Research of OpenMV Intelligent Monitoring and Disease Identification System

WANG Xiu-qing\textsuperscript{1,a}, CHEN Qi\textsuperscript{1,b}, YAO Shi-jia\textsuperscript{1,c} and JIA Xiao-yun\textsuperscript{1,d}

\textsuperscript{1}College of Electronic Information and Automation, Tianjin University of Science & Technology, Tianjin 300222, China

E-mail: \textsuperscript{a}lxqly@163.com; \textsuperscript{b}18322031620@163.com; \textsuperscript{c}yaoshijia6@163.com; \textsuperscript{d}1094622439@qq.com

Abstract. Aiming at the requirement of video surveillance and disease identification in modern greenhouse, this paper designed a system consisted by greenhouse intelligent monitoring and disease identification system. At first, this paper using OpenMV camera designed a disease monitoring system which could recognize and collect disease images automatically by multi-color thresholds tracking. Then, this paper designed a cuckoo search and BP neural network collaborative search (CSBP-CS) algorithm, the algorithm combined the global search capability of Cuckoo Search(CS) and back-propagation algorithm of BP algorithm to optimize weights and thresholds collaboratively. this paper took three tomato diseases and normal leaves as research objects, firstly step was to separate the disease spots from disease images, then was to extract 56 classification features and select 47 excellent classification features by relief F to construct CSBP-CS RF classifier. Finally, this paper compared the classification accuracy of CSBP-CS with CSBP-CS RF network and analyzed the effectiveness of relief F. The simulation results showed that the average correct recognition rate of the CSBP-CS RF was approximately equal to CSBP-CS under the same conditions, but CSBP-CS RF is more simple, so Relief F can help to speed up the efficiency of CSBP-CS.

1. Introduction

Tomato diseases can lead to a decline in crop yield and quality seriously. Therefore, when diseases occurred, agricultural producers need to respond quickly to employ pesticide rationally. so it is indispensable to monitoring greenhouse diseases and research about disease classification models.

In recent years, OpenMV intelligent monitoring system is mainly used in car tracking\cite{1,2}. In the field of video surveillance in Greenhouse, Liu Wenhui et al\cite{3} combined cloud computing technology with video surveillance technology to realize the storage and calculation of massive video files, but plant diseases do not occur all the time, it resulting in the waste of memory resources.

On the other hand, the research on crop disease identification model has attracted more and more scholars' attention\cite{4-5}. Sun Chen\cite{6} put forward the method of CS-BP to forecast the stock market. MA Wei\cite{7} combined cuckoo with SVM, and proposed CS-SVM algorithm for human body recognition. Therefore, the Cuckoo Search(CS) has less optimization parameters and strong global search ability, but easy to fall into local optimization and too slow to convergence in the later stage. So this paper constructed the cooperative search algorithm of CS and BP network (CSBP-CS) and used it for tomato disease identification models.

Firstly, this paper separated the disease spots and extracted 56 classification features as a feature
description. Then, Relief F was used to select 47 excellent classification features to construct CSBP-CS classifier. Besides, this paper designs a disease monitoring system based on OpenMV intelligent monitoring system, which can automatically recognize and collect disease image. The collected disease image is recognized by classifier. Finally, this paper compared the recognition accuracy rate of CSBP-CS with CSBP-CS RF network and analyzed the effect of relief F.

2. Disease Image Monitoring
In this paper, OpenMV intelligent monitoring system was applied to the field of greenhouse disease monitoring to realize the function of automatically tracking and collecting disease images with multi-color thresholds tracking.

Because of the disease leaves generally consist of normal disease-free green and brown or green and white or green and yellow spots, this paper combined this characteristic, and used the camera to collect diseased leaves by double color recognition, when tracing the coexistence of green and yellow, or green and brown, or green and white in the surveillance video, the LED warning light turned bright, and the camera automatically took photos and saves pictures. At the same time, the platform was controlled by the steering gear to rotate continuously for 180° from top to bottom and from right to left. Finally, the experiment achieved the effect of the steering gear to rotate continuously and multi-color recognition to take pictures and save pictures. The image copy is saved to the file system in the path by the automatic accumulation method of the image name taken each time and the collected disease images will be recognized by the classifier the sort of disease.

3. Image Segmentation and Feature Extraction

3.1 Disease Image Acquisition
This paper collected 260 images including 65 powdery mildew disease images, 65 gray mould disease images, 65 late blight disease images and 65 normal leaves from the network as training set of classifier, and the disease category in training set was verified by experts. Then the collected disease images by OpenMV intelligent monitoring module will be the test set for testing the effect of classifier, the test set includes 35 powdery mildew disease images, 35 gray mould disease images, 35 late blight disease images and 35 normal leaves, and they are also verified by experts. Most of the selected images are natural backgrounds, so the clipping of complex natural backgrounds when segmenting diseased parts is more likely to cause errors, but it is more suitable for practical applications.

3.2 Disease Image Segmentation
This paper separated the disease spots by the difference of chroma information between disease spots and green chroma based on the RGB-HSI transformation algorithm. HSI model reflects the way human visual system observes color, and separate the color image to chrominance (H), saturation (S), and intensity (I), so it has great advantages in the field of image segmentation. The conversion formula of RGB coordinates to HSI coordinate chrominance H is as follows:

\[
H = \begin{cases} 
\theta, & G \geq B \\
2\pi - \theta, & G < B 
\end{cases}
\]  

in formula (1):

\[
\theta = \cos^{-1}\left(\frac{(R-G) + (R-B)}{2\sqrt{(R-G)^2 + (R-B)(G-B)}}\right)
\]  

Then the chrominance component is normalized to a chromaticity adjustment factor h. By adjusting the range of h, the paper separated the disease spot. This paper selected \(h\in[1/6,0.5]\) as the basic value, and fine-tune between \([1/6,0.6]\) when the effect is poor. The segmentation result is shown in Figure 1.
It can be seen from the figure that HSI model can quickly and excellently segment the three kinds of disease images in this paper. By experiment, the algorithm only takes 2.036 seconds to divide 100 disease images in MATLAB, it can be known that the adaptability of HSI model is wide and it is suitable for quickly extracting a large number of diseased leaf images.

3.3 Feature Extraction
In this paper, 400 images were segmented for extracting features set. Firstly, 9 color features was extracted including the R, G, and B color components of the color histogram RGB, the hue, saturation, and brightness of the color space HSV, and the chroma, saturation, and intensity of the color model HSI. Then, the gray-gradient concurrence matrix method[8] was used to extract Gray Level Average, Inverse Gap, Inertia, Mixed Entropy, Gradient Entropy, Gray Level Entropy, Relevance, etc, respectively in RGB histogram, the H, S and V three channels of HSV color space, and H, S, I three channels of HSI color model, a total of 45 texture features. Finally, the tamura method[9] was used to extract the roughness and contrast of the disease plot in RGB coordinates. In total, the 56 features of every samples were extracted as the feature set of the dimension space [400,56], then relief F algorithm was used to select excellent features.

3.4 Excellent Feature Selection by Relief F
The relief proposed at the earliest time is only can deal with two labels of data. So Kononeill extends it and gets the relief F algorithm which can deal with multi-class data. When dealing with multi-class problems, each time a sample R is randomly extracted from the training sample set, then K near-Hits of R are found from the same sample set as R, and K near-Hits are found from the different sample sets of each R. Then the weight of each feature is updated by formula, which is as follows:

\[
W(A) = \frac{\sum_{i} \text{diff}(A, R, H_i) \sum_{C \neq \text{class}(R)} \left[ 1 - p(\text{class}(R)) \right] \sum_{j} \text{diff}(A, R, M_j(C))}{mk}
\]

In the formula above, M_j(C) denotes the jth nearest neighbor sample in class C.

This paper choose k = 20 and executes 40 times to get the average weight of each weight, then used 0.05 as the threshold value. Then, 9 redundant features were removed. They are H component, gray level heterogeneity, heterogeneity of gradient distribution in HSV color space; gray level heterogeneity, heterogeneity of gradient distribution, Inertia in RGB color space; gray level heterogeneity, heterogeneity of gradient distribution, inverse difference moment in HSI color space. Finally, the excellent feature set of dimension space [400,47] is obtained.

4. The Detailed Design of CSBP-CS
4.1 Cuckoo Search
The standard cuckoo search algorithm is a group intelligent search algorithm inspired by the behavior of brood parasitism and designed by Xin-She Yang and SuashDeb. Each nest represents a candidate solution, which is based on random walk mode of levy flight (Equation 4). The random walk mode searches for a new solution instead of the old solution, and finally searches for the optimal solution in the nest group.
4. In (4), $x_n$ is old nest, $x_0$ is current optimal solution, levy flight step is (5)

\[ s = \frac{u}{|v|^{1/\beta}} \]  

The variable is given by $\beta \in [0.3, 1.99]$, $u$ and $v$ subject to normal distribution:

\[ \{ u \sim N(0, \sigma_u^2) \}, \{ v \sim N(0, \sigma_v^2) \} \]

4.2 BP Neural Network

The BP neural network is a Multi Layer Perceptron (MLP) consisting of multiple neurons, which is trained by the BackPropagation algorithm. It consists of input layer, hidden layer and output layer. The neural network model constructed in this paper is shown in Figure 2:

![Figure 2 Model of BP](image)

In Fig. 2, in output layer L, for each training set output label value $y$, the error is $\delta^L = a^L - y$. Layer $l$ error is calculated by back propagation algorithm as follows:

\[ \delta^l = ((\delta^T) \delta^{l+1}) \ast a^l \ast (1 - a^l) \]  

At this time, the weight updating formula for each neuron is as follows:

\[ W_{ij}^l = W_{ij}^l + \lambda a_i^l \delta_i^{l+1} + \mu C_{ij} \]  

Where $W_{ij}$ denotes weights between $i, j$ of $l$ layer neurons. In order to speed up the learning efficiency, the correction matrix mechanism is introduced as follows:

\[ C_{ij} = a_i^l \delta_i^{l+1} \]  

4.3 The Design of CSBP-CS

Because of convergence speed of CS algorithm is slower when in later stage and easy to fall into local minimization, so this paper combined the back propagation ability of BP neural network with global search capability of CS to propose a cuckoo search and BP cooperative search algorithm (CSBP-CS). CSBP-CS can help CS to jump out of the local minimum when in the later stage of optimization and decrease the disadvantage of easily falling into the local minimum in CS.

In the collaborative network, this paper combined the thresholds and weights $\{w1\}{b2}\{w2\}{b3}$ (Fig. 4) of three-layer BP network into a nest vector $x^{bp}$ by formula (9), where $\{b2\}\{b3\}$ are the thresholds of hidden layer and output layer respectively,

\[ x_{bp} = [w_1^1, b_2^1, w_2^2, b_3^3] \]  

Then compared the best fitness value with $x^{cs}$ that is searched by CS algorithm and judged what was better nest according to formula (10):

\[ x^b = \begin{cases} x_{bp} & \text{if } f(x_{bp}) < f(x^{cs}) \\ x_{bc}^{cs} & \text{if } f(x_{bc}^{cs}) < f(x_{bp}) \end{cases} \]  

So the better solution vector $x^b$ was obtained. Then iterative updating of $x^b$ from formula (4) (7) and continuous optimization until the end condition, finally, the optimal solution vector in the CSBP-CS network is searched.
5. The Result and Analysis of Classification

5.1 Simulation Experiment
In order to analyze the Relief F's effectiveness in feature selection, the simulation experiment is implemented in MATLAB 2014. The training set is features of 260 images, The test set is 140 images that is collected by OpenMV. The optimal bird's nest x^6 obtained by cooperative search is used as the weight threshold of the CSBP-CS recognition model. Under the same conditions, Then the CSBP-CS network and CSBP-CS RF network are compared to analyze the effectiveness of Relief F algorithm. The number of iterations is all 50, the search nest of CSBP-CS is 30, but the search nest of CSBP-CS RF is only 25.

5.2 Result and Analysis
The two algorithm classification models were successively implemented for many consecutive independent experiments, this paper mainly from convergence time and tomato disease recognition accuracy rate to analysis the performance of RF. The model fitness value changes with the iteration times as shown in Figure 3. It can be seen that the fitness value of CSBP-CS is reduced to the lowest after 27 iterations, but CSBP-CS RF reduced to the lowest only after 19 iterations, which is faster then the speed of CSBP-CS, and the fitness value of CSBP-CS RF and CSBP-CS has little difference, so reducing feature dimension by Relief F has no effect on recognition efficiency, but accelerate convergence rate.

![Figure 3](image)

The training network is saved after training, and the test set [140, 56] is taken as input to calculate the recognition accuracy of the two networks. Perform continuous testing and save the test results for 5 consecutive times. The statistics of the CSBP-CS identification results in this paper are shown in Table 1. The recognition results of the CSBP-CS RF are shown in Table 2.

**Table 1** Recognition results of CSBP-CS network

| classifier | powdery mildew | grey mold | late blight | normal leaf | recognition accuracy | average time (sec) |
|------------|----------------|-----------|-------------|-------------|----------------------|-------------------|
| 32         | 31             | 31        | 34          | 91.1%       |
| 30         | 30             | 33        | 31          | 91.4%       |
| CSBP-CS    | 31             | 28        | 32          | 33          | 90.9%                |
| 30         | 28             | 31        | 32          | 91.1%       |
| CSBP-CS    | 31             | 28        | 32          | 31          | 90.1%                |

**Table 2** Recognition results of CSBP-CS RF

| classifier | powdery mildew | grey mold | late blight | normal leaf | recognition accuracy | average time (sec) |
|------------|----------------|-----------|-------------|-------------|----------------------|-------------------|
| 31         | 28             | 34        | 34          | 90.7%       |
| 30         | 32             | 33        | 35          | 92.0%       |
| CSBP-CS RF | 31             | 28        | 30          | 33          | 90.7%                |
| 30         | 29             | 30        | 34          | 88.8%       |

It can be seen from the Table.1 and Table.2 that the average accuracy of CSBP-CS is same as CSBP-CS RF, but the number of nest needed by CSBP-CS RF to research optimal is five fewer than CSBP-CS’s because the input of CSBP-CS is 56, but CS BP-CS RF is only 47, so the average iteration time of CSBP-CS RF is less than CSBP-CS. So we can know that reducing feature dimension by Relief F makes the network run faster and won’t reduce the recognition accuracy. In conclusion, RF algorithm can be used in this network to accelerate the speed of optimization without affecting the recognition accuracy, and RF can simplify the model.
6. Summary
In this paper, CSBP-CS RF was designed to be a disease classifier. 260 disease images were used for training set and OpenMV intelligent monitoring system was used to collect 140 disease images for testing set. The RGB-HSI conversion algorithm was selected to realize the segmentation of the disease spot, and 56 features such as color features and texture features were extracted from the disease spot, the relief F was used to select excellent features with large weights and the excellent feature set was used to construct the disease classification model. Than CSBP-CS RF network was proposed to be disease classifier. Simulation experiments showed that the CSBP-CS RF network performance was better than CSBP-CS network, the use of relief F is effective.

References
[1] ZHUANG Qiongyun. Design and Implementation of the Auto-Tracking Smart Car Based on OpenMV, J. Journal of Liming Vocational UniversityActa Agriculturae Zhejiangensis, 2018(4).
[2] MEI Yanpin, FU Rong. Research on the 3D Identification System Based on OpenMV, J. New Technology & New Process, 2018.
[3] Liu Wenhui, Wang Zhe. Design and implementation of green house video monitoring system based on cloud computing, J. Application of Electronic Technique, 2015(6):139-142.
[4] WANG Lin-ming, GAO Xiao-yang, LI Hong-ling,el. Barley diseases discrimination based on neural network, J. Journal of Gansu Agricultural University, 2015,50(02):173-176.
[5] Soni P, Chahar R. A segmentation improved robust PNN model for disease identification in different leaf images[C]// IEEE International Conference on Power Electronics. IEEE, 2017.
[6] Sun Chen, Li Yang, Li Xiaoge, Yu Jiaoyan. Stock forecasting model based on optimising BP neural network with cuckoo search, J.Computer Applications and Software,2016,33(02),276-279.
[7] MA Wei. Human Behavior Recognition Based on Cuckoo Search Algorithm Optimizing Features and Classifier Parameters[J].Microelectronics & Computer, 2016, 33(5):102-105.
[8] XIAO Da, WANG Runmin, ZOU Xiao,el. A Noninvasive Temperature Measurement Based on Gabor Transform and Gray Level Gradient Co-Occurrence Matrix Using Ultrasound, J. Chinese Journal of Sensors and Actuators, 2017, 30(11):1684-1688.
[9] Xiao-Qi L, Jin-Ge G, Yu-Hong Z. Research and realization of Tamura texture feature extraction method based on image segmentation, J. Chinese Journal of Tissue Engineering Research, 2012.