s_g is the estimated number of features selected; n_2 is the yellow feature value; k_2 is the green value.

For feature texture, select the optimal texture feature by measuring the lesion circumference, lesion area, lesion roundness, shape factor, and related information measurement factors.

$$f_2(x) = \frac{s_o}{s_g} x - \alpha_2 * \frac{k_2}{n_2} \quad (19)$$

Where, s_o is the texture feature with the largest feature value; s_g is the average value of texture features; n_2 is the total number of original features, and k_2 is the expected number of selected features;

The objective function is

$$\begin{aligned} \max f_m(x) \quad & m=1,2 \\ & k_1 < n_1 \\ & k_2 < n_2 \end{aligned} \quad (20)$$

(2) Algorithm flow based on CKMOPSO feature optimization

1) Initial population POP

FOR $i = 0$ TO n ; # n =number of particles

Initial POP $[i]$; # Use color features and texture features as particles

2) Initialize the velocity of each particle

FOR $i = 0$ TO n ;

$X_i(k) = 0$; # The velocity of the k -generation particle i

- 3) Initialize the memory of each particle (this memory plays a guiding role in the search space particle flight. The memory is also stored in the repository REP):

FOR $i = 0$ TO n ;

PBESTS $[i] = \text{POP}[i]$;

- 4) To correct the target space is to estimate the value of the objective function of each particle in the population POP;
- 5) The objective function value is used for non-dominated sorting, and the archive controller is used to determine whether to store the position of a certain non-dominated particle in REP.
- 6) The corrected target space is gridded to generate a grid space for exploring the current search space, and these grids are used as a coordinate system to locate particles. In this coordinate system, the coordinate system of each particle is defined according to the value of its objective function. The global optimal particle optimal feature combination is selected through the grid space.

2.5. Analysis of Experiment Results

Based on the optimal feature group obtained by the CKMOPSO feature optimization algorithm, the Bayes-based discriminant method is used to predict the health situations of citrus leaves, so as to achieve the purpose of identifying citrus Huanglongbing. We uses Python3.7 software to simulate and test 56 citrus leaf samples. The unhealthy leaves in the sample database are marked as Positive, the healthy leaves are marked as Negative, and the confidence of the prediction results is marked as Confidence. The classification result is shown in Figure 1, the confusion matrix represents the recognition result, and provides a data basis for indicator performance analysis. We uses indicator confusion matrix, precision rate, recall rate and F value to measure the effectiveness of the proposed algorithm. The experimental results are compared with the image recognition algorithm based on genetic algorithm (Karnakar et al. 2017) and the image recognition algorithm based on differential evolution algorithm (Zilinskas et al. 2015) (Notice: citrus Huanglongbing recognition based on feature optimization algorithm based on CKMOPSO, the citrus Huanglongbing recognition algorithm based on genetic algorithm feature optimization and the citrus Huanglongbing recognition algorithm based on differential evolution algorithm feature optimization are represented by CR, GR and DR respectively).

- 1) The confusion matrix: it is a table that describes the results of classification in detail, whether it is classified correctly or incorrectly, and distinguishes between different categories.
- 2) Accuracy: it refers to the proportion of the number of positive samples correctly classified by the classifier to the number of all positive samples classified by the classifier.

$$precision = \frac{TP}{FP + TP} \quad (21)$$

Where, TP represents the right positive sample, and FP represents the wrong positive sample.

Table 1. Confusion matrix

	Predicted as a positive sample	Predicted as a negative sample
Label as positive sample	35	2
Label as negative sample	4	31

3) Recall rate: it refers to the proportion of the number of positive samples classified correctly by the classifier to the number of all positive samples with the original label.

$$recall = \frac{TP}{FN + TP} \quad (22)$$

Where, TP represents the right positive sample, and FN represents the wrong negative sample.

4) F1 value: it is the harmonic average of the precision rate and the recall rate, and measures the similarity between the predicted value and the correct value.

$$F1 = \frac{2 * precision * recall}{precision + recall} \quad (23)$$

2.5.1. Confusion Matrix

Figure 1 shows the classification result of the picture. The prediction results of unhealthy citrus leaves are: the prediction accuracy rate is 89.7%, and the average reliability percentage is 99.8%; the prediction results of healthy leaves are: the prediction accuracy rate is 92.5%, and the average reliability percentage is 91.02%. The confusion matrix obtained in light of Figure 1 is shown in Table 1.

According to Table 1, the true rate is 0.945, the false negative rate is 0.054, the false positive rate is 0.114, the true negative rate is 0.885, the accuracy rate is 0.916, and the average accuracy rate is 0.915. According to these data, the citrus Huanglongbing recognition algorithm based on CKMOPSO is effective in recognition and has certain practicability.

2.5.2. Precision Rate-Recall Rate

The precision-recall results of the three algorithms are shown in Table 2 and Table 3. CR has the highest precision and GR has the lowest precision. This shows that the number of correct positive samples identified by the DR is the largest; the number of correct positive samples identified by the DR is the smallest. At the same time, it can be seen from Table 3 that CR has the highest recall rate and GR has the lowest recall rate. This shows that the ratio of the positive samples with the correct label to the positive samples identified by the CR is the largest; the ratio of the positive samples with the correct positive samples identified by the GR to the positive samples with the label is the smallest. This verifies that the proposed algorithm has high accuracy and certain practicality.

2.5.3. F1 Value

The experimental results of the F1 value are shown in Table 4. The similarity between the result of the CR and the real sample is the highest, and the similarity between the result of the GR and the real

Figure 1. Identification results of citrus leaf Huanglong disease



Table 2. Comparison of the accuracy of the three algorithms

Algorithm	CR	GR	DR
Precision rate	0.897	0.815	0.852

Table 3. Comparison of the recall rates of the three algorithms

Algorithm	CR	GR	DR
recall rate	0.945	0.895	0.901

Table 4. Compare the F1 values of the three algorithms

Algorithm	CR	GR	DR
F1value	0.92	0.853	0.876

sample is the lowest. This shows that the recognition effect of the proposed recognition algorithm is better.

3 CONCLUSION

We applied the feature optimization method based on CKMOPSO to the identification of citrus Huanglongbing, and proposed a citrus Huanglongbing recognition algorithm based on CKMOPSO feature optimization. Through the gray-scale statistical histogram, local binary mode and local similarity mode, the color feature and texture feature are extracted from the preprocessed citrus Huanglongbing image. The feature selection method based on CKMOPSO is used to select feature groups, which will effectively make the recognition accuracy higher. Experiments show that the recognition accuracy of the proposed recognition algorithm reaches 99.99%, which shows that the proposed algorithm can effectively identify citrus Huanglongbing.

ACKNOWLEDGEMENT

This paper is supported by Science and Technology Ph.D. Research Startup Project, China (SZIIT2021KJ001).

REFERENCES

- Fowler, K. R. (2004). Automatic gain control for image-intensified camera. *Instrumentation and Measurement [J]. IEEE Transactions on Instrumentation and Measurement*, 53(4), 1057–1064. doi:10.1109/TIM.2004.831494
- Gottwald, T. R. (2010). Current epidemiological understanding of citrus Huanglongbing. *Annual Review of Phytopathology*, 48(1), 119–139. doi:10.1146/annurev-phyto-073009-114418 PMID:20415578
- Grafton-Cardwell, E. E., Stelinski, L. L., & Stansly, P. A. (2013). Biology and management of Asian citrus psyllid, vector of the huanglongbing pathogens. *Annual Review of Entomology*, 58(1), 413–432. doi:10.1146/annurev-ento-120811-153542 PMID:23317046
- Karmakar, G. C., & Dooley, L. S. (2002). A generic fuzzy rule based image segmentation algorithm [J]. *Pattern Recognition Letters*, 23(10), 1215–1227. doi:10.1016/S0167-8655(02)00069-7
- Li, J., Pan, Q., & Xie, S. (2012). An effective shuffled frog-leaping algorithm for multi-objective flexible job shop scheduling problems [J]. *Applied Mathematics and Computation*, 218(18), 9353–9371. doi:10.1016/j.amc.2012.03.018
- Liang, J. J., Qin, A. K., Suganthan, P. N., & Baskar, S. (2006). Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. *IEEE Transactions on Evolutionary Computation*, 10(3), 281–295. doi:10.1109/TEVC.2005.857610
- Liu, D. S., Tan, K. C., Huang, S. Y., Goh, C. K., & Ho, W. K. (2008). On solving multiobjective bin packing problems using evolutionary particle swarm optimization. *European Journal of Operational Research*, 190(2), 357–382. doi:10.1016/j.ejor.2007.06.032
- Liu, J., & Lampinen, J. (2005). A Fuzzy Adaptive Differential Evolution Algorithm. *Soft Computing*, 9(6), 448–462. doi:10.1007/s00500-004-0363-x
- Liu, Z., & Wang, Y. (2019). Handling Constrained Multiobjective Optimization Problems With Constraints in Both the Decision and Objective Spaces. *IEEE Transactions on Evolutionary Computation*, 23(5), 870–884. doi:10.1109/TEVC.2019.2894743
- Maitra, M., & Chatterjee, A. (2008). A hybrid cooperative–comprehensive learning based PSO algorithm for image segmentation using multilevel thresholding. *Expert Systems with Applications*, 34(2), 1341–1350. doi:10.1016/j.eswa.2007.01.002
- Mishra, A. R., Karimi, D., & Ehsani, R. (2012). Identification of citrus greening (HLB) using a VIS-NIR spectroscopy technique. *Transactions of the ASABE*, 55(2), 711–720. doi:10.13031/2013.41369
- Pourreza, A., Lee, W. S., & Raveh, E. (2014). Citrus Huanglongbing Detection Using Narrow-Band Imaging and Polarized Illumination. *Transactions of the ASABE*, 2014, 259–272.
- Sankaran, S., Maja, J., Buchanon, S., & Ehsani, R. (2013). Huanglongbing (Citrus Greening) Detection Using Visible, Near Infrared and Thermal Imaging Techniques. *Sensors (Basel)*, 13(2), 2117–2130. doi:10.3390/s130202117 PMID:23389343
- Wang, H., Li, K., & Pedrycz, W. (2020). An Elite Hybrid Metaheuristic Optimization Algorithm for Maximizing Wireless Sensor Networks Lifetime With a Sink Node. *IEEE Sensors Journal*, PP, 20(99), 1–1. doi:10.1109/JSEN.2020.2971035
- Xiao, H. (2019). *Research on Diagnosis Method of Citrus Huanglongbing Based on Hyperspectral Imaging Technology*. East China Jiaotong University.

Hui Wang is currently with School of Software, Shenzhen Institute of Information Technology. She obtained her PhD from the South China Agricultural University, Guangzhou, China. Her research interests include, evolutionary algorithm, multi-objective optimization algorithm, granular computation, and evolutionary computing.

Te Cai is a professor, and the Dean of the Software School of Shenzhen Institute of Information Technology. Research fields: information processing and pattern recognition, machine learning, artificial intelligence.

Cao Wei received his doctor in computer science and technology, from National University of Defense Technology. His research interests focus on high performance parallel computing, machine learning development and application.